CCI Toolbox
Carsten Brockmann
Brockmann Consult

with contributions from Land Cover CCI, Ocean Colour CCI, SST CCI and SEWG
CCI Datasets

Facilitators

Transformation & Exploitation Toolbox

Data Access

Visualisation

User Communities
CCI modeller
CCI EO teams
General ECV data users
Directions

Data to users

Data Exploration

Software to data
Tools to Ease Data Usage

- **Converting format**
- **Transforming and supplementing information**
  - Unit conversion, scaling
  - Derived products
  - Temporal and spatial binning
  - Mapping
  - Common preprocessing: flagging, cloud screening, ...
- **Combining information from different sources**
  - Different products of a single CCI
  - Multiple CCIs
  - CCI and other sources
- **Uncertainties exploitation and propagation**
• **User Requirements**
  - Users need Land Cover product(s) in
    - Different spatial resolutions
    - Different projections
    - Spatial subsets
  - Modellers need Plant Functional Types (PFT) rather than Land Cover classes
  - Aggregation by biome

• **LC products are generated as**
  - Global product(s)
  - Lat/Lon projection

=> need for a user tool to convert LC standard product into individual user products
Rules for Aggregation of LC map product:

1. Fractional area of each LC class
2. Ranking of LC class by fractional area in target cell; first n entries of sorted list are written to n bands (n is user parameter, called majority classes)
3. Fractional area of each PFT

Rules for Aggregation condition product:

1. If processed_flag == processed AND current_pixel_state == clear_land OR clear_water then process_pixel()
2. accuracy = median (algorithmic_confidence_level)

<table>
<thead>
<tr>
<th>Class</th>
<th>Area</th>
<th>Majority class</th>
</tr>
</thead>
<tbody>
<tr>
<td>class a</td>
<td>5/35</td>
<td>3</td>
</tr>
<tr>
<td>class b</td>
<td>0/35</td>
<td>4</td>
</tr>
<tr>
<td>class c</td>
<td>19/35</td>
<td>1</td>
</tr>
<tr>
<td>class d</td>
<td>11/35</td>
<td>2</td>
</tr>
</tbody>
</table>
## System Requirements

### PFT conversion

<table>
<thead>
<tr>
<th>ID</th>
<th>LC CCI class description</th>
<th>Tree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Broadleaf</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Evergreen</td>
</tr>
<tr>
<td>0</td>
<td>No data</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Cropland, rainfed</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Herbaceous cover</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Tree or shrub cover</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Cropland, irrigated or post-flooding</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Mosaic cropland (&gt;50%) / natural vegetation (tree, shrub, herbaceous cover) (&lt;50%)</td>
<td>5</td>
</tr>
<tr>
<td>40</td>
<td>Mosaic natural vegetation (tree, shrub, herbaceous cover) (&gt;50%) / cropland (&lt;50%)</td>
<td>5</td>
</tr>
<tr>
<td>50</td>
<td>Tree cover, broadleaved, evergreen, closed to open (&gt;15%)</td>
<td>90</td>
</tr>
<tr>
<td>60</td>
<td>Tree cover, broadleaved, deciduous, closed to open (&gt;15%)</td>
<td>70</td>
</tr>
<tr>
<td>61</td>
<td>Tree cover, broadleaved, deciduous, closed (&gt;40%)</td>
<td>70</td>
</tr>
<tr>
<td>62</td>
<td>Tree cover, broadleaved, deciduous, open (15-40%)</td>
<td>30</td>
</tr>
</tbody>
</table>
CCI LC User Tool
Implementation

- BEAM operator
- Command line interface
- Graphical user interface
Input: Original CCI Land Cover state map

CCI land cover state map for the epoch 2003-2007
CCI land cover state map for the epoch 2003-2007
• algorithmic confidence level
CCI land cover state map for the epoch 2003-2007

- Subset: Western Europe and Mediterranean Basis
CCI land cover state map for the epoch 2003-2007
- majority class 1
- Gauss-Grid N320
Output: Corresponding Confidence map @ N320

CCI land cover state map for the epoch 2003-2007
- Gauss-Grid N320
Output: Areal coverage of grassland class @ 10km

CCI land cover state map for the epoch 2003-2007
• aggregated ~9.8km/ pixel
• area of CCI LC class – 130 – grassland
CCI land cover PFT for the epoch 2003-2007
- aggregated ~9.8km/ pixel
- area of PFT - tree broadleaf evergreen
Example SST CCI

- **Regional averaging**
  - Purpose: spatio-temporal averaging of several years of daily SST CCI data products (L3 or L4, resolution 0.05°) for user specifiable regions and temporal resolutions
  - Result: time series ‘plot‘ for each region specified, with temporal resolution specified

- **Re-gridding**
  - Purpose: re-gridding of daily SST CCI data products (L3 or L4, resolution 0.05°) to lower spatial and temporal resolutions
  - Result: series of re-gridded data products
Tools for exploitation and propagation of uncertainties

• Different types of uncertainties in source data
  • Large scale correlated uncertainty
  • Synoptically correlated uncertainty
  • Uncorrelated uncertainty
  • Adjustment uncertainty

• Coverage uncertainty (due to incomplete spatio-temporal coverage in source data) is introduced

• Different types of uncertainties in CCI ECVs
  • Defined by each CCI individually
  • Comparability, interoperability need to addressed in Phase 2 and supported by proper tools

• Regional averaging and re-gridding propagate each type of uncertainty separately
Tools to Support Validation

- **User Requirements**
  - **Within EO team:**
    - comparison of EO data with reference data
    - Satellite intercomparison
  - **Modellers**
    - comparison with model output
    - Systematic comparison of different configurations
  - **All**
    - Extraction of point data
    - Filtering of data

- **Match-up generation and analysis**
  - Several tools exist: SST-CCI MMDB, OC-CCI, Felyx, MERMAID ...

- **Time series generation and analysis**
Example SST CCI

Multi-sensor Match-up Database
Using MMDB

AVHRRMTA_G SST \text{skin} \text{ versus drifter: SST\_CCI v1 (dt3)}

Day time

Night time

Carsten Brockmann
courtesy G. Corlett
ECV interoperability

Cloud CCI
Cloud Fractional Coverage
July 2008

OC CCI
Cloud Flag Occurrence
July 2008
Offline Toolbox Concept

- Sentinel 1-2-3 Toolboxes
  - Continuation of successful ENVISAT toolboxes
  - Common software platform starting from BEAM
  - Developer forum & community approach
  - Tools for data visualisation, analysis and processing
  - Development platform
  - Python support and other widely used languages

- Sentinel Atmospheric Missions Toolbox

- Model for a CCI Toolbox
  - CCI ECVs take the role of sensor products
  - Partly similar requirements
  - Partly quite different requirements
• From SST-CCI:
  • Re-gridding the complete SST CCI Climate Data Set is a considerable task. It is not something a user really wants to conduct
  • For Phase II tools are planned to be available as ‘online facilities’ integrated into the data retrieval interface

• If data are central (Sentinel, CEMS, Collab GS)
  • Associated tools can be offered
  • Develop once, use multiple
  • Faster response by caching
Example Calvalus

- Hadoop based system for concurrent processing of full mission datasets
- Workflows for
  - L2 and L3 production
  - Match-up analysis
  - Time series generation
- User can
  - Upload own processing code (versioned)
  - Upload own reference data
Calvalus portal

- input set selection
- processor versions
- processing parameters
- in-situ data for matchup analysis
- variables for aggregation
- trend analysis
### Level 3 Parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expression</th>
<th>Aggregator</th>
<th>Weight</th>
<th>Fill</th>
</tr>
</thead>
<tbody>
<tr>
<td>conc_tsm</td>
<td></td>
<td>AVG</td>
<td>1</td>
<td>NaN</td>
</tr>
<tr>
<td>conc_chl</td>
<td></td>
<td>AVG</td>
<td>1</td>
<td>NaN</td>
</tr>
</tbody>
</table>

**Good-pixel expression:** !2w_flags!INVALID

- **Stepping period:** 30 days
- **Spatial resolution:** 9.28 km/pixel
- **Compositing period:** 30 days
- **Supersampling:** 1 pixels
- **Number of periods:** 1
- **Target width:** 399 pixels
- **Target height:** 302 pixels
Processor integration

- Adapter for Unix executables (C++, Fortran, Python, ...)
- Adapter for BEAM GPF operators
- Concurrent processor versions in the system
- Automated deployment of processor bundles at runtime
CCI Toolbox Requirements - Summary

• **Toolbox users**
  - CCI EO teams
  - CCI modelers
  - General EO and ECV data user community

• **Ease data exploitation by scientists ("data to users")**
  - Converting format
  - Transforming and supplementing information
  - Combining information from different sources,
  - Facilitate across CCI data exploitation
  - Uncertainties exploitation and propagation
  - Data extraction, match-up and time series generation
  - Visualisation for exploration
  - Support Data Analysis and Validation

• **Bringing the scientific code to the CCI data**
  - Upload and processing environment (language independent)
  - Application Programming Interface (API)

• **Online and offline processing tools**