Algorithm selection for the operational production of crop maps in the frame of the Sentinel-2 Agriculture project

J. Inglada, B. Tardy, S. Valero, M. Arias, D.Morin, O. Hagolle, G. Dedieu, S. Bontemps, G. Sepulcre, P. Defourny

The Sentinel-2 Agriculture Project

The **Sentinel-2 Agriculture** project, aims at showing on a **large scale** project, the capabilities of Sentinel-2 mission for **agriculture monitoring**, by providing an **open source processing software** to generate, among other products, **crop type maps**.

Project Consortium

- Université Catholique de Louvain
- CESBIO
- C-S
- C-S România

Funded by ESA

http://www.esa-sen2agri.org/

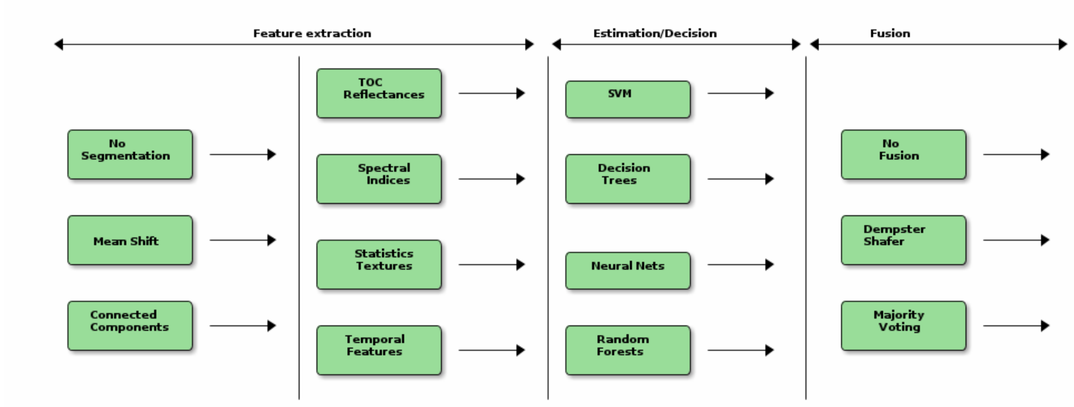
The Crop Type Product

- This product will consist of a map of the main crop types or crop groups in the given region.
- The main crop types are defined as those covering a minimum area of 10% of the annual cropland and for which the cumulated area reaches more than 75% of the annual cropland in the region.
- A maximum of 5 crop types will be considered per site.
- The 4 key crops in the GEO Global Agricultural Monitoring (GEOGLAM) initiative and the Agricultural Market Information System (AMIS) will be prioritized whenever possible: wheat, maize, rice and soybean.
- The distinction between rainfed and irrigated crops will also be included as an additional attribute.
- The delivery time of the first product is set-up to 2 weeks after the first half of the season.
- The last (and most accurate) crop type map will be delivered 2 weeks after the end of the season.
- The crop types maps will be provided on a regular grid at 10m resolution.

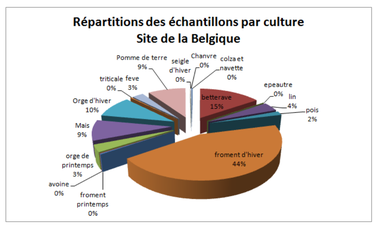
Algorithm exploration

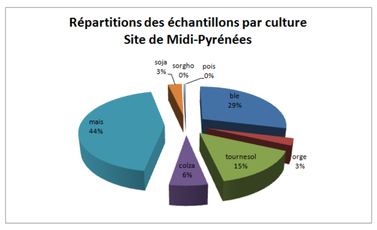
Goal: Select 5 algorithms prior to a benchmark over 12 sites.

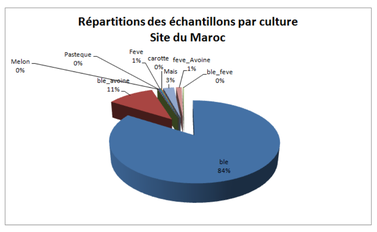
Algorithm choices:



3 different sites among 12: climatic and crop type variability



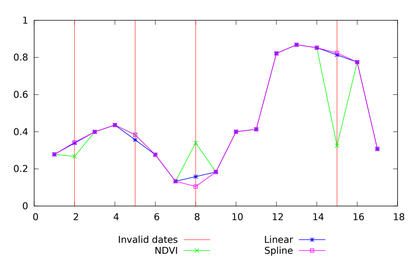


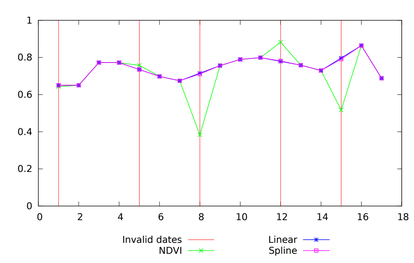


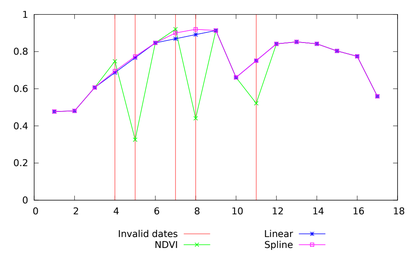
SPOT4(Take5) time series completed with Landsat-8 images.
https://www.ptsc.fr/fr/produits/spot4-take5

Time series pre-processing:

- Comparisons between raw L2 data, raw L2 data plus masks and gapfilled.
- Gapfilled time series allow **regular temporal resampling for interannual** – and between adjacent orbits – **supervised learning**.
- Linear and cubic spline interpolation have been compared.







Classification performances

κ	Linear	No gapfill
Linear	0.9360	0.6855
No gapfill	0.7195	0.9483

κ	Linear	Spline
Linear	0.9360	0.9274
Spline	0.9289	0.9356

Gapfilling is needed, but linear is enough.

Algorithm comparison setup

Goal: explore a large number of combinations between features, classifiers and their parameters before selecting 5 algorithms for a thorough benchmark.

Classifiers:

Statistical

- Kernel Methods
  - Linear SVM
  - RBF SVM
- Neural Networks
  - Mulit-layer perceptron

Trees

- Decision Trees
- Gradient Boosted Trees (GBT)
- Random Forests (RF)

Features:

- Surface Reflectances
- Tasseled Cap Transformation
- NDVI-like indices for different band combinations

$$\frac{B_i - B_j}{B_i + B_j}$$

- Up to 26 features (time series)
- Feature selection approaches implemented

Metrics for the comparison:

- Kappa Index, Overall Accuracy, FScore, Computation Time

General conclusions:

- RF and GBT have similar performances and better than classical Decision Trees.
- RBF SVM is better than Linear SVM and close to RF, but much slower
- Neural Networks have bad performances and their architecture is difficult to tune.

Selected 5 Algorithms for Benchmarking

Input data: linearly gapfilled L2 series. TOCRefl, NDVI, NDWI, Brightness.

Algorithms:

1. Random Forest classifier
2. RBF-SVM classifier
3. Best classifier with Mean-shift filtering
4. Best classifier with temporal regular resampling
5. Dempster-Shafer fusion of the previous approaches

Next Steps

- Benchmark the 5 selected algorithms on 12 sites spread over Africa, Asia, Europe and the Americas.
- Define the system specifications for the operational processing chains using the results of the benchmark.
- A similar approach is also applied for the other target products of the Sentinel-2 Agriculture project: cloud-free composites, binary crop mask and vegetation status indicators.