Investigating operational country-level crop monitoring with Sentinel 1 and 2 imagery

Nicolas David, Sébastien Giordano, Clément Mallet

ACCEPTED TO REMOTE SENSING LETTERS

Investigating operational country-level crop monitoring with Sentinel 1 and 2 imagery

Nicolas David, Sébastien Giordano and Clément Mallet. Email: clement.mallet@ign.fr

Univ. Gustave Eiffel, IGN-ENSG, LASTIG, Saint-Mandé, France

ARTICLE HISTORY
Compiled August 2, 2021

ABSTRACT
In this paper, we propose an operational solution for the yearly classification of crop parcels at national scale (namely France) for Land Parcel Identification System updating, under the Common Agricultural Policy (CAP) umbrella. Our pipeline is based on the ι^2 open source framework and fed with both time series of Sentinel-1 radar and Sentinel-2 optical images, with complementary contributions. Three conceivable scenarios are investigated with two sets of nomenclatures (17 and 43 classes): early, on-line, and late classifications. Experiments performed on 2017 show very satisfactory results (82-97%), locally almost on-par with state-of-the-art deep based methods. We can conclude our framework offers a strong basis for country-scale operational deployment for 2020+ CAP.

KEYWORDS
Sentinel, optical, SAR, time series, crops, classification, operational system, country scale, LPIS, Common Agricultural Policy.

1. Introduction

With almost 30% of the European budget, the Common Agricultural Policy (CAP) has been the prominent European Commission policy for the last 30 years. It supports farmers, ensuring food quality with secured prices. In return, it requires farmers to declare their agricultural land cover/use every year (Land Parcel Identification System, LPIS). The Paying Agency of every member state must supervise at least 5% of the declaration, traditionally performed by field inspection and photo interpretation of Very High Resolution geospatial images. The new regulation for 2020+ CAP payments allows the possibility of integrating Sentinel data for monitoring and cross-compliance assessment (Campos-Taberner et al., 2019). This change of paradigm is in line with the current findings of the literature, which mainly focused on the crop classification task. It highly benefits from the high temporal resolution of Sentinel (5-6 days). Similarly to many countries, France is investigating the potential of Sentinel imagery for crop monitoring so as to define an operational 2020+ CAP workflow. This paper details the framework proposed to the French Paying Agency (Agence de Services et de Paiement, ASP) that addresses various monitoring scenarios jointly built with ASP.

The vast literature dealing with crop mapping has demonstrated the relevance of satellite image time series (STIS) both at the pixel and parcel levels (Belgiu and Csillik, 2018; Sitokonstantinou et al., 2018), especially with the joint exploitation of Synthetic

Aperture Radar (SAR) and optical images (Veloso et al., 2017; Neetu and Ray, 2020). In particular, deep learning techniques have recently shown their suitability to foster temporal information extraction from multi-modal STIS (Zhao et al., 2020; Adrian, Sagan, and Maimaitijiang, 2021) and their high discrimination power for a large range of crop types (Kussul et al., 2017; Ji et al., 2018; Sainte Fare Garnot et al., 2020; Rußwurm and Körner, 2020). However, they have not yet proved to be country-wise compliant. Several methods fit to operational constraints (scalability, automation, computing times), with explainable algorithms (Inglada et al., 2017; Konduri et al., 2020), assuming the desired nomenclature remains simple (<15 classes, e.g., (Pott et al., 2021)). Our contribution both lies on the definition of an operational framework integrating SAR and optical images as well as country-scale validation. Experiments show our solution based on Sentinel-1 (S1) and Sentinel-2 (S2) images with Random Forests provides satisfactory results both for 17 and 43-class crop classifications, on par with various deep-based SITS methods of the literature.

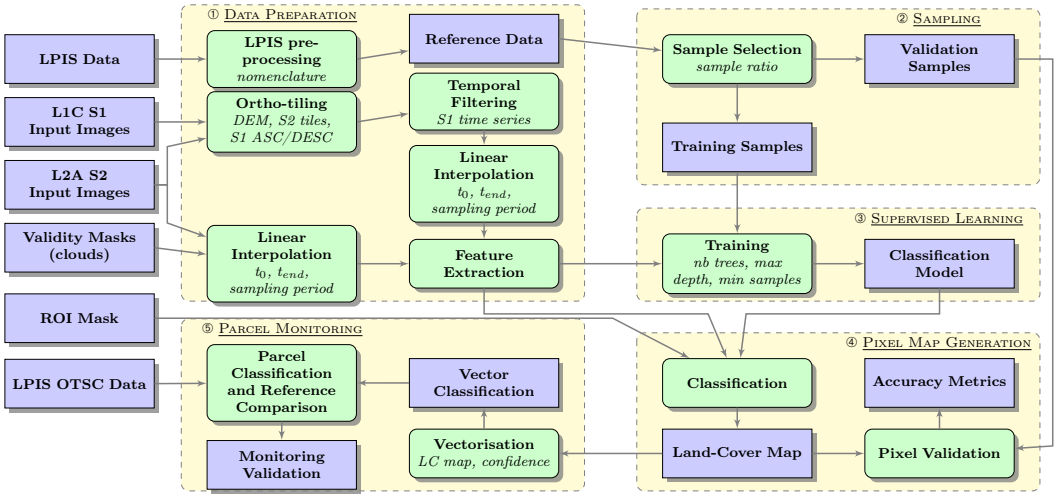


Figure 1. Our crop monitoring targeted workflow adapted from the standard ι^2 processing chain. Data and outputs are in purple while processes are colored in green. Main parameters, input and outputs of each atomic process are indicated in *italics* and detailed in Table 1.

2. Operational workflow and experiments

2.1. *Proposed solution*

We target to discriminate crops every year at the parcel level using freely available Sentinel images. Due to their well documented complementarity, both SAR Sentinel-1 and optical Sentinel-2 images are fed into the pipeline (Section 2.3). We classify these crops according to several use cases (2 nomenclature - 3 time frames, Section 2.2). We have access to the former French LPIS data, our discrimination problem can be cast as a supervised classification task based on a stack of S1+S2 images. For that purpose, we adopt the ι^2 framework (Inglada et al., 2017) based on Random Forests (RF). This is a perfectly tailored open-source solution (`https://framagit.org/iota2-project/iota2`), that fits with operational constraints and with automatic high resolution crop monitoring (Defourny et al., 2019). It has already demonstrated its high efficiency in performing country wide land cover classification tasks with multi-temporal high

resolution Landsat and Sentinel imagery at the pixel level. It is based on Orfeo Tool Box library and python scripts. It can run either on standard desktop computers or High Performance Computing clusters, therefore adapted to any end-user.
To remain scalable, explainable and adapted to parallelization schemes, we adopted the per-pixel mono-modal RF classification scheme augmented with SAR data and simple post-processing steps. Contextual classifiers and/or object-based (deep-based) approaches are known to be efficient but were discarded. We adapted the standard ι^2 workflow by (i) integrating S1 images and (ii) performing parcel monitoring with the generated per-pixel classifications.

Step	Parameter/Input/Output	Explanation
1	nomenclature	Set of classes to be discriminated (17 and 43 for Level 1 and Level 2, respectively)
	DEM	Digital Elevation Model used to orthorectify Sentinel-1 images
	S1 ASC/DESC	Images both from Sentinel-1 Ascending and Descending orbits
	t_0, t_{end}	Predefined starting and ending months for analysis (3 scenarios are tested)
	sampling period	The regular temporal grid on which radar and optical images are matched before feature computation
2	sample ratio	Ratio between learning and testing samples for the randomly selected parcels per class and ratio of pixels randomly kept for each parcel
3	nb trees, max depth	Number of trees of the Random Forest and maximal depth of each tree
	min samples	Minimum number of samples required to split an internal node in the Random Forest
5	confidence	Per-class confidence score computed for each parcel as the ratio of pixels labelled for this class among all pixels

Table 1. Explanation of the inputs, outputs and main parameters for the 5 steps of our framework (see Figure 1 and Section 2.1 for more details).

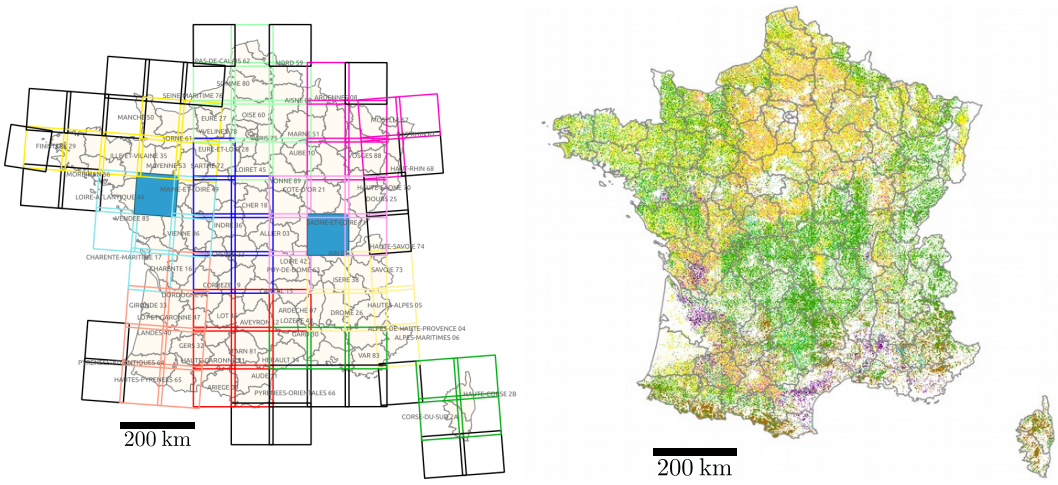
Our approach can be decomposed into five main steps (Fig. 1), our contributions lying in steps (1) and (5):
1. Data preparation. First, LPIS data is processed in order to fit to the desired nomenclature and be ready for step (2). Secondly, in order to extract multi-modal features, S1 images are orthorectified and matched onto S2 tiles (using S1tiling, `http://tully.ups-tlse.fr/koleckt/s1tiling`), filtered and interpolated to be spatially and temporally coherent with S2 images. S2 images are gap-filled with linear interpolation. We work on a 10-day regular temporal grid handling initial clouds and various temporal samplings.
2. Sampling. From reference data, we automatically extract training and validation samples for the classification task. For each class, a random selection of 2,500 parcels was performed. We kept only 50% of the parcels for classes with less than 5,000 objects. The ratio between learning/testing is 70%/30% and is performed at the parcel level.

For each learning parcel, a random subset of 10% of pixels is kept. More pixels would lead to insert redundant information and would not lead to increase the discrimination and generalisation ability of the classifier. No cleansing procedure is performed apart from excluding border pixels of the training polygons.

3. Supervised learning. RF have proved to be efficient for multi-modal and national-scale mapping (Inglada et al., 2017; Defourny et al., 2019) with noisy training data. The unit for the definition of a training model is the S2 tile (see Section 2.4), albeit the S1-S2 coregistration procedure allows to feed such models with both S1 and S2 information. This consists both of raw channels and attributes that have already been assessed to be relevant for multitemporal land-cover classification: VV and VH polarisation channels for S1 (Whelen and Siqueira, 2018), all 10 m+20 m spectral channels, NDVI, and NDWI for S2. We extract 14 features for each epoch. It means that, for a time range of n days, we have $14 \times \left\lfloor \frac{n}{10} \right\rfloor$ features in total.

4. Pixel map generation. The model is applied for each S2 tile, discarding non agricultural areas (Region Of Interest - ROI - mask), using both S1 and S2 information. It allows to generate a 10 m pixel-based crop raster map, which is known to be the most adequate spatial resolution (Defourny et al., 2019; Griffiths, Nendel, and Hostert, 2019).

5. Parcel monitoring. Pixel-based maps are vectorised. A per-class confidence score is computed for each parcel as the ratio of pixels labelled for this class among all pixels. The three dominant labels are kept, and the majority label is given to the parcel (*i.e.*, the class with the highest confidence score). This allows to remove salt and pepper noise due to non contextual classification as well as misclassification at the parcel edges (mixed pixels, Fig. 4). The final decision and the three main labels can then compared with On The Spot Check (OTSC) data available for the year of interest.



(a) A RF model is trained and applied for each S2 tile, using both S1 and S2 features. Each color corresponds to a specific server (10). Black tiles with few data were discarded and the two full blue tiles are the initial tuning areas.

(b) Results for ι^2 OTSC with Level 1 (color code: Table 2).

Figure 2. Classification at national scale.

2.2. *Use cases and implementation*

Three scenarios have been tested with two sets of classes (17 and 43 classes, Table 2), leading to 6 configurations. They correspond to different time ranges within the November (Year (Y-1)) – November (Year Y) time frame for crop monitoring of Year Y. They correspond to use cases formalized by ASP for verification of farmers' declarations with Copernicus data under the new regulation for 2020+ CAP payments.

A. Pre-filling farmers' aid form (denoted ι^2 Early in the following). The aim is to help farmers by decreasing the time needed to fill up the official aid application. The farmers would be only in charge of validating/correcting the proposed crop types. We use Sentinel images from November to April.

B. Guiding On The Spot Checks (ι^2 OTSC). Continuous monitoring can be a solution for Paying Agencies, replacing the responsibility of supervising at least 5% of the received declarations. Sentinels could help checking as fast as possible eligibility criteria and would allow to comply with legal cross compliance control systems with 100% declarations. We consider data from November of July.

C. Seamless claim crop monitoring (ι^2 Late). In such a case, we are moving towards a fully automatic aid application and payment system. The full time frame (November-November) can be fed into our system.

From the 300+ existing codes, two nomenclatures were set up by ASP experts. The sets of classes are based on crop phenology and differences in aid payment: we do distinguish between close crop species which lead to the same amount of CAP aid. At *Level 1*, 17 classes are considered and at the finer-grained *Level 2*, 43 classes (Table 2).

2.3. *Data*

We worked with year 2017. Sentinel-1 images were downloaded from the Copernicus portal at the L1C GRD (Ground Range Detected) level (Ascending and Descending orbits). Images were orthorectified with the SRTM (Shuttle Radar Topography Mission) Digital Elevation Model (DEM). Sentinel-2 images were extracted at level L2A from French Theia webportal. When we started our work, the MAJA (MACCS-ATCOR Joint Algorithm) cloud mask algorithm and radiometric correction techniques available through Theia were of better quality than those from Copernicus (Baetens et al., 2019). S2 images with less than 75% cloud cover were retained. 73 of the 98 Sentinel-2 tiles (110 km×110 km) covering France were processed. We discarded tiles with few data (Fig. 2). In total, this corresponds to 75 TB of data, including 11.5, 20.5, 34, and 8.2 TB of Sentinel archives, uncompressed S1/S2 data, pre-processed S1/S2 data, and classification results, respectively.

The French LPIS is the Référentiel Parcellaire Graphique (RPG), made available every year. It was provided by the French Paying Agency (ASP). Such reference data come from farmers' declarations (both spatial extent and crop types), checked and refined by human operators through visual interpretation of Very High Resolution images and field surveys. Label noise-free data cannot be guaranteed. We assume noise having a limited impact on the learning and the quantitative assessment procedures. For our study, parcels were anonymised, merged according to the selected nomenclatures, their geometry was corrected if necessary, split up into Sentinel tiles. Very small parcels were filtered out from the learning stage due to their limited number of pixels, that mostly fail at the border with other (potentially distinct) crops. Such mixed pixels are likely to deteriorate the learning process A threshold of 10 pixels is selected to avoid

	Class	No. of parcels (France)	Highest F-score (Method, Gain)
Level 1	Predominant herbaceous (1)	5,848,985	96.4 (PSE-TAE, 4.6)
	Cereal (2)	3,134,464	97.7 (PSE-TAE, 7.3)
	Hemp (3)	4,514	85.7 (ι^2 OTSC, 13.7)
	Leguminous fodder (4)	326,175	47.9 (ι^2 OTSC, 2)
	Other fodder (5)	33,297	54.1 (ι^2 Late, 27)
	Oilseed (6)	493,374	97 (PSE-TAE, 9.6)
	Protein crop (7)	139,083	89.9 (PSE-TAE, 14.6)
	Rice (8)	1,904	—
	Industrial cultivation (9)	167,854	52.3 (ι^2 OTSC, 2.1)
	Fruit/vegetable/flower (10)	142,651	64.4 (PSE-TAE, 4.2)
	Aromatic plant (11)	28,050	44.9 (ι^2 OTSC, 19.7)
	Ligneous (12)	201,639	40.9 (PSE-TAE, 9.7)
	Orchard (13)	158,486	15.5 (PSE-TAE, 5.2)
	Vineyard (14)	534,546	95.9 (PSE-TAE, 4.6)
	Copse short rotation (15)	2,595	8.7 (ι^2 Late, 1.6)
	Non agricultural areas (16)	484,595	57.7 (PSE-TAE, 32.1)
	Other (17)	44,525	27.9 (ι^2 OTSC, 0.5)
Level 2	Predominant herbaceous (1.1)	5,848,985	96.6 (PSE-TAE, 5.3)
	Spring durum wheat (2.1)	2,181	—
	Winter durum wheat (2.2)	93,670	72.7 (ι^2 Late, 26.8)
	Spring cereal (2.3)	150,208	72.7 (ι^2 OTSC, 7.2)
	Summer cereal (2.4)	968,608	97.3 (PSE-TAE, 8.4)
	Winter cereal (2.5)	1,837,711	96.5 (PSE-TAE, 7.6)
	Maize zea (2.6)	13	—
	Sorghum millet (2.7)	38,269	47.7 (ι^2 Late, 6.6)
	Other cereals (2.8)	43,804	46.1 (ι^2 Late, 16.5)
	Hemp (3.1)	4,514	92 (ι^2 OTSC, 13.6)
	Buckwheat (3.2)	18,036	42.4 (ι^2 Late, 17.1)
	Leguminous fodder (4.1)	326,175	48.2 (PSE-TAE, 6.2)
	Other fodder (5.1)	33,297	53.2 (ι^2 Late, 18.2)
	Winter rapeseed (6.1)	304,166	98.4 (PSE-TAE, 10.4)
	Spring rapeseed (6.2)	976	—
	Sunflower (6.3)	183,390	91.9 (PSE-TAE, 4.7)
	Other oilseed (6.4)	4,842	81.8 (ι^2 OTSC, 24.2)
	Soybean (7.1)	45,343	95.2 (PSE-TAE, 17.9)
	Other protein crop (7.2)	93,740	67.3 (ι^2 Late, 13.8)
	Rice (8.1)	1,904	—
	Beet (9.1)	81,308	—
	Flax (9.2)	26,951	65.3 (ι^2 Late, 20.5)
	Potato (10.1)	57,828	51.2 (PSE-TAE, 2.5)
	Tobacco (10.2)	1,767	—
	Dried vegetables (10.3)	18,840	—
	Other fruit/flower/veg. (10.4)	123,811	68.6 (PSE-TAE, 8.6)
	Lavender (11.1)	16,734	—
	Other aromatic plant (11.2)	11,316	39.1 (ι^2 Late, 27.6)
	Ligneous (12.1)	201,639	41.2 (PSE-TAE, 7.2)
	Oliver grove (13.1)	20,210	—
	Other orchard (13.2)	138,276	21.1 (PSE-TAE, 10.2)
	Vineyard (14.1)	534,546	96.6 (PSE-TAE, 5.2)
	Copse short rotation (15.1)	2,595	24 (ι^2 Late, 14.9)
	Non agricultural areas (16.1)	484,595	57.5 (PSE-TAE, 29.8)
	Building (17.1)	573	—
	Biomass (17.2)	3,469	68 (GRU, 2.9)
	Woods (17.3)	12,510	23.1 (PSE-TAE, 3.4)
	Hops (17.4)	638	—
	Water marsh (17.5)	348	—
	Other (17.6)	8,951	35.7 (ι^2 Late, 25)

Table 2. The two nomenclatures (Levels 1 and 2). The top F-scores for each class are reported for `T31FM` tile, with the corresponding method (see Table 3 and Section 3.2) and the increase in F-score with respect to the best solution of the other family (RF $\leftrightarrow$ deep learning). The mapping between classes of Levels 1 and 2 is provided between brackets (For Level 2, the number "$X.Y$" refers to the Y^{th} sub-class of the class X of Level 1

).

irrelevantly selecting a single pixel in a given parcel (with the 10% training rate).

2.4. *Moving towards country-wide classification*

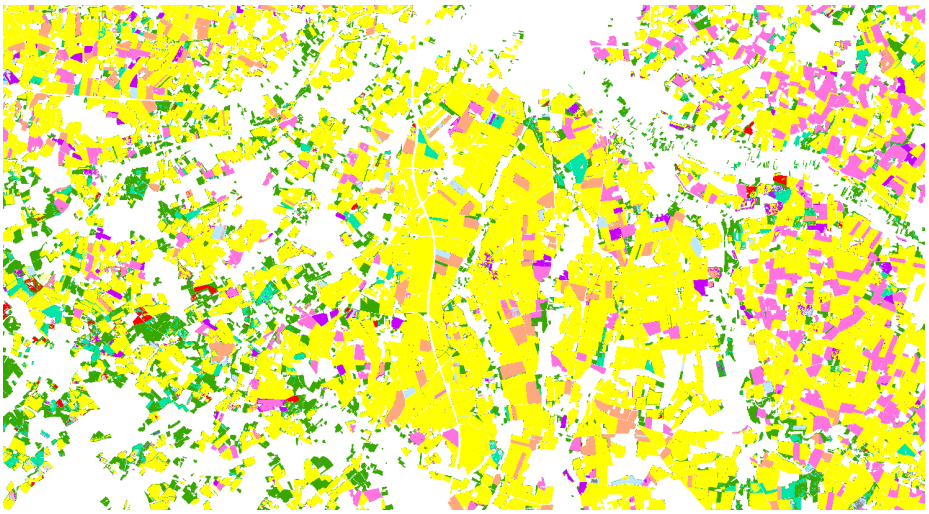
The proposed workflow is straightforwardly made scalable by considering each Sentinel-2 tile independent and subsequently the unit for parallel computation. Therefore, all steps are performed by tile (training, prediction, and evaluation), following Inglada et al. (2017). Instead of standard batch processing, here, streaming computation is enabled thanks to the possibility of chaining Orfeo ToolBox applications in RAM instead of handling temporary files. This allows to efficiently cope with the vast amount of Sentinel data with significant pre-processing steps. The open source OpenMPI (Message Passing Interface) library is adopted for distributing the jobs to the available 10 servers (Section 2.2).
We performed initial tests on 2 tiles to tune our processing chain (most suitable parameters and server configuration, Fig. 2): the sample ratios for Step (2), the Random Forest hyperparameters (100 trees, 25-level depth, min. 5 samples per node), the strategy for vector classification. A computing server with 12 CPU, 2 TB, 24GB RAM was used for that purpose. For large scale experiments, we worked with 10 servers with at least 8 vCpu, 5-8 TB, 16GB RAM each (5 from IGN-France and 5 from Copernicus Data and Information Access Services).

3. Results

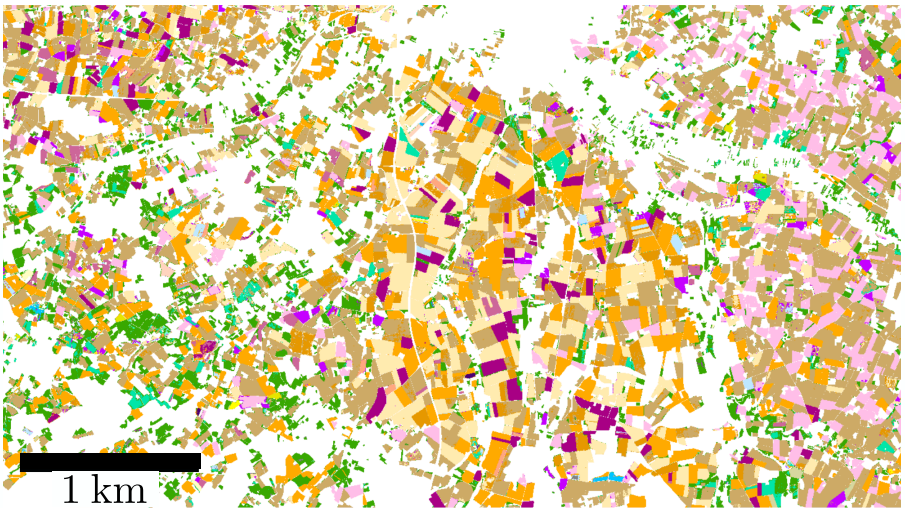
3.1. *Quantitative and qualitative assessment*

For a tile, 4 days are necessary to pre-process Sentinel and reference data for a full year while 7-8 days suffice for performing the 6 classification tasks. The label with the largest number of pixels is affected to the parcel. Confusion matrices, Precision, Recall, F-score metrics were computed for each tile for each use case using farmers' declarations (Fig. 2 and 5). Initial experiments performed over `T31FM` tile for the ι^2 OTSC permits to assess the relevance of Sentinel-1 images. Feeding our pipeline with only S2 images lead to an Overall Accuracy of 83.25 and 82.8%, for Levels 1 and 2, respectively. This is slightly inferior to our S1+S2 scenario with Overall Accuracy (OA)=87.97% and% 86.59%.
At country-scale, high scores (82-97%) are obtained for most dominant classes (*predominant herbaceous*, *cereal*, *oilseed*, *protein crop*), as well as rare classes with specific signatures (*vineyard*, *rice*, *hemp*, *lavender*). *Orchards* and *Fruits* are often misclassified due to the poor texture and limited spatial resolution of S1/S2 images. The *predominant herbaceous* class exhibits high scores but we noted a significant confusion with *fodder* and *protein crop*. Due to its dominance, misclassification less affects *herbaceous* recall than *fodder/protein crop* precision. Better results are obtained for flat landscapes and in areas with many crops, compared to mountainous areas. Eventually, ι^2 Early results are worse that ι^2 OTSC and ι^2 Late. ι^2 Late does not show genuine superior performance compared with ι^2 OTSC, several classes even show lower accuracy scores (Fig. 5), OTSC (until July) could be considered as a valid scenario resulting in a smaller amount of data to ingest. It varies with the crop type and can be explained with the phenological stage (You and Dong, 2020) or the emergence of another crop type in November.

The final vectorisation step allows to remove local errors either due to border (mixed)

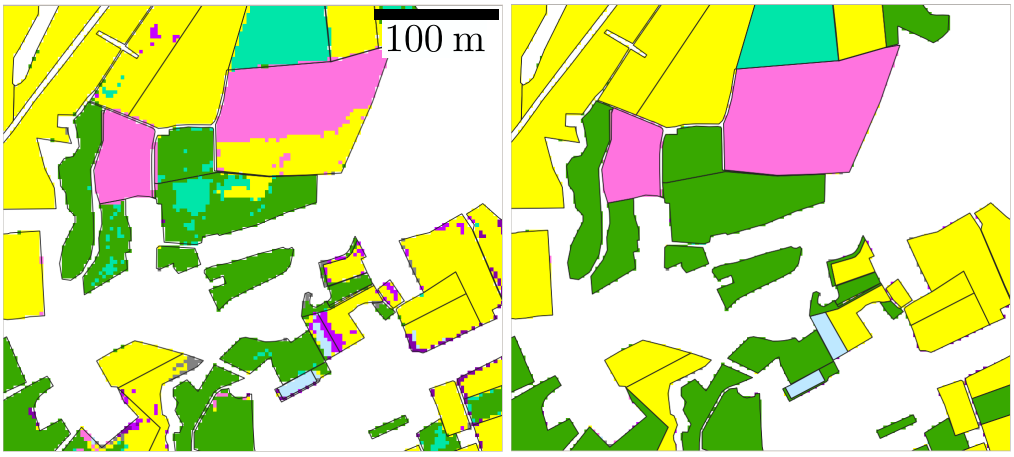


(a) Level 1 with 17 classes.



(b) Level 2 with 43 classes.

Figure 3. ι^2 OTSC results. Color code is provided in Table 2.



(a) Initial per-pixel classification.

(b) One label per parcel, based on the dominant class. We can observe less noise but intra-parcel variability is lost.

Figure 4. Advantages and limitations of the object-based approach. Same behaviour is noticed for the three investigated scenarios (ι^2 Early, ι^2 OTSC, ι^2 Late). Color code is provided in Table 2.

	Method	**Precision** (%)	**Recall** (%)	***F*-Score** (%)	**OA** (%)
Level 1	ι^2 Early (✱)	74.20	43.92	43.32	75.46
	ι^2 OTSC (✱)	**81.06**	52.27	54.89	87.97
	ι^2 Late (✱)	80.28	50.15	53.35	87.86
	GRU (✲)	50.26	52.85	50.94	92.97
	LSTM (✲)	49.54	52.32	50.51	92.96
	ConvLSTM (S2, ✻)	64.00	44.50	49.20	93.10
	PSE-TAE (S2, ★)	62.65	**54.15**	**56.9**	**94.05**
Level 2	ι^2 Early (✱)	79.05	34.29	39.07	82.84
	ι^2 OTSC (✱)	**85.18**	47.18	51.97	86.59
	ι^2 Late (✱)	81.23	53.24	55.89	84.71
	GRU (✲)	46.15	51.96	48.37	92.08
	LSTM (✲)	45.95	52.99	48.61	92.08
	ConvLSTM (S2, ✻)	56.98	48.50	50.94	92.50
	PSE-TAE (S2, ★)	67.35	**58.01**	**61.35**	**94.2**

Table 3. Performance of various existing solutions on the `T31FM` tile for the 2 nomenclatures. S2: classification performed only on Sentinel-2 images. Values in **bold** correspond to the highest value among all methods for each of the four criteria. ✱, ✲, ✻, ★ indicate the learning strategy (details are provided in (Sainte Fare Garnot et al., 2019)). Direct comparison is not possible due to varying learning schemes.

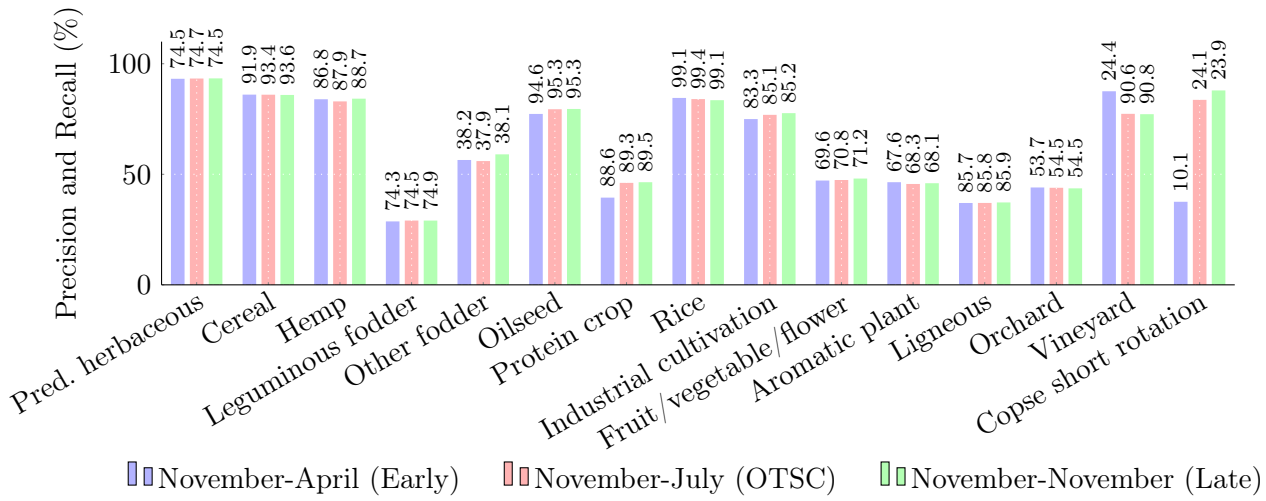


Figure 5. Accuracy for the three ι^2 scenarios for the 17-class problem (Full France). Bars correspond to Precision and Recall is reported on top.

pixels or irrelevant objects within parcels. This also suppresses multiple crop types within one parcel (Fig. 4) and potentially highly relevant information about the intra-parcel variability. Keeping the 3 dominant classes is bound to be relevant if further refinement is required by the Paying Agency.

3.2. *Performance with respect to existing solutions*

Our RF-based solutions are tested against several deep-learning approaches, using our own available implementations and prediction at the parcel level on `T31FM` tile: Recurrent Neural Networks (GRU, LSTM (Ienco et al., 2017), ConvLSTM (Rußwurm and Körner, 2018), for Gated Recurrent Unit, Long Short-Term Memory, and Convultional LSTM, respectively) and a hybrid Spatio-Temporal attention-based architecture (PSE-TAE, for Pixel-Set Encoder Temporal Attention Encoder) exhibiting state of the art results (Sainte Fare Garnot et al., 2020). We selected such methods since the temporal dimension is superior to the spatial one for Sentinel-based crop mapping (Sainte Fare Garnot et al., 2019) and since recurrent and attention-based mechanisms are superio to convolutional approaches (Rußwurm and Körner, 2020). Similarly to us, GRU and LSTM solutions are fed with S1+S2 handcrafted features (28: mean and standard deviation for the 10 S2 spectral bands, NDVI, HH, VV, VV/HH) to assess the impact of keeping the temporal structure of the data. ConvLSTM and PSE-TAE are fed only with S2 images but are adopted to evaluate the relevance of end-to-end learning. Direct comparison is not possible: features, training procedure (number of samples, split train/test/validation) and validation are not aligned. In particular, ι^2 solutions were tested on all parcels, including training ones. For classes with limited samples (<5,000), this can lead to a slightly positive bias for mean accuracy metrics. PSE-TAE discard classes with less than 100 samples leading to 20 classes for Level 2.
Performances are provided in Table 3. RF-based solutions give higher precision scores while deep-based ones exhibit higher OA values. PSE-TAE performs better but without a significant gap and without considering rare classes. In such a case, our solution shows higher precision/recall/F-Score values for both levels, validating its relevance at the local scale. Table 2 shows also that PSE-TAE is prone to provide a better accuracy

even without S1 imagery. It better exploits the temporal coherence, lost with a RF-based approach. Still, almost 50% of the classes of both levels are better discriminated with our workflow, even sometimes with a large margin (*protein crop*, *winter durum wheat*, *other fodder* and *oilseed* for the most significant classes). This again shows the validity of our solution in a strongly unbalanced and large-scale context and that the temporal structure of time-series of images is not always a meaningful information. However, we can only conclude that our framework is on par with the state of the art with a majority of crop classes and perform worse for several key dominant classes. We therefore validate the conclusions of (Xu et al., 2020) at larger scale and with more classes.

4. Conclusion

We satisfactorily provided an operational framework for country-wide crop classification. It is based on an improved version of the open source ι^2 solution with Sentinel imagery. Results show high accuracy for the main classes, locally close to state-of-the-art methods. This validates both the proposed nomenclatures and the remote sensing capability to fulfill CAP 2020+ requirements. Improvement lies in the design of a more hierarchical strategy that handles separately dominant and rare classes at the tile and national levels, respectively.

Acknowledgment

The authors are partly supported by the French Paying Agency (ASP) and by the French National Research Agency under the grant ANR-18-CE23-0023. We thank Vivien Sainte-Fare-Garnot for the implementation and the experiments in Section 3.2.

References

Adrian, J., V. Sagan, and M. Maimaitijiang. 2021. "Sentinel SAR-optical fusion for crop type mapping using deep learning and Google Earth Engine." *ISPRS Journal of Photogrammetry and Remote Sensing* 175: 215–235.

Baetens, L., et al. 2019. "Validation of Copernicus Sentinel-2 Cloud Masks Obtained from MAJA, Sen2Cor, and FMask Processors Using Reference Cloud Masks Generated with a Supervised Active Learning Procedure." *Remote Sensing* 11: 433.

Belgiu, M., and O. Csillik. 2018. "Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis." *Remote Sensing of Environment* 204: 509–523.

Campos-Taberner, M., et al. 2019. "A Copernicus Sentinel-1 and Sentinel-2 Classification Framework for the 2020+ European Common Agricultural Policy: A Case Study in València (Spain)." *Agronomy* 9 (9): 556.

Defourny, P., et al. 2019. "Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world." *Remote Sensing of Environment* 221: 551–568.

Griffiths, P., C. Nendel, and P. Hostert. 2019. "Intra-annual reflectance composites from

Sentinel-2 and Landsat for national-scale crop and land cover mapping." *Remote Sensing of Environment* 220: 135–151.

Ienco, D., R. Gaetano, C. Dupaquier, and P. Maurel. 2017. "Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks." *IEEE Geoscience and Remote Sensing Letters* 14 (10): 1685–1689.

Inglada, J., et al. 2017. "Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series." *Remote Sensing* 9 (1): 95.

Ji, S., C. Zhang, A. Xu, Y. Shi, and Y. Duan. 2018. "3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images." *Remote Sensing* 10 (2): 75.

Konduri, V.S., et al. 2020. "Mapping crops within the growing season across the United States." *Remote Sensing of Environment* 251: 112048.

Kussul, N., M. Lavreniuk, S. Skakun, and A. Shelestov. 2017. "Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data." *IEEE Geoscience and Remote Sensing Letters* 14 (5): 778–782.

Neetu, and S. S. Ray. 2020. "Evaluation of different approaches to the fusion of Sentinel -1 SAR data and Resourcesat 2 LISS III optical data for use in crop classification." *Remote Sensing Letters* 11 (12): 1157–1166.

Pott, L.P., et al. 2021. "Satellite-based data fusion crop type classification and mapping in Rio Grande do Sul, Brazil." *ISPRS Journal of Photogrammetry and Remote Sensing* 176: 196–210.

Rußwurm, M., and M. Körner. 2018. "Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders." *ISPRS International Journal of Geo-Information* 7 (4): 129.

Rußwurm, M., and M. Körner. 2020. "Self-attention for raw optical Satellite Time Series Classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 169: 421–435.

Sainte Fare Garnot, V., L. Landrieu, S. Giordano, and N. Chehata. 2019. "Time-Space Tradeoff in Deep Learning Models for Crop Classification on Satellite Multi-Spectral Image Time Series." In *IGARSS*, 6247–6250.

Sainte Fare Garnot, V., L. Landrieu, S. Giordano, and N. Chehata. 2020. "Satellite Image Time Series Classification with Pixel-Set Encoders and Temporal Self-Attention." In *CVPR*, 12325–12334.

Sitokonstantinou, V., et al. 2018. "Parcel-Based Crop Identification Scheme Using Sentinel-2 Data Time-Series for the Monitoring of the Common Agricultural Policy." *Remote Sensing* 10: 911.

Veloso, A., et al. 2017. "Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural application." *Remote Sensing of Environment* 199: 415–426.

Whelen, T., and P. Siqueira. 2018. "Time-series classification of Sentinel-1 agricultural data over North Dakota." *Remote Sensing Letters* 9 (5): 411–420.

Xu, J., et al. 2020. "DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping." *Remote Sensing of Environment* 247: 111946.

You, N., and J. Dong. 2020. "Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth Engine." *ISPRS Journal of Photogrammetry and Remote Sensing* 161: 109 – 123.

Zhao, W., Y. Qu, J. Chen, and Z. Yuan. 2020. "Deeply synergistic optical and SAR time series for crop dynamic monitoring." *Remote Sensing of Environment* 247: 111952.