Land-Cover Mapping of Agricultural Areas Using Machine Learning in Google Earth Engine

Florencia Hastings[1,2(✉)], Ignacio Fuentes[3], Mario Perez-Bidegain[1], Rafael Navas[4], and Angela Gorgoglione[5]

[1] School of Agronomy, Universidad de la República, Av. Gral. Eugenio Garzón 780, Montevideo, Uruguay
fhastings@mgap.gub.uy
[2] Directorate of Natural Resources, Ministry of Agriculture, Livestock and Fisheries, Av. Gral. Eugenio Garzón 456, Montevideo, Uruguay
[3] School of Life and Environmental Sciences, University of Sydney, Sydney, NSW 2006, Australia
[4] Programa Nacional de Investigación en Producción y Sustentabilidad Ambiental, Instituto Nacional de Investigación Agropecuaria, Montevideo, Uruguay
[5] School of Engineering, Universidad de la República, Julio Herrera y Reissig 565, Montevideo, Uruguay

Abstract. Land-cover mapping is critically needed in land-use planning and policy making. Compared to other techniques, Google Earth Engine (GEE) offers a free cloud of satellite information and high computation capabilities. In this context, this article examines machine learning with GEE for land-cover mapping. For this purpose, a five-phase procedure is applied: (1) imagery selection and pre-processing, (2) selection of the classes and training samples, (3) classification process, (4) post-classification, and (5) validation. The study region is located in the San Salvador basin (Uruguay), which is under agricultural intensification. As a result, the 1990 land-cover map of the San Salvador basin is produced. The new map shows good agreements with past agriculture census and reveals the transformation of grassland to cropland in the period 1990–2018.

Keywords: Land-cover map · Supervised classification · Google earth engine · Agricultural region

1 Introduction

Land-use/land-cover (LULC) change plays a critical role in the global change study. Often, land cover is the result of human actions to modify the natural environment into a "customized environment". Such customization takes place to increase agriculture production, satisfy domestic needs, or restore ecosystem services. In the last decade, several studies have demonstrated that environmental problems are often related to LULC change [1,2]. LULC and human/natural

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12252, pp. 721–736, 2020.
https://doi.org/10.1007/978-3-030-58811-3_52

modification have impacted negatively on the degradation of water quality, loss of biodiversity, deforestation, and global warming [3]. Therefore, characterizing and mapping land cover is essential for planning and managing natural resources, including agricultural fields [4]. Furthermore, LULC is a major data source of hydrologic models [5,6]. For this purpose, the implementation of effective land-cover mapping requires advanced remote-sensing methodologies able to provide accurate, inexpensive, and on-demand land-cover products using free cloud-based data processing platforms and free and open-access information.

Based on these considerations, it is essential to process big earth data (e.g., satellite images) in the classification procedure over a large watershed. Numerous platforms, such as Google Earth Engine (GEE), Amazon's Web Services, Microsoft's Azure, and National Aeronautics and Space Administration (NASA) Earth Exchange have been created to tackle this issue and support processing and analysis of big earth data [7,8]. In particular, GEE is an application that processes big geospatial data and classifies land cover over vast areas [9]. In this platform, the open-source satellite images can be efficiently imported and processed in the cloud without the need for downloading the data to local computers. Furthermore, many remote-sensing algorithms (e.g., classification algorithms and cloud masking techniques) and several image-driven products are available in this platform and are included and editable in user-defined algorithms [9,10]. To date, numerous studies about classifications using GEE in large areas have been conducted. Belgiu et al. investigated the ability of a machine learning technique in land-cover mapping in different agro-ecological areas of the world [11]. Many efforts with remote sensing have been carried out to overcome the challenges of producing less costly (or free) and more time-efficient land-cover mapping around the world [12]. Project MapBiomas is a good example; it shows how collections of LULC maps could be integrated into a web platform to aid planning and managing natural resources in Brazil [13].

Based on these considerations, this study explores a methodology that uses a cloud-based free open-source database to contribute to current land-cover mapping efforts. Particularly in Uruguay, only LULC maps for the years 2000, 2008, 2011, 2015 [14] and 2018 [15] are available (online). The primary objective is to explore machine learning in GEE and its accuracy for historical land-cover mapping of areas mainly characterized by agricultural land use. The specific objectives of this study are: i) evaluate the utilization of GEE for feature extraction, with its advantages and disadvantages; ii) assess the performance of supervised machine-learning techniques to classify the land surface; iii) obtain a historical land-cover map for an agricultural watershed by coupling GEE and machine learning methods. This methodology was applied to the San Salvador watershed, an agricultural basin located in Uruguay (South America).

2 Materials

This section first presents a description of the study area, and then the data used to develop the land-cover map. The data used includes satellite data and field

data from the General Census of Agriculture (GCA). The phenological data of the surveyed crops is also presented.

2.1 Study Area

The San Salvador basin is located in Uruguay, in the Soriano department, and it covers an area of 2,390 390 km^2 (Fig. 1).

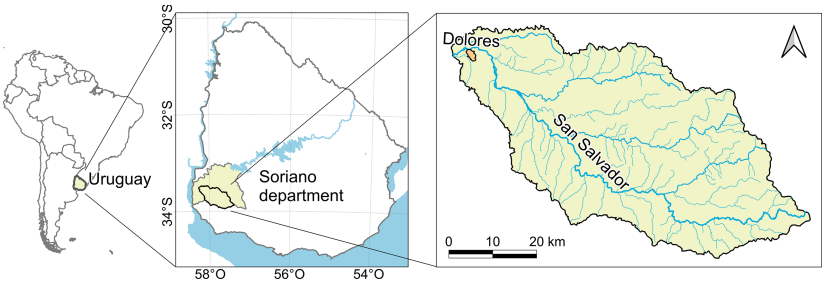

Fig. 1. Location of San Salvador basin (Soriano department, Uruguay). Coordinate reference system: world geodetic system 1984 (WGS84).

Uruguay has a humid subtropical climate. According to the Köppen climate classification, it is Cfa, C = temperate, moderate, rainy, f = fully humid, a = mean temperature in the warmest month is 22 °C or higher [16]. The mean total annual precipitation is 1100 mm and the temperature can range between –7.9 °C and 40.4 °C [16]. The region is characterized by a landscape of smooth hills and, the study area, has an average slope equal to 2.3%.

The San Salvador basin is located in the west part of the country where the most suitable soils for agriculture are [17]. In the decade of 1990, the agriculture practices included crop-pasture rotation, associated with livestock, with tillage practices[18]. Since the 21st century, agriculture suffered important changes due the incorporation of no-tillage practices, transgenic, continuous cropping, and economic freedom [18]. In a short period, soybean become the main grain crop, increasing the planted area from 12,000 ha in 2000/2001 to 1,300,000 ha in 2014/2015, total cropland extension were 426,000 ha and 1,500,000 ha respectively [19]. Also, production forest had an important increase from 186,277 ha planted in 1990 [20] to 1,000,190 ha in 2018 [21] because of the development in the forest-wood sector. Land-cover map of 2018 shows that 68% of Uruguay surface is covered with native grassland associated with livestock and 18% with croplands [15]. It is worth to mention that the 8.4% of the national gross domestic product (GDP) (2018) is from agricultural products and associated industries [22].

The above-mentioned increase in cropland also took place in the San Salvador basin. The land-cover map of 2018 shows that 62.4% of the basin is covered by cropland. Figure 2 and Table 1 show the basin land-cover of 2018.

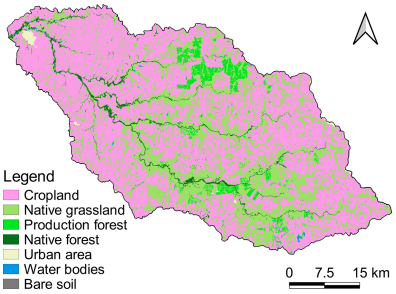

Fig. 2. Land-cover map of San Salvador basin in 2018, classification of Sentinel-2 Level 2A scenes (resolution 10m/pixel) [15].

Table 1. Area of land-cover classes in 2018 in the San Salvador basin.

Class	Area (ha)	%
Cropland	149,255	62.4%
Native grassland	74,845	31.3%
Production forest	7,943	3.3%
Native forest	5,691	2.4%
Urban area	927	0.4%
Water bodies	361	0.2%
Bare soil	27	0.0%
Total	239,047	100.0%

Taking into account the strategic importance of San Salvador watershead for the economy of the country, it is fundamental to have accurate historical land-cover mapping of this area to detect any land-cover change that can affect the agricultural production. Furthermore, it is worth mentioning the importance of the distribution of croplands in 1990 to study the land-cover change in this area. Based on these considerations, we selected San Salvador and the year 1990 as our case study.

2.2 Satellite and Field Data

Surface Reflectance Tier 1 dataset, which represents the atmospherically corrected surface reflectance from the Landsat 5 Enhanced Thematic Mapper (ETM) sensor, was used in this study. Landsat 5 was operational in the '90s (launched in 1984 and decommissioned in 2013) and its images are freely available. The satellite had a 16-day repeat cycle and a spatial resolution of 30 m.

The field data available for 1990 is the database from the General Census of Agriculture (GCA) that provides aggregated data to validate the results of the imagery classification. Furthermore, the phenological data of the surveyed crops were used for the visual interpretation of the satellite images.

General Census of Agriculture

In Uruguay, the GCA is conducted by the Ministry of Agriculture, Livestock and Fisheries (MGAP) every ten years. Census data is reported, to preserve people's privacy, in aggregated geographical units called census tracts (CTs).

In 1990, for the Soriano department, the census registered an area dedicated to winter crops six times larger than the one dedicated to summer crops. The main crops are wheat in winter and sorghum in summer. Among winter crops, wheat covers 72% of the planted area, barley covers 21%, and oat (grain) 7%. The main winter forage crop is oat (sowed pasture), 80% of the oat area is for

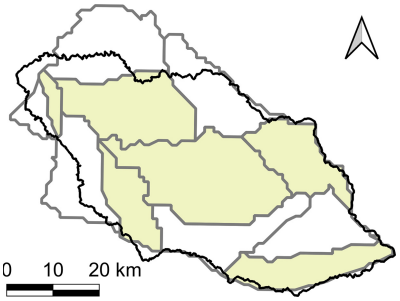

Fig. 3. Selected census tracts of the General Census of Agriculture 1990.

Table 2. Land-cover area of the selected census tracts of the General Census of Agriculture 1990.

Class	Area (ha)	%
Native grassland	73,525	51.7%
Cropland	34,563	24.3%
Sowed pasture	30,856	21.7%
Native forest	2,490	1.7%
Production forest	923	0.6%
Total	142,357	100.0%

grazing (sowed pasture) and 20% is for grain. In summer sorghum, corn, and soybean cover 54% , 33% and 12% of the planted area respectively.

The data from GCA disaggregated per CTs was obtained from the GCA database facilitated by agricultural statistics of MGAP. The database showed that some CTs had under-coverage and over-coverage of the total area. For this reason, only were considered those CTs that had a difference smaller than ± 15% in the coverage-area. So, even though the San Salvador watershed has 12 CTs in common with the Soriano department, only 6 CTs were taken into account. The selected CTs and the land-cover area obtained from GCA is presented in Fig. 3 and Table 2.

Phenological Data

The phenological data used was taken from two sources. For wheat, barley, and oat, phenological data was obtained from the reports on the evaluation of cultivars (results of 2003) [23]. In the case of sorghum, corn, and soybean, a phenologic model developed on the base of experimental data ran in 1990 [24] was considered. Table 3 presents the phenology of crops surveyed in the 1990 census. The phenological phases considered were the early growth stage (between emergence and flowering) and peak on leaf area index season (after flowering).

3 Methods

This section outlines the conceptualization of the methodology approach and the phases implemented to develop the land-cover map.

3.1 Conceptualization of the Methodology Approach

According to the major steps of image classification identified by Lu and Weng [25], in this study, five phases were implemented: (1) imagery selection and preprocessing, (2) selection of the classes and training samples, (3) classification

Table 3. Average dates of phenological phases. *Average date of harvest. **Average date of the last cut.

Crop	Sowing	Emergence	Flowering	Maturity
Wheat [23]	13-jun	23-jun	22-oct	12-dic*
Barley [23]	10-jul	19-jul	10-oct	2-dic*
Oat [23]	12-mar	18-mar		17-nov**
Sorghum [24]	15-oct	26-oct	9-ene	13-feb
Corn [24]	7-oct	17-oct	24-dic	24-feb
Soybean [24]	21-oct	2-nov	31-dic	16-abr

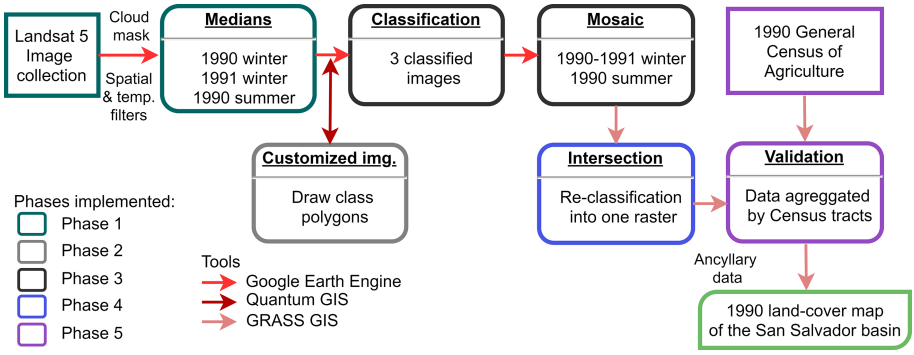

Fig. 4. Conceptualization of the methodology approach.

process, (4) post-classification process, and (5) validation. An overall flowchart of the methodology approach showing each phase is presented in Fig. 4.

GEE is a cloud-based platform that includes a wide geospatial dataset, as well as the entire Landsat archive, and algorithms that facilitate spatial analysis [10]. Image selection and supervised classification techniques were implemented in GEE which was coupled with Geographic Information Systems (GIS) environment. Quantum GIS 3.4.15 (QGIS) was used to customize images and create the polygons classes, and GRASS GIS 7.4.2 was adopted for the post-classification phase.

3.2 Phase 1: Imagery Selection and Pre-processing

The satellite images were collected between December 1989 and August 1991. The criterium adopted for imagery selection was to cover completely the study area considering the crop peak growth phenological stage of winter and summer crops. According to the wheat phenology, the main winter crop, winter peak growing season ranges from October 22nd to December 12th. Sorghum and corn were the main crops for summer season, so, in average, peak growing season

ranges from December 15th to February 15th. Furthermore, a quality filter was applied to mask pixels with clouds, cloud shadows or high reflectance (reflectance greater than 6000). Considering the scarcity of quality scenes that covers the entire study area, scenes acquired in different dates were used. To minimize the phenology variation among the scenes, a period smaller than a month was chosen within each growing season. To perform the classification, the images selected were reduced to one per season by the median. Thus, when images had overlapping pixels, the median was computed independently in each band.

3.3 Phase 2: Selection of Classes and Training Samples

The supervised classification approach uses the training samples signatures to classify the image and define the land-cover categories [25]. A successful classification depends on a proper class selection and its representative sampling [25]. Considering that the classes are defined by visual interpretation of satellite images, discernible land-cover units were selected as native grasslands, croplands, and water. Cropland class was also disaggregated into a hierarchical system of subclasses by the post-classification process. Three spectral bands (red, green and near infrared) of surface reflectance were used for the class-visual interpretation.

Some other classes as urban area, production forest, and native forest were considered. These classes were not included in the classification because of mixed pixel problems in remote sensing technology within non-cropland [26]. Instead, some ancillary data were used. For urban area, from the land-cover map of Uruguay of 2000 was used [14]. For production forest, data were mapped from Landsat imagery as there were only four polygons of this class. For native forest, data from the national forest inventory (2018) [21] was used. A visual comparison between the data of the national forest inventory (2018) and aerial photos, taken in 1966 by the Military Geographic Service [27] was made, and no substantial difference was observed. Furthermore, native forest represents about 2% of the study area according to 1990 census data, so it does not represent a significant contribution to the land-cover map.

3.4 Phase 3: Classification Process

A supervised classification approach was performed using a Classification and Regression Tree (CART) classifier implemented in GEE. CART is a binary decision tree. It uses a sample of training data for which the correct classification is known. It recursively splits the data based on a statistical test ending in terminal nodes associated with entities correctly classified [28,29]. One advantage of selecting non-parametric classifiers, like decision trees, is that they do not suppose a normal distribution of the dataset, which is difficult to achieve especially in complex landscapes [25]. Furthermore, visual interpretation-sampling data were considered.

The polygons with data classes of 1990 winter and 1990 summer were randomly split keeping on average 70% of the polygons for training and 30% for

validating the classifier. Also, were consider polygons with data classes of 1991 winter and 1991 summer to perform a cross-year validation. As a result, two classified images were obtained representing 1990 winter and 1990 summer. All spectral bands were used as inputs to classification, four visible and near-infrared (VNIR) bands, two short-wave infrared (SWIR) bands processed to orthorectified surface reflectance, and one thermal infrared (TIR) band processed to orthorectified brightness temperature.

An error matrix was employed to assess the accuracy of the classification and cross-year validation. The overall accuracy (OA), Kappa coefficient (K), consumer's accuracy (CA), and producer's accuracy (PA) coefficients were calculated from the error matrix. OA is the rate of the total correctly classified field samples [30]. K is a measure of overall statistical agreement of an error matrix, which takes into account non-diagonal elements [30]. CA is the conditional probability that the pixels classified as category i was assigned as category i by the reference data [30]. PA is the conditional probability that the pixels assigned as category j by the reference data are classified as category j [30].

3.5 Phases 4 and 5: Post-classification Process and Validation

The post-classification process consisted of observing the sequence winter-summer land cover and making a re-assignation of classes resulting in the final land-cover map. For that, the two classified images were imported as a raster image into GRASS GIS. Then, the two raster images were intersected, creating a new cross-raster map representing all unique combinations of classes and its associated cross-matrix. Finally, the classes were re-assigned.

After the classes re-assignation, a comparison to census data was made to evaluate the accuracy of the resulting map. The map validation is performed since training and validating data were not surveyed in the field [25].

4 Results and Discussion

In this section, the results of each methodological phase to develop the land-cover map are presented. Furthermore, a discussion of the main findings of the study reported in this article is included.

4.1 Phase 1 Results: Imagery Selection and Pre-processing

To create a cloud-free image of 1990, winter scenes corresponding to the early growth stage of winter crops were selected due to the lack of scenes during the phenological phase of peak growth of winter crops. However, due to the quality filters (to mask pixels with clouds, cloud shadows, or high reflectance) applied to the images, the coverage of the studied area was 83% . Thus, to complete the coverage of the entire study area, images of 1991 winter were used. For summer period, scenes during the phenological phase of the peak growth of summer crops were selected. Also, images of 1991 winter and 1991 summer were selected

Table 4. Scenes used from Landsat 5 (Enhanced thematic mapper sensor, surface reflectance Tier 1 dataset). *Scene ID refers to path, row, and acquisition date from the complete ID.

	ID*	Cloud cover	Acquisition date
1990 winter	224083_19900804	0%	8/4/1990
	224084_19900804	0%	8/4/1990
	225083_19900811	28%	8/11/1990
	225084_19900811	41%	8/11/1990
	224083_19900820	0%	8/20/1990
	224084_19900820	0%	8/20/1990
	225084_19900827	58%	8/27/1990
1991 winter	224084_19910807	96%	8/7/1991
	225083_19910814	0%	8/14/1991
	225084_19910814	0%	8/14/1991
1990 summer	224083_19891223	1%	12/23/1989
	224084_19891223	0%	12/23/1989
	225083_19891230	7%	12/30/1989
	225084_19891230	27%	12/30/1989
	225084_19900115	4%	1/15/1990
	225083_19900115	0%	1/15/1990
1991 summer	224083_19910111	10%	1/11/1991
	224084_19910111	12%	1/11/1991
	225083_19910102	5%	1/2/1991
	225084_19910102	3%	1/2/1991

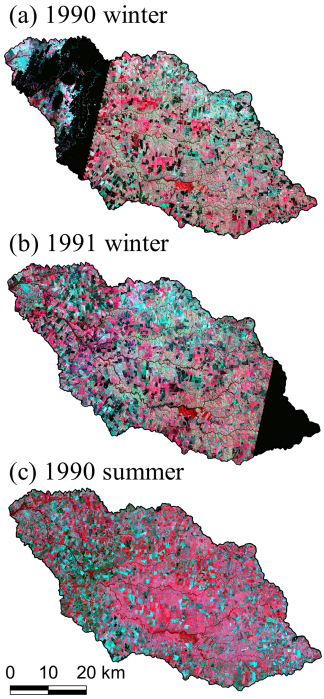

Fig. 5. Landsat median images of 1990 winter, 1991 winter and 1990 summer (false colour composite Bands 4, 3, 2).

to perform a cross-year validation. A summary of the considered scenes for each season, its cloud coverage, and acquisition date is presented in Table 4. In Fig. 5, the selected scenes, reduced by the median, representing (a) 1990 winter, (b) 1991 winter, and (c) 1990 summer are shown.

4.2 Phase 2 Results: Selection of Classes and Training Samples

A first search was made in USGS Landsatlook to recognize how crops looked during August and define the land-cover units [31]. A 1989 sequence of images showing the different phenologic phases of winter crops was found. Wheat and barley (winter crops) are discern in May image (Fig. 6a) respectively highlighted in red and yellow but not distinguished in the image from August onward (Figs. 6b–d). Figure 6c shows that winter crops achieved peak growth and Fig. 6d shows they were harvested. From August image (Fig. 6a) oat was recognized, highlighted in green, this enable to assign a subclass of winter crops to oat. Native grasslands were recognized and showed in black with no significant variation observed into the four images.



(a) Date 5/20/1989. (b) Date 8/1/1989. (c) Date 10/27/1989. (d) Date 11/28/1989.

Fig. 6. Landsat 5 images showing the evolution of winter crops (red circle: wheat, yellow square: barley, green square: sowed pasture, black circle: grassland). (Color figure online)

Table 5. Land-cover classification classes.

1st level	2nd level	Descriptions
Water bodies	WATB	Streams, rivers, lakes and reservoirs
Grassland	GRAS	Native grasslands or improved grasslands
	FLDP	Native grasslands placed in floodplains
Cropland	PAST, PAST2	Sowed pasture in two phenological phases (only for winter images)
	CRPw	Winter cropland. Includes: wheat and barley
	CRPs	Summer cropland. Includes: sorghum, corn and soybean
	FLLW	Summer fallow
	BARE	Bare or tilled soil

In this study, ten classes were defined to perform the classification. For floodplains (FLDP) a new class was added, as it was observed that grassland kept misclassified as cropland. A list of the defined classes with their description is presented in Table 5.

4.3 Phase 3 Results: Classification Process

For each image, an average of 20 polygons were defined per class, and area used to train and validate the classifiers. This resulted in an average of 1500 ha for training, 500 ha for validation, and 1000 ha for cross-year validation (6). The supervised classification was applied to the three median images described in Sect. 4.1. Finally, in the case of winter images, a mosaic was made to couple the missing area of the winter image of 1990 with data of the winter image of 1991.

The statistical accuracy assessment was applied to train and validate datasets of each image. Table 6 presents the results of statistics OA, K, CA and PA calculated. For the training, validating and cross-year validating dataset, OA and K vary in a range of 0.96 to 1.00 and 0.95 to 1.00 respectively. These results indicate that the three classified images achieved high overall accuracies. Furthermore, CA and PA coefficients show the performance per class. In general, PA and CA vary in a range of 0.83 to 1.00. However, floodplain class had low accuracy, PA and CA vary in a range of 0.44 to 1.00. As mentioned before, it

was detected by visual inspection that grassland was misclassified as cropland by the classifier in this area. The post-classification process allowed separating cropland and grassland from the floodplain class.

Table 6. Classification performance and area used for training and validation. CA: consumer's accuracy, PA: producer's accuracy, OA: overall accuracy, K: Kappa.

	Winter						Summer					
	1990				1991		1990				1991	
	Training		Validation		Validation		Training		Validation		Validation	
	CA	PA	CA	PA	CA	PA	CA	PA	CA	PA	CA	PA
Water bodies	1.00	1.00	1.00	1.00	—	—	1.00	0.99	0.83	0.99		
Native grassland	1.00	1.00	0.99	0.98	0.99	0.95	1.00	0.97	0.99	0.97	0.99	1.00
Floodplain	1.00	0.98	1.00	0.95	0.62	0.97	0.98	0.90	0.44	0.90	1.00	0.94
Sowed pasture	1.00	1.00	0.99	1.00	0.99	0.87	—	—	—	—	—	—
Sowed pasture2	1.00	1.00	1.00	1.00	1.00	0.99	—	—	—	—	—	—
Winter cropland	1.00	1.00	1.00	0.99	0.96	1.00	—	—	—	—	—	—
Fallow	—	—	—	—	—	—	1.00	0.99	1.00	0.99	1.00	1.00
Summer cropland	—	—	—	—	—	—	1.00	0.99	1.00	0.99	1.00	0.99
Bare soil	1.00	1.00	0.97	1.00	0.96	0.98	1.00	1.00	1.00	1.00	1.00	1.00
OA	1.00		0.99		0.96		1.00		0.98		1.00	
K	1.00		0.99		0.95		1.00		0.98		0.99	
Area (ha)	1548		454		1,016		1,411		615		1,066	

4.4 Phase 4 Results: Post-classification Process

The seasonal variation of each class was taken into account to better distinguish the class of interest and do the class re-assignation. The intersection of the two classified images (winter and summer) resulted in a cross-raster map and its associated cross-matrix. The five classes considered to construct the final map are water bodies, native grassland, sowed pasture, winter croplands, summer croplands and double cropping lands.

In Table 7, the cross-matrix shows the area (ha) distribution of the intersection between the classified winter image and the classified summer image. Each row represents the area of winter classes desegregated into summer classified classes. Similarly, each column represents the area classified on summer image desegregated into winter classified classes.

The cross-matrix, presented in Table 7, shows that 84% of the area classified as GRAS on winter image stayed as GRAS on the final map, and the remaining 16% was assingned to a cropland class. Furthermore, 76% of the area classified as FLDP on winter was re-assigned to the class GRAS and the remaining 24% was re-allocated to a cropland class (CRPw). This was the main reason for defining the specific FLDP class, as discussed in Sect. 4.3.

Table 7. Cross-matrix expressing the area (ha) distribution of the intersection of the classified winter image with the classified summer image.

		Area (ha) classified on summer					
	Class	Native grassland	Flood-plain	Fallow	Summer cropland	Bare soil	Total
Area (ha) classified on winter	Native grassland	63,942	25,435	6,772	3,835	6,565	106,549
	Floodplain	12,228	29,703	9,279	5,001	5,147	61,358
	Sowed pasture	1,814	5,208	3,033	678	1,843	12,576
	Sowed pasture2	2,625	5,613	2,594	1,846	1,917	14,595
	Winter cropland	2,703	3,384	5,077	706	7,033	18,903
	Bare soil	7,203	5,174	4,634	1,405	5,551	23,967
	Total	90,515	74,517	31,389	13,471	28,056	237,948

4.5 Phase 5 Results: Validation

A comparison between 1990 GCA and the new land-cover map was made to validate the results of the final land-cover map. GCA data was grouped into classes of interest. So, the total area surveyed in GCA as vegetables, row crops, tilled soil, and fallow was compared with the total area of CRPw, CRPs and CRPWs classes; native forest was compared with MONT class area; production forest surface was compared with the FRST surface; the total area covered by pasture and annual forages was compared with the PAST area; and the total area dedicated to improved grassland and native grassland was compared with the GRAS area. The area of the six CTs selected was aggregated to reduce the geographic resolution errors and compared with the same extension of the land-cover map.

The area of the land-cover map classes is very similar to surface reported in the GCA. In Table 8 is presented the areas per class obtained from the land-cover map and from the GCA, and the difference between the two areas. In general, the area of the new-defined categories is greater than the GCA area due to the difference between the surveyed and the measured area that is equal to 6,574 ha. This difference is not valid for pasture and production forest classes. The challenge in classifying pasture class is represented by the fact that perennial pastures can be confused with grasslands. Furthermore, dairy production land use occupies an area where there are mixed portions of pasture and crop that are difficult to separate. About production forest area, an error in the CTs borders may explain the misclassified area since an area of production forest was identified next to the CTs borders.

Figure 7 shows the final land-cover map of the San Salvador basin in 1990 and, Table 9 presents the area covered by each class. The three main land covers identified in the basin were native grasslands (GRAS) with a surface equal to 56.8%, croplands (CRPw, CRPws and CRPs) with a total area of 30.0%, and sowed pastures with a surface of 9.8%.

As expected, the results showed a significant land-cover change between 1990 and 2018. Cropland surface increased from 95,258 ha to 149,255 ha (57%) and grassland reduced from 135,826 ha to 74,845 ha (45%). Also, the production forest area increased from 1,185 ha to 7,943 ha (570%). The observed trends are

Table 8. Comparison of the areas per class of the land-cover map and the General Census of Agriculture 1990 (GCA).

Class	Classified (ha)	%	GCA (ha)	%	Difference (ha)
Native grassland	84,097	57%	73,525	52%	10,572
Native forest	3,520	2%	2,490	2%	1,030
Cropland	43,805	29%	34,563	24%	9,242
Sowed pasture	17,509	12%	30,856	21%	–13,347
Production forest	0	0%	923	1%	–923
Total	148,931	100%	142,357	100%	6,574

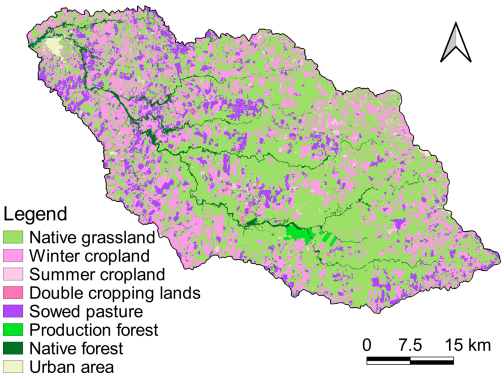

Fig. 7. Land-cover map of San Salvador basin in 1990.

Table 9. Area of land-cover classes of San Salvador basin in 1990.

Class	Area (ha)	%
Native grassland	135,826	56.8%
Winter cropland	66,830	28.0%
Summer cropland	4,314	1.8%
Double cropping	662	0.3%
Sowed pasture	23,452	9.8%
Native forest	5,683	2.4%
Production forest	1,185	0.5%
Urban area	991	0.4%
Water bodies	127	0.0%
Total	239,070	100.0%

related to agriculture and forestry expansion since 1990. Uruguay met a challenge for sustainable intensification as committed, in 2015, in the 21st Session of the Conference of the Parties (COP21) [32]. Since 2013, a public policy towards soil conservation, based on the Universal Soil Loss Equation (USLE), is being implemented [32]. This policy, designed by the MGAP, is recognized as a good example by the Food and Agriculture Organization (FAO) [33].

The main source of error identified in this study might be related to the scarcity of scenes on the dates and area of interest. For instance, the winter crop season was characterized only based on the available images in August, when winter crops were on an early growth stage. Another source of error may be introduced by the training and validation polygons created by the visual interpretation of images and its sampling design. The low accuracy achieved on the floodplain class may occur due to: (1) reflectance similarities between vegetation on floodplains and crops, especially in the summer season, and (2) the heterogeneity of the vegetation on the surface this class represents. However, the low accuracy in the floodplain class was improved, based on visual validation, by using a multi-seasonal post-classification approach.

The methodology of this study can be used to develop a series of historical annual land-cover maps and might be applied for mapping other basins with similar surface features. Landsat scenes are freely available since 1984. Yet, their frequency increase since 1987–1988, so temporal series from this period to the present can be achieved. For validation, data from GCA is available every ten years and data of agricultural annual surveys (DIEA-MGAP) can also be employed. Furthermore, since 2000 onwards other satellite missions, such as MODIS (2000), CEBERS-4 (2014) and Sentinel (2014), provide free satellite imagery. Therefore, the presented methodology might be improved using a multi-satellite approach.

5 Conclusions

This study presented an efficient approach to estimate the land-cover of an agricultural basin, using free open-source tools. GEE showed excellent potential in land cover mapping with high processing efficiency. Machine-learning algorithms developed in GEE are easily applied and a wide range of free satellite imagery is available.

As a case study, we estimated the land-cover in the San Salvador basin in 1990 based on Landsat 5 imagery and GEE processing platform. High accuracy was achieved in the supervised classification according to the error matrix analysis. Additionally, map validation showed good agreements with GCA data. The methodology presented was capable to deal with the scarcity of historical quality scenes.

The methodology presented in this article can be improved by a multi-satellite approach. Temporal map series of the same watershed can be developed as well as other catchments. The above mentioned are issues that can be explored in future research.

Map availability. The San Salvador basin land-cover map of 1990 can be freely downloaded from https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/politicas-y-gestion/mapa-cobertura-utilizando-google-earth-engine-cuenca-del-rio-san-salvador-1990.

Funding. This research was supported by the National Research and Innovation Agency (ANII) under contract number FSA_PI_2018_1_148628.

Acknowledgements. We thank DIEA-MGAP that kindly made available the GCA data and DGRN-MGAP for their comments and suggestions.

References

1. Giri, S., Qiu, Z.: Understanding the relationship of land uses and water quality in Twenty First Century: a review. J. Environ. Manag. **173**, 41–48 (2016)
2. Rodríguez, J., Rico, A., Mendoza-Martínez, E., Gómez-Ruiz, A., Sedeño-Diaz, J., López-López, E.: Impact of changes of land use on water quality, from tropical forest to anthropogenic occupation: a multivariate approach. Water **10**, 1–16 (2018)
3. Dwivedi, R., Kandrika, S., Ramana, K.: Land-use/land-cover change analysis in part of Ethiopia using landsat thematic mapper data. Int. J. Remote Sens. **26**, 1285–1287 (2005)
4. Gomez, C., White, J., Wulder, M.: Optical remotely sensed time series data for land cover classification: a review. ISPRS J. Photogramm. **116**, 55–72 (2016)
5. Figorito, B., Tarantino, E., Balacco, G., Fratino, U.: An object-based method for mapping ephemeral river areas from WorldView-2 satellite data. Proc. SPIE **8531**, 85310B (2012)
6. Aquilino, M., Tarantino, E., Fratino, U.: Multi-temporal land use analysis of an ephemeral river area using an artificial neural network approach on Landsat imagery. ISPRS - Int. Arch. Photogramm. **XL–5/W3**, 167–173 (2013)
7. Amani, M., et al.: A generalized supervised classification scheme to produce provincial wetland inventory maps: an application of Google Earth Engine for big geo data processing. Big Earth Data **3**(4), 378–394 (2019)
8. Xiong, J., et al.: Automated cropland mapping of continental Africa using Google earth engine cloud computing. ISPRS J. Photogramm. **126**, 225–244 (2017)
9. Kumar, L., Mutanga, O.: Google earth engine applications since inception: usage, trends, and potential. Remote Sens. **10**, 1509 (2018)
10. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R.: Google earth engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. **202**, 18–27 (2017)
11. Belgiu, M., Csillik, O.: Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. Remote Sens. Environ. **204**, 509–523 (2018)
12. Mumby, P., Green, E., Edwards, A., Clark, C.: The cost-effectiveness of remote sensing for tropical coastal resources assessment and management. J. Environ. Manag. **55**, 157–166 (1999)
13. MapBiomas Project - Collection v4.0 of the Annual Land Use Land Cover Maps of Brazil. http://mapbiomas.org. Accessed 7 Feb 2020
14. National Directorate of Land-use Planning - Ministry of Housing, Land Planning and Environment (DINOT-MVOTMA): Land-cover map of Uruguay. Technical report, Montevideo, Uruguay (2014). (in Spanish)
15. Petraglia, C., Dell'Acqua, M., Pereira, G., Yussim, E.: Integrated Land Cover/Use Map of Uruguay of 2018. Gráfica Mosca, Montevideo (2019). (in Spanish)
16. Uruguay Meteorological Institute (INUMET). https://www.inumet.gub.uy/index.php/clima. Accessed 3 Jan 2020
17. Agricultural Statistics - Ministry of Agriculture, Livestock and Fisheries (DIEA-MGAP): Agricultural Regions of Uruguay. Technical report, Montevideo, Uruguay (2015). (in Spanish)
18. Arbeletche, P., Coppola, M., Paladino, C.: Analysis of agro-business as a form of business management in South America: the Uruguayan case. Agrociencia Uruguay **16**, 110–119 (2012). (in Spanish)

19. Couto, P.: Recent trends in rainfed agriculture. Plan Agropecuario **161**, 64–68 (2017). (in Spanish)
20. Agricultural Statistics - Ministry of Agriculture, Livestock and Fisheries (DIEA-MGAP): General Census of Agriculture of 1990. Editorial MGAP, Montevideo, Uruguay (1994). (in Spanish)
21. Forestry Directorate - Ministry of Agriculture, Livestock and Fisheries (DGF-MGAP): Results of the national forest mapping 2018. Technical report, Montevideo, Uruguay (2018). (in Spanish)
22. Agricultural Statistics - Ministry of Agriculture: Agricultural statistical yearbook 2019. Technical report, Montevideo, Uruguay (2019). (in Spanish)
23. Castro, M., Pereyra, S., Stewart, S., Germán, S., Vázquez, D.: Experimental results of the national cultivar evaluation. Technical report, National Institute of Agricultural Research (INIA) and National Institute of Seeds (INASE), Montevideo, Uruguay (2003). (in Spanish)
24. Fassio, A., et al.: Prediction of phenological states for soy, sunflower, corn, sorghum (grain, forage, sweet and silage). Technical report, National Institute of Agricultural Research (INIA), Montevideo, Uruguay (2014). (in Spanish)
25. Lu, D., Weng, Q.: A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. **28**, 823–870 (2007)
26. Liu, J., et al.: Spatial and temporal patterns of China's cropland during 1990–2000: an analysis based on landsat TM data. Remote Sens. Environ. **98**, 442–456 (2005)
27. Military Geographic Service (SGM): Aerial photographs of Uruguay 1966/67, scale 1 / 40,000. http://visualizador.sgm.gub.uy/gmaps/index.html. Accessed 3 Jan 2020. (in Spanish)
28. Bittencourt, H., Clarke, R.T.: Use of classification and regression trees (CART) to classify remotely-sensed digital images. In: IGARSS 2003, 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477), vol. 6, pp. 3751–3753 (2003)
29. Friedl, M., Brodley, C.: Decision tree classification of land cover from remotely sensed data. Remote Sens. Environ. **61**(3), 399–409 (1997)
30. Stehman, S.: Selecting and interpreting measures of thematic classification accuracy. Remote Sens. Environ. **62**, 77–89 (1997)
31. U.S. Geological Survey (USGS) Landsatlook. https://landsatlook.usgs.gov/viewer.html. Accessed 15 Jan 2020
32. Ministry of Agriculture: Livestock and Fisheries (MGAP): Uruguay agribusiness. The challenges for sustainable development. Technical report, Montevideo, Uruguay (2017). (in Spanish)
33. Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils: Status of the World's soil resources - main report. Technical report, Rome, Italy (2015)