In situ/laboratory classification of phytoplankton types - database: efforts and goals

Lesley Clementson (CSIRO), Ray Barlow (BCRE) and Toru Hirawake (HU)
IOCS meeting, Darmstadt, 06-08 May 2013

CSIRO MARINE AND ATMOSPHERIC RESEARCH
www.csiro.au

International Working Group for PFT Algorithm Development
First Meeting  
- Sapporo, Japan, November 2011

Tasked with establishing a dataset to be used by scientists developing a PFT algorithm.

The rationale was that PFT algorithms could then be compared on a “level playing field” – tested on a common dataset.

It was acknowledged that a significant effort had already been made to construct various phytoplankton databases by individual institutions and international science programmes (notably by NASA), however the workshop participants believe added phytoplankton-specific parameters would only enhance the information available to researchers.

The aim was to have more than just pigment data.
Since the first meeting

The MARine Ecosystem DATa (MAREDAT) initiative

pigment data from 136 surveys around the globe (35,634 values)


But remember the aim of The International Working Group for PFT Algorithm Development was to have more than just pigment data
Why more than pigments

Two key uncertainties are highlighted, related to the use of 
(1) fucoxanthin (fucox) as a diagnostic pigment for diatoms and 
microphytoplankton (>20 µm)
Fucox is primarily associated with diatoms but can also be the dominant pigment 
in some dinoflagellates and prymnesiophytes. Furthermore, some 
fucox-containing small diatoms exist in the nano size range (<20 µm).

(2) 19’-hexanoyloxyfucoxanthin (hex) as a diagnostic pigment for the 
nano-eukaryotes.
Hex is recognised as a diagnostic pigment for prymnesiophytes, however, this 
group exists in both the nano and pico size ranges, the latter comprising a 
significant proportion of the pico-eukaryotes.
The wish list

Data types

_in situ_ Radiometry
HPLC
IOPs – absorption, backscatter
Fractionated chlorophyll
Microscopy
Flow cytometry

microscopic counts give a measure of the number of cells rather than the contribution of cells to the total chlorophyll concentration, and these measures are not equivalent. Using the locations where both sets of measurements exist to better understand the magnitude of the many methodological uncertainties and there is no established protocol for uncertainties taking these measurements. Stacked filters vs parallel; pressure vs vacuum filtering; fluorimetric vs HPLC
Bio-optical algorithm data sets
- SeaBASS (taken from Werdell, OOW, Glasgow, Scotland, UK, 07 October 2012)

Rrs

Rrs & Chl

Rrs & Chl & absorption

Rrs & Chl & absorption & backscattering
Challenges

NOMAD (Feb 2012)

\(a (a_p, a_g, a_d)\) (20 \(\lambda\)s); \(bb\) (fitted - 20 \(\lambda\)s); \(bbr\) (original - 14 \(\lambda\)s); fluorometric chl-a; HPLC pigments; \(Kd\) (20 \(\lambda\)s); \(Lw\) (20 \(\lambda\)s);

4459 samples – 1381 pigment results Not all samples have all parameters

AMT
Great for pigment data
Not so great for coincident data

Rrs & Chl & absorption & backscattering
<table>
<thead>
<tr>
<th>Cruise code</th>
<th>start date</th>
<th>end date</th>
<th>start location</th>
<th>end location</th>
<th>Chl-a (bottle)</th>
<th>Chl-a (uway)</th>
<th>Chl-a size fraction</th>
<th>HPLC pigments (bottle)</th>
<th>HPLC pigments (uway)</th>
<th>Phytoplankton ID</th>
<th>Flow cytometry</th>
<th>FlowCAM</th>
<th>Primary production</th>
<th>CDOM</th>
<th>apH</th>
<th>Particle multiplier</th>
<th>SPM conc</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMT5</td>
<td>14-September-1997</td>
<td>17-October-1997</td>
<td>UK</td>
<td>Falklands</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT6B</td>
<td>04-April-1998</td>
<td>04-May-1998</td>
<td>Falklands</td>
<td>UK</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT6B</td>
<td>14-May-1998</td>
<td>15-June-1998</td>
<td>Sth Africa</td>
<td>UK</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT7</td>
<td>14-September-1998</td>
<td>25-October-1998</td>
<td>UK</td>
<td>Falklands</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT8</td>
<td>25-April-1999</td>
<td>07-June-1999</td>
<td>Falklands</td>
<td>UK</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT9</td>
<td>15-September-1999</td>
<td>13-October-1999</td>
<td>UK</td>
<td>Montevideo</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT10</td>
<td>12-April-2000</td>
<td>08-May-2000</td>
<td>Montevideo</td>
<td>UK</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT11</td>
<td>12-September-2000</td>
<td>11-October-2000</td>
<td>UK</td>
<td>Montevideo</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT12</td>
<td>12-May-2003</td>
<td>17-June-2003</td>
<td>Port Stanley</td>
<td>UK</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT13</td>
<td>10-September-2003</td>
<td>14-October-2003</td>
<td>UK</td>
<td>Port Stanley</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT14</td>
<td>28-April-2004</td>
<td>01-June-2004</td>
<td>Port Stanley</td>
<td>UK</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT15</td>
<td>17-September-2004</td>
<td>29-October-2004</td>
<td>UK</td>
<td>Cape Town</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT16</td>
<td>20-May-2005</td>
<td>29-June-2005</td>
<td>Cape Town</td>
<td>UK</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT17</td>
<td>15-October-2005</td>
<td>28-November-2005</td>
<td>Sth Africa</td>
<td>UK</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td>Y</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT18</td>
<td>10-October-2008</td>
<td>10-November-2008</td>
<td>UK</td>
<td>Falklands</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT19</td>
<td>13-October-2009</td>
<td>28-November-2009</td>
<td>UK</td>
<td>Chile</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT20</td>
<td>12-October-2010</td>
<td>25-November-2010</td>
<td>UK</td>
<td>Chile</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td>X</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMT21</td>
<td>29-September-2011</td>
<td>16-November-2011</td>
<td>UK</td>
<td>Chile</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td>Y</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The collection of coincident data is not an easy task as the different types of data are often collected on the one voyage by different scientists, not necessarily from the same institute.

this is certainly part of the issue with the AMT data

HPLC is the most commonly used data source in the parameterisation of algorithms.

- relatively large number of data points available in all ocean environments

However there uncertainties involved in the HPLC data

- not all HPLC data the same
- Sea-HARRE reports (NASA)
Challenges continued

Also with the pigment data, care must be taken to avoid validating against parameterisation data.

Should we re-parameterise the HPLC-based algorithms with a common HPLC dataset to ensure this does not overlap with the validation dataset?

**The real challenge is to pull together coincident data**
An interrogative database of bio-optical parameters for Australian waters is being established by the AEsOP project, funded by the EOI-TCP. This database will provide in situ data for the development of robust regional algorithms, the validation of PFT algorithms and, the enhancement of standard global algorithms in the future.
The dataset will include multiple coincident parameters such as HPLC pigments (including size fractionated pigments where available), pigment concentration and composition, absorption coefficients ($a_{ph}$, $a_d$, $a_g$), total suspended mater (TSM), Secchi depth and processed data from radiometers, hydrosclat, ac-9, ac-s and other instruments from 1997 to the present day. It is envisaged that the first version of this database will available by mid 2013.

Should we attempt to use this database design for a global database/set

How will this be maintained?
Thank you

CSIRO Marine and Atmospheric Research
Lesley Clementson

t  +61 3 6232 5337
e  lesley.clementson@csiro.au
w  www.csiro.au