

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

Land cover mapping at national scale with Sentinel-2 and LUCAS: a case study in Portugal

Benevides, Pedro, Silva, Nuno, Costa, Hugo, Moreira, Francisco, Moraes, Daniel, et al.

Pedro Benevides, Nuno Silva, Hugo Costa, Francisco D. Moreira, Daniel Moraes, Mauro Castelli, Mário Caetano, "Land cover mapping at national scale with Sentinel-2 and LUCAS: a case study in Portugal," Proc. SPIE 11856, Remote Sensing for Agriculture, Ecosystems, and Hydrology XXIII, 1185606 (12 September 2021); doi: 10.1117/12.2598789

SPIE.

Event: SPIE Remote Sensing, 2021, Online Only

Land cover mapping at national scale with Sentinel-2 and LUCAS: a case study in Portugal

Pedro Benevides^{*a}, Nuno Silva^b, Hugo Costa^{a,b}, Francisco D. Moreira^a, Daniel Moraes^a, Mauro Castelli^b, Mário Caetano^{a,b}

^aDireção-Geral do Território, 1099-052 Lisboa, Portugal; ^bNOVA Information Management School (NOVA IMS), Universidade Nova Lisboa, Campus de Campolide, 1070-312 Lisbon, Portugal

ABSTRACT

Experiments were carried out to investigate the use of Land Use and Coverage Area frame Survey (LUCAS) dataset and Sentinel-2 imagery to produce a land cover map in Portugal through automated supervised classification. LUCAS is a free land cover land use (LCLU) dataset based in Europe, while Sentinel-2 satellites provide also free images with short revisit frequency. The goal was to evaluate if LUCAS dataset from 2018 can be used as a single reference dataset for land cover classification at national level. The Random Forest (RF) algorithm was used. Some processing steps were undertaken to use LUCAS as reference dataset. The original LUCAS LCLU nomenclature was modified into a new nomenclature composed of 12 and 6 level-2 and level-1 map classes, respectively. Filtering was performed on LUCAS metadata, reducing the initial number of LUCAS points over Portugal from 7168 to 4910. Monthly composites of Sentinel-2 images acquired between October 2017 and September 2018 were used. To reduce the imbalance in LUCAS training points, an oversampling technique based on Synthetic Minority Over-Sampling Technique (SMOTE) was used. An independent validation dataset was produced with 600 points. RF shows an overall accuracy (OA) of 57% for level-2 and 72% for level-1 nomenclatures. When using the oversampling technique, the OA accuracy increases by 3% for level-2 and 2% for level-1. The preliminary results of this experiment show that LUCAS dataset used in supervised machine learning classification has potential to produce a reliable land cover map at national scale.

Keywords: LUCAS survey, Land Cover, National mapping, Random Forest, Oversampling, Sentinel-2

1. INTRODUCTION

The Land Use and Coverage Area frame Survey (LUCAS) is a land cover land use (LCLU) European database promoted by the European Union (EU) under the Eurostat organization. The survey is conducted every 3 years since 2006, being the latest published in 2018, providing more than 300 000 points spread all over the EU countries. Its distribution follows a 2 km regular grid of points with data being collected in situ by surveyors. Each point contains information regarding its positional accuracy, land cover and land use characterization, soil and other environmental features. Considering the land cover information each point describes a circle area of 1.5 m radius as a rule, but wider radius of 20 m is used if the point is located in heterogeneous landscapes. Points which are not verified in the field campaign are filled by interpretation of satellite images and orthophotos. Although the main LUCAS purpose is to provide harmonized LCLU statistics among the EU, its European coverage with detailed information and its free access makes it attractive for land cover map production through supervised classification of remote sensing data.

The EU also provides Earth Observation images from the Sentinel-2 satellites which are freely accessible through the Copernicus program. Sentinel-2 delivers multi-spectral images with detailed spatial resolution at a five day revisit frequency, making it suitable for land cover applications. The global coverage, large spectral information and high temporal frequency of the Sentinel-2 combined with the computational advances allowed larger amount of data to be handled, leading to land cover classification on extensive areas at regional, national or even continental scales^{1,2}. However, another main challenge for land cover classification still remains, which is the existence of a reliable training dataset to ensure that a quality map is obtained³.

*pbenevides@dgterritorio.pt; phone (+351) 21 381 96 00; <https://www.dgterritorio.gov.pt/>

In recent years LUCAS database has been used in several studies to perform supervised classification of land cover information^{1,3,4,5,6,7}. Some of them are focused on national land cover (LC) classification^{3,4,6} and others on broader areas such as continental Europe⁷. A considerable part of these studies contemplate LUCAS as the unique training source for classification^{1,4,5,7}. LUCAS survey attributes can be used to select suitable data for land cover classification, avoiding points which may have mixed LCLU characteristics. Some studies have used this approach following diverse quality criteria to filter reference LUCAS points^{4,6,7}. Another way to condition the selection of information is to aggregate the original LUCAS LCLU nomenclature into a smaller number of classes. Several studies have followed this strategy reducing the number of classes to no more than 12¹, adapting the nomenclature to specific investigation interests, different landscape locations and mapping areas of various dimensions. Some authors have also reported that the LUCAS data is generally fairly unbalanced for land cover classifications at regional or national scale, even after class reduction^{3,4,7}. One of the applied strategies to mitigate this condition is based on the manual correction of LUCAS samples, relabeling the class or moving the point location to the correct LC^{3,5}. More sophisticated approaches involving oversampling minority classes through replication to match the quantity of majority classes have also been explored⁸. Synthetic Minority Over-Sampling Technique (SMOTE) is one of the most used oversampling techniques in machine learning classification and it is based on the creation of artificial data points along the lines that connect the k-nearest neighbors on the minority classes⁹.

Random Forest (RF) is one of the machine learning algorithms most widely used when it concerns supervised classification for land cover mapping applications¹⁰. RF is able to produce high classification accuracies, with similar results when comparing to other more complex and recent algorithms, but with a lower computational cost regarding the choice of parameters. It can be trained in a high dimensional dataset without overfitting and is somewhat robust to outliers and noise present in the input data¹¹.

In this study experiments were carried out to investigate the use of LUCAS dataset from 2018 and Sentinel-2 imagery to produce a land cover map at a national scale, in Portugal, through automated supervised classification. In particular, we are interested in evaluating if LUCAS can be used as a single reference dataset for land cover classification at the national level. Some processing steps were applied to LUCAS dataset to remove observations with potentially misleading information. Class nomenclature reclassification to 12 mapping classes was performed and point filtering based on LUCAS survey attributes information was implemented. Monthly composites of Sentinel-2 images acquired over Portugal during the 2018 agricultural year were used in this study. Additional derived data was also computed, including spectral indices and spectral-temporal metrics. To reduce the unbalance in training, an oversampling technique based on SMOTE was used. A comparison between two classification scenarios, LUCAS training and LUCAS training with SMOTE, was performed. An independent validation dataset was produced with 600 points manually labelled. Accuracy assessment was done with overall accuracy estimation using a confusion matrix, along with the estimation of users and producers accuracy by class.

2. DATA AND METHODS

2.1 LUCAS processing

The study area chosen for this experiment was continental Portugal, which includes 7168 points of the LUCAS 2018 dataset (Figure 1a). To use LUCAS as a reference source for land cover (LC) classification, some processing steps were undertaken. The LUCAS LCLU nomenclature is divided into 8 main classes and 88 subclasses concerning LC and 4 main classes and 47 subclasses related to land use (LU)¹². In order to adapt the original nomenclature for supervised land cover classification, reclassification from LUCAS subclasses was employed to a nomenclature closer to the Copernicus Corine Land Cover map (CLC). Table 1 indicates the resulting classes obtained from the reclassification and their LUCAS class original correspondence. The defined nomenclature was composed by 20 LC classes. Some of the original LUCAS classes were excluded from the correspondence due to their potential thematic ambiguity, which makes them problematic for training⁷. After the reclassification process the number of LUCAS points decreased to 6646, representing a reduction of about 7% (Figure 1b).

Thereafter, an additional filtering step was performed considering LUCAS survey attributes information. Since the main purpose of the survey is not land cover classification, some of the observations can have mixed class information, such as transitional areas, points falling in narrow linear or small features, or non agricultural areas susceptible to land cover variation during the year. With this in mind, some LUCAS attributes were considered in order to exclude or include observations for training. Two criteria were followed considering the spatial dimension and LCLU special remarks.

Spatial dimension criterion included the selection of observations that were observed directly on point, located in areas larger than 0.1 ha, located in features with width equal or larger than 20 m and with percentage of area covered larger than 50%. The second criterion was related to points with the attribute “LC_LU_Special_Remark”, generally including points with “Harvested field”, “Tilled/sowed” or “No remark”, excluding broadleaved points with “Harvested field” (avoiding possible indication of clear cut) and excluding spontaneously re-vegetated points with “Harvested field” or “Tilled/sowed (avoiding possible confusion with permanent bare soil).

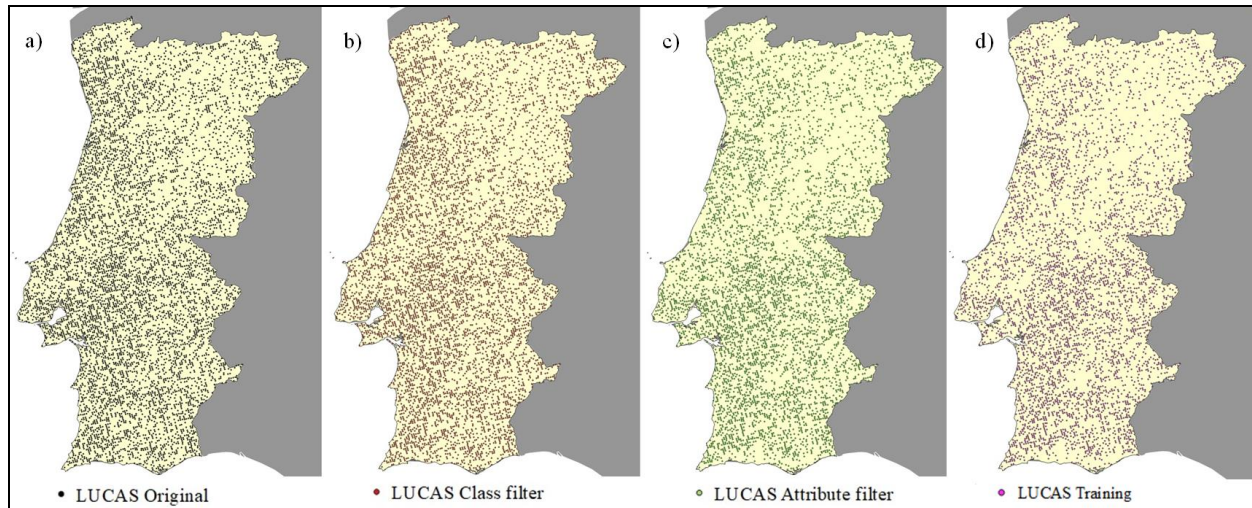


Figure 1. LUCAS 2018 points over the study area of continental Portugal. Black dots represent all the LUCAS points without filter (a), red dots the points after class reclassification (b), green dots the points after attribute filtering (c) and purple dots the points after final nomenclature definition and class removal (d).

Table 1. Class nomenclature defined as “LCLU map”, including the corresponding LUCAS classes. The number of points at each filter step is also presented, together with the number and percentage of excluded points processing steps.

LCLU map Level-2	LUCAS corresponding classes	Class filter points	Attribute filter points	Points excluded	Points excluded (%)
Urban	A11, A12	183	85	98	54
Non built-up area features	A21	165	100	65	39
Rainfed	B11, B12, B13, B14, B15, B18, B52,B54	179	165	14	8
Irrigated	B16, B21 B31, B41, B42	150	57	93	62
Rice	B17	26	19	7	27
Temporary pastures	B55	100	81	19	19
Orchards	B71, B72, B73, B74, B75, B76, B77	96	40	56	58
Olive groves	B81	161	72	89	55
Vineyards	B82	80	45	35	44
Broadleaved	C10	2494	2304	190	8
Coniferous	C21, C22, C23	657	601	56	9
Mixed	C32, C33	202	181	21	10
Shrubland	D10, D20	718	563	155	22
Permanent pastures	E10 w/ U111, E20 w/ U111	787	699	88	11
Natural herbaceous	E10 w/ U415 or U420, E20 w/ U415 or U420	37	24	13	35
Spontaneous re-vegetated surfaces	E30	325	198	127	39
Bare soil permanent	F10, F20	46	37	9	20
Bare soil temporary	F40	140	112	28	20
Water	G11, G21	85	62	23	27
Wetlands	H11, H21	15	15	0	0
Total		6646	5460	1186	18%

The impact of these filtering steps can be visualized in Figure 1c. Table 1 shows the filter effect by class, where the number of points available per class after the nomenclature reclassification (Class filter points) is compared with the number of points that resulted after the attribute filter application (Attribute filter points). The number of points and percentage of points excluded from one filter to another is also presented. Overall a significant exclusion of points per class was not verified, resulting in a total percentage of point exclusion of 18%, and a total number of 5460 points. The most affected classes were the ones related to urban and some types of agriculture crops (Irrigated, Orchards, Olive groves) with percentage of points excluded larger than 50%. However, most of the classes with less number of points (Rice, Natural herbaceous or Bare soil permanent) were only marginally affected, not increasing too much the already unbalanced status of the LUCAS dataset.

Initial classification tests revealed that some of the LCLU map classes presented in Table 1 were not suitable for LC mapping. The classes related to permanent crops (Orchards, Olive groves, Vineyards) failed to be classified with LUCAS training probably due to different maturity of trees, causing variable spectral responses that can be confused with many other land cover types (Broadleaved, Natural herbaceous or other types of Agriculture). Other less pure classes were also removed, including Mixed forest, which is composed by Broadleaved and Coniferous tree types, Non built-up area features, which by class definition can include very distinct artificial structures (including yards, dams or docks) and Bare soil temporary, because this class also includes soil types with very different spectral properties (arable land, clear cut in forests and burnt areas)¹². The LCLU map previous nomenclature was therefore updated with the removal of the aforementioned classes. Table 2 shows the final LCLU nomenclature which is defined by a total of 12 classes in Level-2 and 6 classes in Level-1. Classes in Level-2 have direct correspondence with training points from classes in Table 1, with the exception of Spontaneous re-vegetated surfaces that was aggregated into Natural herbaceous class and both Temporary and Permanent pastures that were aggregated to Pastures class.

Table 2. LCLU map final class nomenclature.

LCLU map Level-1	LCLU map Level-2
Urban	Urban
Agriculture	Rainfed
	Irrigated
	Rice
	Pastures
Natural herbaceous and Shrubland	Natural herbaceous
	Shrubland
Forest	Broadleaved
	Coniferous
Bare soil	Bare soil permanent
Water and Wetlands	Wetlands
	Water

After the class removal in the map final nomenclature the number of LUCAS points available for training was reduced again, resulting in a total of 4910 points (Figure 1d). This corresponds to a decrease of about 10% after the attribute filter, but the percentage reduction was significantly higher when comparing to the initial LUCAS points available in continental Portugal (about 30%). However, Figure 1 shows that the overall representation of points is not significantly affected and its even distribution by the whole territory is maintained.

The final filtered LUCAS points were used for training the supervised land cover classification experiment. These points can be considered the most reliable LUCAS sampling units for land cover supervised classification. However, they are significantly unbalanced as it can be observed in the training point representation by class in Table 3. In order to minimize the unbalanced condition, an oversampling technique based on SMOTE is used. A second training dataset based on LUCAS final training points was generated from the application of the SMOTE oversampling using a minimum threshold of 150 points per class. This increased the original sampling units in 6 of the 12 Level-2 classes (Table 3), resulting in a LUCAS based training dataset with 5730 points, which represents an increase of about 14%. Still, the sampling units for the most abundant class (Broadleaved) are 15 times larger than the least abundant classes (150 sampling units). Tests were made by applying undersampling techniques to the Broadleaved class, but classification results showed no improvements (not shown). Attempting to balance the training sampling units by class is important,

since it has been identified that classification with LUCAS dataset tends to slightly underpredict classes with smaller sample size and overpredict classes with larger sample size^{4,7}.

Table 3. Training points per class for the final LUCAS dataset and for the final LUCAS dataset with SMOTE.

LCLU map Level-2	LUCAS training points	LUCAS training points after SMOTE
Urban	85	150
Rainfed	165	165
Irrigated	57	150
Rice	19	150
Pastures	780	780
Natural herbaceous	222	222
Shrubland	563	563
Broadleaved	2304	2304
Coniferous	601	601
Bare soil permanent	37	150
Wetlands	15	150
Water	62	150

2.2 Sentinel-2 processing

The remote sensing data used in this supervised classification experiment were Sentinel-2 images acquired over Portugal between October 2017 and September 2018. This annual period, the 2018 agricultural year, represents the agricultural cycle of crops usually observed in continental Portugal and it is important to ensure spectral separability in classification between vegetation classes which are affected by phenological variability^{1,7}. The dataset used in this study has been utilized in several studies^{13,14,15,16}. Level-2A images were acquired from the Theia Land Data Centre, which are atmospherically corrected and already include cloud, shadow and topography corrections. Images with less than 50% cloud cover were downloaded from a total of 17 Sentinel-2 tiles that cover continental Portugal (Figure 2). These images were composed by 10 bands, which were all disaggregated to 10 m pixel size. Multi-temporal Sentinel-2 monthly composites were generated by calculating the median value for each band. Pixels affect by cloud or shadow mask were converted to missing data and a subsequent process of linear interpolation was applied to fill the data gaps. A total of 12 image monthly composites were produced. Additionally, 5 spectral indices were processed for each composite including NDVI, NBR, NDWI, NDBI and NDMIR. Spectral-temporal metrics were also computed from both monthly composites and spectral indices, based in 7 quantiles that included the 10th, 25th, 50th, 75th and 90th quantiles along with the difference between the 75th and 25th and the 90th and 10th. The input remote sensing data available for this classification experiment was composed of 285 variables, which are summarized in Table 4.

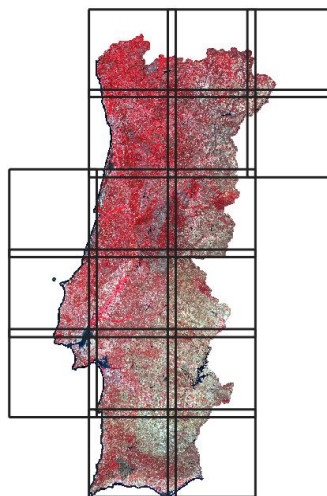


Figure 2. Sentinel-2 tiles covering continental Portugal superimposed to August 2018 monthly composite.

Table 4. Variables used in the supervised classification for this experiment.

Data type	Number of images	Input variables
Monthly composites	10 bands x 12 months	120
Spectral indices	5 indices x 12 months	60
Metrics composites	10 bands x 7 metrics	70
Metrics indices	5 indices x 7 metrics	35
Total	-	285

2.3 Random Forest classification

Random forest machine learning algorithm was selected for this supervised classification experiment, using LUCAS processed dataset as training and Sentinel-2 multi-temporal processed data as input dataset. Even though the RF algorithm is not significantly affected by its parameterization¹¹, the following two main parameters have been adjusted, namely the number of trees to 500 and the number of randomly selected features to perform each split as the square root of the total number of variables to $\sqrt{285}$. The supervised classification was performed for the two final LUCAS training datasets from Table 3, in order to evaluate the benefit of using oversampling techniques. All the classification steps were performed using Scikit Learn library from Python language¹⁷.

2.4 Validation database

An independent validation dataset was created to assess the performance of this classification experiment. The validation dataset is composed by 600 points that were evenly stratified by class using the LCLU map Level-2 nomenclature, corresponding to 50 points per class. The stratification was performed using the most recent LCLU official Portuguese map from 2018, Carta de Uso e Ocupação do Solo (COS). Afterwards, a manual labelling process was executed by expert technicians in LCLU using orthophoto maps from 2018, with a 25 cm pixel resolution. For each LUCAS dataset classification, confusion matrices were produced along with the estimation of users and producers accuracies (UA and PA) for each Level-1 and Level-2 LCLU map nomenclature, as well as the overall accuracy (OA).

3. RESULTS AND DISCUSSION

The OA results for this experiment are presented in Table 5, where Level-1 and Level-2 are shown together with the RF classification results using LUCAS training data with and without SMOTE application. The LUCAS training classification without oversampling presented an OA of 56.7% for Level-2 and 71.3% for Level-1, showing a significant increment of 14.6%. As for the LUCAS training dataset with SMOTE the obtained OA is 60.3% for Level-2 and 73.8% for Level-1, showing a slightly smaller increment of 13.5%. Comparing both LUCAS datasets, the RF classification with SMOTE training showed improvements of 3.6% for Level-2 and of 2.5% for Level-1. The OA improvements observed for the LUCAS training with SMOTE demonstrated that the oversampling techniques can be useful to further enhance the RF classification performance.

Table 5. Overall accuracy (OA) results for LCLU map classification experiments for Level-1 and Level-2.

RF classification	LCLU map Level-2	LCLU map Level-1
LUCAS training	56.7	71.3
LUCAS training with SMOTE	60.3	73.8

A comparison of the UA and PA obtained for the Level-2 LCLU map classes, for both LUCAS datasets classifications with and without application of oversampling, can be observed in Table 6. Classes with both higher UA and PA accuracy were Rice and Urban. Classes with only high UA were Rainfed, Irrigated, Bare soil permanent, Natural herbaceous (with SMOTE) and Wetlands. Classes with only high PA were Pastures, Broadleaved and Water. Extreme values of accuracy reaching 100% were found for several classes (e.g. Rice). On the other hand, values close to 1% were reported for Natural herbaceous PA. Generally, improvements have been observed for both PA and UA when evaluating results with SMOTE classification. Most notable improvement were found for Irrigated PA (16% increase), Bare soil permanent PA (39% increase) and Natural Herbaceous UA (67% increase). PA accuracy decreases in SMOTE dataset can also be found

but with a much lower magnitude, in Urban (2% decrease), Natural herbaceous (1% decrease) and Shrubland (5% decrease). As for UA, the decrease in SMOTE results were found in classes that reached 100% accuracy without oversampling, but still remaining high, with Wetlands decreasing 8% (92% UA) and Bare soil permanent almost 30% (72% UA).

Table 6. UA and PA obtained for LCLU map Level-2 for both LUCAS training datasets classifications.

LCLU map Level-2	LUCAS training points		LUCAS training points after SMOTE	
	Users accuracy	Producers accuracy	Users accuracy	Producers accuracy
Urban	80.0	85.7	87.5	83.3
Rainfed	66.7	15.4	71.4	20.0
Irrigated	100.0	44.4	100.0	60.6
Rice	100.0	100.0	100.0	100.0
Pastures	57.7	78.9	59.4	78.9
Natural herbaceous	33.3	1.9	100.0	1.0
Shrubland	41.1	60.7	42.7	56.1
Broadleaved	41.6	96.6	43.4	98.9
Coniferous	63.6	56.0	75.0	60.0
Bare soil permanent	100.0	22.7	72.0	62.1
Wetlands	100.0	35.8	91.7	41.5
Water	62.7	90.2	63.8	90.2

Table 7. Confusion matrix obtained for LCLU map Level-2 for LUCAS training dataset classification with SMOTE. The number of points (NP) of the validation dataset is presented together with the classes UA and PA.

	Ur	Ra	Ir	Ri	Pa	Nh	Sr	Br	Co	Bs	We	Wa	UA (%)	NP
Urban	35	1								4			88	40
Rainfed		5			2								71	7
Irrigated			20										100	20
Rice				31									100	31
Pastures	1	11	5		60	22				1	1		59	101
Natural herbaceous						1							100	1
Shrubland			2			32	32	1	1	2	4	1	43	75
Broadleaved	4	7	6		14	41	23	86	9	3	4	1	43	198
Coniferous						1	2		15		2		75	20
Bare soil permanent	2	1				4				18			72	25
Wetlands											22	2	92	24
Water										1	20	37	64	58
PA (%)	83	20	61	100	79	1	56	99	60	62	42	90		
Number points (NP)	42	25	33	31	76	101	57	87	25	29	53	41		

The confusion matrix for Level-2 classes obtained for the classification experiment with SMOTE oversampling is presented in Table 7. The matrix was calculated from the validation dataset points that were classified (rows) versus the validation points labelling or ground truth (columns). The total number of points (NP) for both classified (last column) and ground truth (last row) are also presented. The best accuracy was obtained for Rice, with no omission or commission errors. Irrigated classification has no commission errors but has some omission, failing to classify 13 out of 33 points. Natural herbaceous was another class that reached 100% UA, but it only classified correctly 1 point out of 101, resulting in large omission error. Indeed, Natural herbaceous was responsible for some of the largest errors, being mainly confused with Pastures (22 points), Shrubland (32 points) and Broadleaved (41 points). Similar confusions have been reported in other studies⁷. This may be related with the LUCAS survey point nature, in which most of them are located in areas with Shrubland or Broadleaved that are scattered or opened, and thus being mixed with ground vegetation. Sentinel-2 images

are also not capable to separate these classes when they are clustered at a resolution lower than the sensors capacity. Another significant confusion was obtained between Broadleaved and Shrubland (23 points), being the former characterized by commission error through almost all classes. Broadleaved is one of the most abundant classes and its omission error was very low, clearly benefiting from being the class with the largest training sampling units. Still, high commission error may result from a considerable portion of the Broadleaved LUCAS training points being surveyed in areas where trees are mixed within several types of LC. The omission error in Wetlands with Water was also large, with 20 points classified as Water and labelled as Wetland. This is also a common confusion due to the amount of time an area is inundated leading the classifier to oscillate between Water and Wetland with high uncertainty. Generally, most of the observed confusions between classes were expected, such as Natural herbaceous with Pastures and Shrubland, Broadleaved with Shrubland and Water with Wetlands.

Some studies have produced national scale LC maps with results that cover the continental Portugal area, using similar methodologies of supervised classification with remote sensing data. Therefore, a comparison between the OA of the results obtained in this experiment is carried out for two previous studies^{2,3}. The first one reaches an OA of 67% using 13 classes, being about 7% higher than our comparable result (60%, Level-2, 12 classes). However, in this study the training data is extracted from CLC and High Resolution Layers (HRL) from Copernicus, requiring more complex and time-consuming processing. Moreover, validation is based on CLC samples (not entirely independent) which were extracted from only one tile located in Portugal. Our validation dataset covers the entire study area and it was based on manual labelling, being totally independent from training data. The second study reaches 44.3% of OA using only 11 classes, thus being about 16% lower than the 60% obtained in our experiment. This study is more directly comparable since training data is also obtained from LUCAS. Similar accuracy results (59% OA) are only reported in this study when using additional training samples also extracted from CLC. However, validation here is also not independent, being performed with a small subset of corrected training samples.

4. CONCLUSIONS

This study presented a supervised classification experiment using training extracted uniquely from LUCAS 2018 dataset for land cover classification at the national scale. Multi-temporal Sentinel-2 remote sensing imagery during one year was also used in classification together with the machine learning RF algorithm. SMOTE oversampling technique was applied to reduce the unbalance in LUCAS training classes. An independent validation dataset with 600 points was created and manually labeled through orthophotos. OA of 57% for Level-2 and 72% for Level-1 were obtained. Using SMOTE improved the OA accuracy to 60% and 74%, Level-2 and Level-1 respectively. Results of OA are consistent with previous results reported by other studies in the same study area, being superior by 16% when using uniquely LUCAS data for training. The processing filters applied to LUCAS data have made the data well-suited for national scale LC classifications. The application of SMOTE also showed small but considerable improvements in the classification results, demonstrating that the use of oversampling techniques, in particular for small and unbalanced training datasets such as LUCAS, can be of great benefit. Confusion observed between classes was also in line with similar studies (e.g. Natural herbaceous with Pastures or Broadleaved with Shrubland). However, improvement in LUCAS data filtering concerning the trees and soil properties is needed to reduce confusion between vegetation classes. The exploration of oversampling and undersampling techniques to further balance training sampling units by class and therefore improve the classification accuracy, can also be an investigation topic for future work. The results of this experiment have demonstrated that LUCAS data can be used as a single training source for supervised classification at national scale, even in the fairly large and diverse landscape of continental Portugal.

ACKNOWLEDGMENTS

This work has been supported by project SCAPEFIRE (PCIF/MOS/0046/2017), project foRESTER (PCIF/SSI/0102/2017), and by Centro de Investigação em Gestão de Informação (MagIC), all funded by the Portuguese Foundation for Science and Technology (FCT). Value-added data processed by CNES for the Theia data centre www.theia-land.fr using Copernicus products. The satellite image pre-processing uses algorithms developed by Theia's Scientific Expertise Centers.

REFERENCES

- [1] Weigand, M., Staab, J., Wurm, M. and Taubenböck, H., “Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data”, *International Journal of Applied Earth Observation and Geoinformation*, 88, 102065 (2020).
- [2] Malinowski, R., Lewiński, S., Rybicki, M., Gromny, E., Jenerowicz, M., Krupiński, M., Nowakowski A., Wojtkowski C., Krupiński M., Krätzschmar E. and Schauer, P., “Automated production of a land cover/use map of Europe based on Sentinel-2 imagery”, *Remote Sensing*, 12(21), 3523 (2020).
- [3] Leinenkugel, P., Deck, R., Huth, J., Ottinger, M. and Mack, B., “The potential of open geodata for automated large-scale land use and land cover classification”, *Remote Sensing*, 11(19), 2249 (2019).
- [4] Mack, B., Leinenkugel, P., Kuenzer, C. and Dech, S., “A semi-automated approach for the generation of a new land use and land cover product for Germany based on Landsat time-series and Lucas in-situ data”, *Remote Sensing Letters*, 8(3), 244-253 (2017).
- [5] Close, O., Benjamin, B., Petit, S., Fripiat, X. and Hallot, E., “Use of Sentinel-2 and LUCAS database for the inventory of land use, land use change, and forestry in Wallonia, Belgium”, *Land*, 7(4), 154 (2018).
- [6] Griffiths, P., Nendel, C. and Hostert, P., “Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping”, *Remote sensing of environment*, 220, 135-151 (2019).
- [7] Pflugmacher, D., Rabe, A., Peters, M. and Hostert, P., “Mapping pan-European land cover using Landsat spectral-temporal metrics and the European LUCAS survey”, *Remote sensing of environment*, 221, 583-595 (2019).
- [8] Mellor, A., Boukir, S., Haywood, A. and Jones, S., “Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin”, *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, 155-168 (2015).
- [9] Fernández-Navarro, F., Hervás-Martínez, C. and Gutiérrez, P. A., “A dynamic over-sampling procedure based on sensitivity for multi-class problems”, *Pattern Recognition*, 44(8), 1821-1833 (2011).
- [10] Gómez, C., White, J. C. and Wulder, M. A., “Optical remotely sensed time series data for land cover classification: A review”, *ISPRS Journal of Photogrammetry and Remote Sensing*, 116, 55-72 (2016).
- [11] Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., and Rodes, I., “Operational high resolution land cover map production at the country scale using satellite image time series”, *Remote Sensing*, 9(1), 95 (2017).
- [12] EUROSTAT, “LUCAS 2018 (Land Use/Cover Area Frame Survey) – Technical reference document C3 Classification (Land Use and Land Cover)”, Technical Report, Eurostat, European Commission (2018), <https://ec.europa.eu/eurostat/documents/205002/8072634/LUCAS2018-C3-Classification.pdf>
- [13] Hernandez, I., Benevides, P., Costa, H. and Caetano, M., “Exploring Sentinel-2 for Land Cover and Crop Mapping in Portugal. The International Archives of Photogrammetry”, *Remote Sensing and Spatial Information Sciences*, 43, 83-89 (2020).
- [14] Moraes D., Benevides P., Costa H., Moreira F.D., Caetano M., "Influence of Sample Size in Land Cover Classification Accuracy using Random Forest and Sentinel-2 Data in Portugal", *International Geoscience and Remote Sensing Symposium IGARSS 2021*, (2021) (In Press).
- [15] Costa H., Machado I., Moreira F.D., Benevides P., Moraes D., Caetano M., "Exploring the Potential of Sentinel-2 Data for Tree Crown Mapping in Oak Agro-Forestry Systems", *International Geoscience and Remote Sensing Symposium IGARSS 2021*, (2021) (In Press).
- [16] Benevides P., Costa H., Moreira F.D., Moraes D., Caetano M., "Annual Crop Classification Experiments in Portugal using Sentinel-2", *International Geoscience and Remote Sensing Symposium IGARSS 2021*, (2021) (In Press).
- [17] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, Blondel, M., Prettenhofer P., Weiss, R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M. and Duchesnay E., “Scikit-learn: Machine learning in Python”, *the Journal of machine Learning research*, 12, 2825-2830 (2011).