IDEAS + Task3
Sentinel 2 / Landsat Fusion Technics
06/04/2017

Béatrice Berthelot, Germain Salgues
Project context

• 2 main objectives:
  – Consolidation of S2 time series for thematic analysis and support land science
  – Densification of S2 time series to fill gaps and densify
  – The processing is performed at BOA level

• Thanks to S2B, density will be improved but even if revisit is high, there still are long period of time without any data
What is consolidation?

- A method to fill gap inside a time series
Illustration of the Methodology
How?

• Using innovative technics of distance based methods (vs interpolation) to fill the gaps:
  – Step of consolidation

• Combining multi spectral information close to reference sensors (MSI) to densify the time series:
  – Step of fusion

Application to landsat 8

The works was splitted in 3 steps:

Mono sensor Time series consolidation
Multi sensor Time series fusion/densification
Multi sensor Time series evolution
Since previous meeting

Previous status
- Consolidation approach was tested using L8 as reference sensor, and SPOT5 foreseen for the densification.

Updated status:
- Thanks to the ESA/SciHub, S2 data collection started in Fall 2016 to have more than 1 year of acquisitions (update on-going)
- Consolidation approach reviewed and tested onto S2 datasets
- Fusion approach implemented and applied onto L8 datasets
Presentation overview

1. Consolidation method description

2. Dataset description

3. Results

4. Upcoming activity
Phase 1

Mono sensor Time series consolidation
Mono sensor Time series consolidation

The objective of this step is the production of consolidated time series using DTW based method

Consolidation refers to the recovery of pixel values which are occluded at a given date in the time series.

The proposed approach has been evaluated against Interpolation based methods with:

- Linear interpolation/extrapolation
- Smoothing filtering (Savitsky-Golay)
Mono sensor Time series consolidation - DTW

- Minimization
- Alignment
- Distance Computation

N – DTW Metrics

Descriptors

Rebuild Time Series

Selected Time Series

Aggregation With TS to be rebuilt

Consolidated Temporal Block

Collection of Neighborhood temporal pixel

Collection of temporal pixel to be recovered

Time Series to be recovered

N - Candidates Time Series

Time series + Mask
Principle of DTW (Dynamic Time Wrapping)

- The Dynamic Time warping is a distance measurement between two time series.

For each pixels \((x_0, y_0, t)\)
- DTW is able to find a re-align pixel match within time series
- This property is then used for gap filling purpose
Cost function

• Used usefull bands (BOA):

\[ \text{dist}_k = \text{dist}(i, i') = \sqrt{\sum_{j=1}^{m} (C_{i,j} - Q_{i',j})^2} \]

• Path cost:

\[ DTW = \sum_{k=1}^{l} \text{dist}_k \]

• Minimisation:

\[ DTW(TPx(x_0, y_0), TPx(x_k, y_l)) = \min_{(x_k, y_l) \in N} DTW(TPx(x_0, y_0), TPx(x_k, y_l)) \]
The evaluation of this processing was evaluated on Sentinel 2 time series according the following approach:

- Evaluation of the consolidation apply directly on the time series by selecting a subset of cloud-free pixels
- Part of them are manually masked/degraded
- Consolidation method is applied
- The re-build pixels are compared to their actual value
Consolidation metrics

- Deviation from ground truth estimation

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy (A) / Mean Error</td>
<td>[ A = \frac{1}{N} \times \sum_{i=1}^{N} \varepsilon_i ]</td>
</tr>
<tr>
<td>Precision (P) / Standard deviation (STD)</td>
<td>[ P = \sqrt{\frac{1}{N-1} \times \sum_{i=1}^{N} (\varepsilon_i - A)^2} ]</td>
</tr>
<tr>
<td>Uncertainty (U) / Root Mean Square Error (RMSE)</td>
<td>[ U = \sqrt{\frac{1}{N} \times \sum_{i=1}^{N} \varepsilon_i^2} ]</td>
</tr>
</tbody>
</table>
• 4 sites
• 2 new added since previous meeting to start a collaboration with IDEAS task 3
## Site cloud coverage characterisation

<table>
<thead>
<tr>
<th>Site Name</th>
<th>% acquisitions where CC &lt; 20%</th>
<th>% acquisitions where CC &lt; 50%</th>
<th>% acquisitions where CC &lt; 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toulouse</td>
<td>55.00%</td>
<td>62.50%</td>
<td>71.25%</td>
</tr>
<tr>
<td>La Crau</td>
<td>54.24%</td>
<td>69.49%</td>
<td>86.44%</td>
</tr>
<tr>
<td>Barrax</td>
<td>61.40%</td>
<td>75.44%</td>
<td>89.47%</td>
</tr>
<tr>
<td>Maricopa</td>
<td>75.00%</td>
<td>83.33%</td>
<td>88.89%</td>
</tr>
<tr>
<td>Garderen</td>
<td>22.64%</td>
<td>43.40%</td>
<td>62.26%</td>
</tr>
<tr>
<td>Ottawa</td>
<td>34.00%</td>
<td>50.00%</td>
<td>70.00%</td>
</tr>
</tbody>
</table>
Toulouse – ROI locations and associated cloud coverage
Barrax – ROI locations and associated cloud coverage
Maricopa – ROI location and associated cloud coverage
La Crau – ROI location and associated cloud coverage
Ottawa – ROI location and associated cloud coverage
Garderen – ROI location and associated cloud coverage
Data pre-processing

Atmospheric corrections:
• L1C → L2A : 1st version using sen2cor v2.2.1
• Update of the full dataset using sen2cor 2.3 with DEM correction application

Data ingestion:
• ROIs definition and extraction (reflectances and quality indicator)
• NDVI computation
• Support change of granule format distribution since January 2017 (affect the on-going dataset extension collection)
Results overview

• Detailed results are shown for one site
  – Profile analysis
  – Statistical results

• Statistics summary over the 6 test sites

• Sensitivity analysis
  – Inputs bands
  – Tests sites
Profiles analysis: Toulouse test site
Profiles analysis: Toulouse test site
Profiles analysis: Toulouse test site

Consolidated profile for band 1

Consolidated profile for band 2

Consolidated profile for band 3

Consolidated profile for band 5

Consolidated profile for band 6

Consolidated profile for band 7
The DTW allows to retrieve pattern that may not be predicted accurately from previous or next acquisition, and in a much consistent way.
Outcomes

• Temporal density of the available data. *If the density is too low the neighborhood requested for linear and SG methods might be not reliable.*

• Frequencies contained in the signal for each band. *The NDVI example illustrates that for low frequencies signal SG approach could be reliable.*

• Method parameters when applicable. *For example, a better parameterization of the SG method would have provided better result than linear in almost any cases.*
Statistical results analysis: Toulouse test site

**Linear**

**DTW**

**SG**

DTW provides lower residues and is more consistent over the different dates.
### Statistical results: Toulouse test site

**TOULOUSE ROI1**

<table>
<thead>
<tr>
<th>Method</th>
<th>Input bands</th>
<th>Params</th>
<th>Nb dates</th>
<th>Mean Accuracy</th>
<th>Mean Precision</th>
<th>Mean Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>14</td>
<td>-0.0031</td>
<td>0.0339</td>
<td>0.0399</td>
</tr>
<tr>
<td>SG</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>14</td>
<td>0.0500</td>
<td>0.0417</td>
<td>0.0862</td>
</tr>
<tr>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>14</td>
<td>-0.0007</td>
<td>0.0169</td>
<td>0.0172</td>
</tr>
</tbody>
</table>

**TOULOUSE ROI2**

<table>
<thead>
<tr>
<th>Method</th>
<th>Input bands</th>
<th>Params</th>
<th>Nb dates</th>
<th>Mean Accuracy</th>
<th>Mean Precision</th>
<th>Mean Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>15</td>
<td>-0.0015</td>
<td>0.0367</td>
<td>0.0408</td>
</tr>
<tr>
<td>SG</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>15</td>
<td>0.0503</td>
<td>0.0439</td>
<td>0.0869</td>
</tr>
<tr>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>15</td>
<td>0.0001</td>
<td>0.0224</td>
<td>0.0227</td>
</tr>
</tbody>
</table>
## Statistical results

### GARDEREN

<table>
<thead>
<tr>
<th>Method</th>
<th>Input bands</th>
<th>Params</th>
<th>Nb dates</th>
<th>Mean Accuracy</th>
<th>Mean Precision</th>
<th>Mean Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>19</td>
<td>0.0085</td>
<td>0.0664</td>
<td>0.0762</td>
</tr>
<tr>
<td>SG</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>19</td>
<td>0.0423</td>
<td>0.0685</td>
<td>0.0977</td>
</tr>
<tr>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>19</td>
<td>0.0097</td>
<td>0.0430</td>
<td>0.0463</td>
</tr>
</tbody>
</table>

### OTTAWA

<table>
<thead>
<tr>
<th>Method</th>
<th>Input bands</th>
<th>Params</th>
<th>Nb dates</th>
<th>Mean Accuracy</th>
<th>Mean Precision</th>
<th>Mean Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>18</td>
<td>-0.0017</td>
<td>0.0349</td>
<td>0.0396</td>
</tr>
<tr>
<td>SG</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>18</td>
<td>0.0220</td>
<td>0.0327</td>
<td>0.0519</td>
</tr>
<tr>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>18</td>
<td>-0.0007</td>
<td>0.0142</td>
<td>0.0143</td>
</tr>
</tbody>
</table>
### Statistical results: La Crau test site

<table>
<thead>
<tr>
<th>Method</th>
<th>Input bands</th>
<th>Params</th>
<th>Nb dates</th>
<th>Mean Accuracy</th>
<th>Mean Precision</th>
<th>Mean Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>31</td>
<td>0.0004</td>
<td>0.0143</td>
<td>0.0241</td>
</tr>
<tr>
<td>SG</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>31</td>
<td>0.0259</td>
<td>0.018634068</td>
<td>0.0509</td>
</tr>
<tr>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>31</td>
<td>-9.5E-06</td>
<td>0.0213</td>
<td>0.0237</td>
</tr>
</tbody>
</table>
Statistical results: Maricopa test site

**Linear**

**DTW**

**SG**

DTW doesn’t provide the best residues but stays consistent over the different dates.
## Sensitivity analysis

### Input bands

<table>
<thead>
<tr>
<th>Method</th>
<th>Input bands</th>
<th>Params</th>
<th>Nb dates</th>
<th>Mean Accuracy</th>
<th>Mean Precision</th>
<th>Mean Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>14</td>
<td>-0.0007</td>
<td>0.0169</td>
<td>0.0172</td>
</tr>
<tr>
<td>DTW</td>
<td>'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>14</td>
<td>0.1313</td>
<td>0.1931</td>
<td>0.2532</td>
</tr>
</tbody>
</table>

### Site sensitivity

<table>
<thead>
<tr>
<th>Site</th>
<th>Method</th>
<th>Input bands</th>
<th>Params</th>
<th>Nb dates</th>
<th>Mean Accuracy</th>
<th>Mean Precision</th>
<th>Mean Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toulouse ROI1</td>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>14</td>
<td>-0.0007</td>
<td>0.0169</td>
<td>0.0172</td>
</tr>
<tr>
<td>Toulouse ROI2</td>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>15</td>
<td>0.0001</td>
<td>0.0224</td>
<td>0.0227</td>
</tr>
<tr>
<td>Garderen</td>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>19</td>
<td>0.0097</td>
<td>0.0430</td>
<td>0.0463</td>
</tr>
<tr>
<td>Ottawa</td>
<td>DTW</td>
<td>1,2,3,4,5,6,7,8,11,12, 'NDVI'</td>
<td>Cloud &lt; 80%</td>
<td>18</td>
<td>-0.0007</td>
<td>0.0142</td>
<td>0.0143</td>
</tr>
</tbody>
</table>
Conclusion on consolidation (1/2)

• The profiles analysis and statistical results over the tests sites confirm that the DTW provide the overall best results.
• The RMS is about twice smaller for the DTW than for any other methods.
• The profile analysis, underline how the DTW approach can provide better results than the linear and sg methods.
• And the DTW asset consists in using actual measurement of the data, because it allows retrieving the specificity of the signal to consolidate that cannot be estimated accurately, by relying only on other acquisitions and area.
Conclusion (2/2)

For an optimal result the DTW method requires:

• Accurate BQA layers; this is the purpose of the assimilation stage.

• A good time sampling of the time series; this is the purpose of the densification module. The Sentinel 2 data provide a 10 days sampling (this will be extended to 5 days after S2B activation) and even more by fusing LS08 data onto S2 time series. And regarding the occurrence of the cloud coverage over the different test sites, it is a crucial aspect of the time series analysis.

• Use the information of as many as possible time serie’s bands
Phase 2

Multi sensor Time series fusion/densification
Requirements for data merging

• With S2A as the reference sensor
  – Need for sensors which spectral responses are closed to S2/MSI
  – Spatial resolution compatible with S2/MSI
  – Which could be cross calibrated with S2/MSI

• Recording at high spatial and high-temporal resolution
  – Surface reflectance needs to be available (or computable)
    ➞Compatible atmospheric correction & cloud/shadow detection
    ➞Adjustments for BRDF (solar, view angle), band pass differences
    ➞Adjustment to common frame, resolution (~20m), regrid
Multi sensor Time series fusion/densification (1)

The two aspects of the products fusion are dealt with separately:

- Radiometric adjustment
- Geometric adjustment

This processing chain is compositing with four different steps:

- First, the pan-sharpening of the Landsat 8 products, it is required to significantly improve the Landsat 8 spectral band resolution.
- Secondly, the spectral adjustment using an approach relying on the sensor relative spectral response to project the Landsat 8 OLI acquired values onto the Sentinel 2 MSI sensor.
- Third, BRDF adjustments
- Finally, the resulting products should geometrically adjust to match the Sentinel 2 time series foot print.
Multi sensor Time series fusion/densification (2)

Stage 1

- MSI Level 2A +Masks +BRDF Coefficient
- Consolidation

Stage 2

- OLI Level 2A +Masks +BRDF Coefficient
- Pan-sharpening
- Band pass adjustment to MSI (SBAF)
- BRDF adjustments
- Re-gridding co registration
- Densification

Data Collection

User Defined ROI

MSI/OLI Simulated Surface Reflectance products
Geometric adjustment

Geometric adjustment required:
• Landsat XS (30m) and PAN (15m)
• Sentinel 2, XS (10m, 20m, 60m)

Solution:
• Working at the Landsat 8 XS 30m resolution:
  – Subsampling the S2 bands at 30m (up-sampling for 60m bands)
• Working at the Sentinel 2 20m resolution:
  – Up-sampling the bands to 20m
  – Using a Pan Sharpening method to increasing the Landsat 8 products resolution to 20 m by relying on actual information included in the panchromatic band

Pansharpening:

\[ XS_{HR} = Pan \odot \frac{\uparrow (XS_{LR})}{(Pan \otimes PSF)} \]
Spatial resampling
The bands used for L8/S2 fusion are selected among the similar bands:
## Multi sensor Time series fusion/densification - SBAF

<table>
<thead>
<tr>
<th></th>
<th>Costal Aerosol</th>
<th>Blue</th>
<th>Green</th>
<th>Red</th>
<th>NIR</th>
<th>SWIR 1</th>
<th>SWIR 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLI</td>
<td>Band 1</td>
<td>Band 2</td>
<td>Band 3</td>
<td>Band 4</td>
<td>Band 5</td>
<td>Band 6</td>
<td>Band 7</td>
</tr>
<tr>
<td>MSI</td>
<td>Band 1</td>
<td>Band 2</td>
<td>Band 3</td>
<td>Band 4</td>
<td>Band 8a</td>
<td>Band 11</td>
<td>Band 12</td>
</tr>
</tbody>
</table>

\[
SBAF = \frac{\overline{\rho_{\lambda(Oli)}}}{\overline{\rho_{\lambda(MSI)}}} = \frac{\left( \int \rho_{\lambda} \ast RSR_{\lambda(Oli)} \, d\lambda \right)}{\left( \int RSR_{\lambda(Oli)} \, d\lambda \right)} \frac{\left( \int \rho_{\lambda} \ast RSR_{\lambda(MSI)} \, d\lambda \right)}{\left( \int RSR_{\lambda(MSI)} \, d\lambda \right)}
\]

\( \rho_{\lambda} \) is at this step estimated from L8 values using all bands.
Will be taken from the DB in step 3  (or MERIS or hyperspectral if existant e.g. Hyperion,)

\[
\overline{\rho_{\lambda(MSI)}}^* = \frac{\overline{\rho_{\lambda(Oli)}}}{SBAF}
\]
Multi sensor Time series fusion/densification - SBAF

Reprojected L8 data on S2 spectral and spatial resolution
BRDF adjustments

• Landsat-8 and Sentinel-2 will have distinct orbit and sun/view geometry.

• Sun/view geometry:
  – Sentinel-2: VZA = +/- 12 deg, Aq. Time ~ 10:30 a.m
  – Landsat-8 : VZA = +/- 7 deg, Aq. Time ~ 10:00 a.m

• On-going activities: BRDF correction study
  – MODIS BRDF
  – PARASOL BRDF
Geometry acquisition for S2 and L8/OLI

- Low $\theta_v$
- High variability of $\theta_s$

**L8 Sun**

**L8 View**

**S2 Sun**

**S2 view**
Multi sensor Time series evolution

Roadmap

- Apply BRDF correction to improve quality
- Evaluation of the overall processing chain
- Create a spectrum database for DTW matching