

A Semi-Supervised Hybrid Approach for Multitemporal Multi-Region Multisensor Landsat Data Classification

Edgar Leonairo Pencue-Fierro, *Student Member, IEEE*, Yady Tatiana Solano-Correa, *Student Member, IEEE*, Juan Carlos Corrales-Muñoz, and Apolinar Figueroa-Casas

Abstract—The classification of land covers is one of the most relevant tasks carried on to understand the state of a certain region. Additional studies about the biodiversity, hydrology, human impact, modeling dynamics, and phenology in the study area, can be carried on. In these cases, a wide temporal series of images need to be considered in order to get the tendencies throughout the years. In some regions, such as the South-West part of Colombia (Andean region), studies over large areas are needed in order to obtain unified and coherent statistics that can be representative of the region. This means that different images, acquired by the same satellite and over different areas, or acquired by different sensors, or at different times, need to be classified. Standard classification methods do not work properly to perform this task, due to the heterogeneity in both land cover and orography. This paper presents a hybrid approach for the classification of multitemporal, multiregion, and multisensor images. Classification and regression trees (CART) decision tree and an SVM-based clustering were used in cascade in order to get the final classification maps. Experimental results carried over three Landsat Path/Rows, three sensors, and six different years, confirm the effectiveness of the proposed approach, where the overall accuracy was of 93% with a kappa factor of 0.92.

Index Terms—Image classification, Multisensor Landsat images, multitemporal data, radiometric indices, remote sensing (RS).

I. INTRODUCTION

CHANGES in land use and land cover (LULC) have a significant relevance for the understanding and study of the Earth natural dynamics. Their relation with the human intervention is represented mainly by deforestation, biodiversity lost, and climate change [1]. These variations have resulted in the development of extreme events, such as floods, landslides, and droughts [2]. The natural systems have lost their capability to give the basic ecosystemic services support. Specifically,

Manuscript received March 8, 2016; revised July 26, 2016 and October 6, 2016; accepted October 25, 2016. Date of current version November 30, 2016. This work was supported in part by the Colciencias and University of Cauca under Project Programa para el fortalecimiento de la Red Interinstitucional de Cambio Climático y Seguridad Alimentaria (RICCLISA) and in part by the University of Cauca and the WP4 team.

E. L. Pencue-Fierro, J. C. Corrales-Muñoz, and A. Figueroa-Casas are with the University of Cauca, Cauca 190001, Colombia (e-mail: leonairo@unicauca.edu.co; jcorral@unicauca.edu.co; apolinar@unicauca.edu.co).

Y. T. Solano-Correa is with the Fondazione Bruno Kessler, Center for Information and Communication Technology and the Department of Information Engineering and Computer Science, University of Trento, Trento I-38123 Italy (e-mail: solano@fbk.eu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>

Digital Object Identifier 10.1109/JSTARS.2016.2623567

they do not have the capacity to store water to be freed slowly during the raining seasons. Therefore, the drought periods are more extreme, establishing a closed cycle that can be noticed at the global level [3]. In most of the cases, the above mentioned events are related to the LCLU changes. Hence, knowing what is happening with the LULC can provide schemes and tools to manage the natural resources [4]. Remote sensing (RS) provides tools to get proper information allowing the study of LULC changes at both, spatial and temporal levels [5]. Nowadays it is possible to access to current and historical information in a freeway [6]. This has increased the possibilities to perform deep analysis over the whole world. The Landsat program is one of the most well-known sources used for this kind of studies, covering a little bit more than four decades of information until the moment. Its spatial resolution (30 m) gives the opportunity to obtain detailed maps for LULC. Its spectral response covers bands suitable for the vegetation study. Therefore, it is possible to analyze areas, such as agricultural ones, forest, páramo, and grassland, among others. These land covers (LC) are of main relevance to study the impact, dynamics, and tendencies of human intervention over the territory. Resulting in the possibility to establish regulations about the proper political management that can guarantee the regional sustainability.

The process to detect the LULC changes consist of two steps: 1) classify the LULC throughout the use of RS information and 2) find the changes in the features of interest between two times in a temporal series [7]. The critical process in this case is the proper classification of the LULC, this given that the wrong classification of a LULC will derive in a wrong temporal analysis. Therefore, it will generate nonreliable information. That is why it becomes important to develop a reliable classification system in order to get relevant conclusions for the natural resources management. The image classification consists in the assignation of a pixel to a certain class or LC. A criteria based on specific characteristics of each class is defined in such a way that classes are differentiable between them. The great amount of processes and information to work with, requires the development of automatic methods for the classification, especially when working with big areas and different sensors and times. Supervised and unsupervised methods have been developed to this end. The former requires the selection of training samples in which there is a total correlation between a specific pixel and the class to which it corresponds. In this case, the precision of the method is correlated to the way in which the training samples have been selected. Some of the widely used supervised methods

include artificial neural networks (ANN), maximum likelihood, minimum distance, k-nearest neighbor, support vector machines (SVM), and classification and regression trees (CART) [8]. The unsupervised classification use clustering mechanism according to a criteria, generally distance or similitude, to establish a set of classes that can describe the whole analyzed set. The main unsupervised methods are ISODATA algorithm, SVM based on clustering (SVM-C) [9], and k-means [10].

Several classification methods have been proposed in RS, all of them taking advantage of different aspects according to the spectral capacity, the spatial resolution or the temporal frequency of the images. Al-Doski *et al.* [11] use a pixel-based approach with similar spectral features that clusters according to a specific statistical criterion. Lu and Weng [12] focused on nonparametric classifiers, given that nonhypothesis about the data is required. In both cases, the problem is complex since it is difficult to integrate auxiliary data and spatial and contextual attributes. Li [13] applies the spatio-contextual analysis techniques in three categories: texture extraction, Markov Random Fields (MRF) and, analysis and segmentation of the images based on object. The problem with MRF is that they have not been fully accepted due to their theoretical and computational complexity. On the other hand, Li *et al.* [14] use the theory and evidences of the decision tree to classify pixels, allowing the measurement of the classification uncertainty. Perumal and Bhaskaran [15] use the Mahalonobis distance to perform the classification, assuming the spectral bands have a normal distribution. The problem of this classifier is that the final result has many unclassified pixels and there are also many overlapping among classes. Gomez *et al.* [16] analyze the methods for supervised and semisupervised classification by means of kernels. With these methods, it is possible to handle the noise in a better way, as well as the use of different sources. At the end, it is possible to improve the results of parametric methods and with neural networks (NN). The problem in this case is that classes should be linearly separable. Prasad *et al.* [17] make a spectral, spatial, and temporal analysis by using the information coming from different sensors. They also take into account the selection of suitable variables and auxiliary data that allows the improvement of the classification. The problem is that the kind of auxiliary data they are using is not always available for all the regions or the temporal period. Other supervised methods can be found in the literature, but there are still many issues regarding the proper selection of training samples, the features selection, the use of auxiliary data, and the computational capacity. These issues are seen in the classification accuracy. Some semisupervised techniques based on SVM have been also proposed in the literature [9], [18], [19]. Among them, the SVM-Clustering (SVM-C) has shown a good performance when different datasets are considered. An SVM-C algorithm does not require *a priori* knowledge of the input classes since the initial training samples are mapped to a different dimensional space in which a Gaussian kernel is used to make the data more separable [18]. The SVM-C has been used mainly for change detection purposes where only two classes are considered, changed and unchanged [20]. Therefore, some adaptations should be carried on for the use with classification of more than

two classes. Several methods are present in the literature, but all of them present many issues with regard to computational cost, selection of training samples, or precision. Therefore, the need to develop an approach that allows the proper classification of LULC over a multitemporal period and with the use of multisensory and multiregion images is required. In this paper, we propose a semisupervised hybrid approach that takes advantage of one supervised and one unsupervised methods to achieve the classification of LULC. The approach is developed such that it is possible to overcome two main problems: 1) to classify with high precision images coming from heterogeneous sensors and acquired over different regions and at different times, and 2) to classify in a time that is fast enough as to be able to deal with big amount of information. Multitemporal classifications can be performed by considering the selection of training samples coming from multiregions and multisensor images.

The remainder of this paper is structured as follows. We recall the classification basics and the Landsat data information in Section II. Section III illustrates the proposed semisupervised hybrid classification for multitemporal multiregion and multisensor Landsat images. Section IV describes the study area and the experimental results obtained on the study area dataset. Finally, conclusions and future work are given in Section V.

II. BACKGROUND ON CLASSIFICATION METHODS AND LANDSAT DATA

Classification algorithms are not supposed to receive data coming directly from the original spectral bands of the satellite sensor. Given the spectral closeness of the different bands, it might imply redundancy on the information content, making the convergence process of the training algorithm difficult. On the other hand, it is also possible that classification results are not optimal. Hence, the use of features derived from the original bands is nowadays proposed for the classification process. Features with a physical meaning, such as radiometric indices [21], [22], or features that modify the original bands in the spatial domain, can be considered. Whether to use spectral or spatial based features is correlated to the use of the same or different sensor for the acquisition of the images and to the spatial resolution of the satellite images. Features with a high correlation among them must be eliminated. Only features showing a higher probability to differentiate the LULC should be kept, so that the different classes can have a specific area in the feature space. Another important point to be considered, is the number of features to be used in order to keep a good tradeoff between complexity and capability to differentiate the LULC. Too many features result in a difficult training process, whereas few of them may not offer enough information for the LULC differentiation. Different techniques, allowing the proper selection of the features and their number, can be found in the literature [23], [24].

In the case of ANNs, they allow the introduction of new features in an easy way. The training process, including possible new features, can be performed with the already training samples to improve the classification result. This is an important characteristic given that it is possible to get a lot of different

features from satellite data, and it is important to be able to perform different trials until the best results are obtained. In the CART case, adding new features can imply an analysis in which the tree is modified from the base, and the classification rules are changed over the tree structure. Nevertheless, the rules built in the different trial can be used to build a knowledge database. This is due to the advantage to establish ranges for each feature that allows the separation of a certain class. Same process cannot be done by using ANNs, because the weights for each neuron are established in accordance to the correlation with the output condition, and not the meaning of a certain feature to differentiate a class. The CART method has the information in the established ranges giving an advantage with respect to the capability of the features to differentiate the LCLU. With this, it is possible to introduce new classes that will be differentiable if additional features are considered. In this context, the capability to discriminate the classes in the features space must be considered.

Generally, one single feature is able to divide the data in few sets that will not allow to differentiate a certain class. For this reason, it is necessary to use more than one feature. In this sense, the features are selected in accordance with the classification method to be used, in order to get a reliable classification between a reasonable time. The previous analysis allows us to understand that it is not possible to fulfill all the requirements by considering one single classification method. In this sense, the use of a hybrid method, exploiting the characteristics of two methods is suggested. In order to do this, the hybrid approach should perform a preliminary classification in which most of the LULC can be differentiated, and the computational cost can be the minimum. A second classification is performed to separate the remaining classes, by using a reduced feature space and number of LULC. According to their characteristics on speed and accuracy, the CART and SVM-C classifiers were found to be the best hybrid combination, further details on the selection of these two classifiers is given in the next section.

III. PROPOSED APPROACH TO MULTITEMPORAL MULTIREGION MULTISENSOR LANDSAT (MMM-L) CLASSIFICATION IN ANDEAN AREAS

Traditionally, classification methods made use of training samples selected only over the specific area of interest. According to the selected features, these training samples may be good enough to apply the resulted classification model to other images acquired over the same area of interest, but in a different time. Some problems are to be expected when moving to a different adjacent area, especially when considering features involving data transformation, or better to say, spatial transformations. This problem becomes more evident or critical while using images coming from different optical sensors, even if the same spatial resolution is kept. Such is the case of the multitemporal images acquired by the Landsat series. In this case, images are acquired under the same spatial resolution and under similar spectral resolution. Nevertheless, the sensors on-board the Landsat satellites are different from one satellite to the other, and have a different physical transformation for the acquired data.

Therefore, the selection of features that can keep the physical meaning of the data over the years is required. On the other hand, proper selection of the training samples should be considered, by extracting them over the whole multitemporal dataset and over different adjacent areas. Fig. 1 shows the general block scheme of the proposed approach for MMM-L classification. The input data for the classification process are images $X_{n,m}$, acquired over different times n and over different regions or Path/Rows m . The images may come from same $X_{n,m}$ or different Landsat sensors, according to the conditions of the study areas. First step consists on the MMM-L images preprocessing, where $X_{n,m}$ are converted to top of atmosphere (TOA). After that, a first feature extraction and proper selection of training samples is carried on. Then, a classification is performed by means of a CART decision tree. An evaluation, by means of a threshold value over the overall accuracy (OA), is then performed to separate differentiable LC-1 from nondifferentiable LC. At this point, a mask is generated from the nondifferentiable LCs in order to be used to extract new reduced features over those problematic areas and then separate the training samples for SVM-C. A classification of differentiable LC-2 is obtained, and the results are added to the differentiable LCs-1, detected with the CART decision tree. Finally, MMM-L classifications are obtained from the proposed hybrid approach. Further details are given in the next sections.

A. Feature Extraction Based on Coordination of Information on the Environment (CORINE) Land Cover Methodology

Let us assume the availability of a time series of optical images $X_{n,m}$, acquired by the Landsat series sensors over adjacent m regions and different n times. Let us also assume that different Landsat sensors, e.g., Landsat 4, Landsat 5, and Landsat 7, acquired the images. Given the use of multiple sensors, a preprocessing step to guarantee the homogeneity in between the $X_{n,m}$ images is required. This process consists of two steps [25]: 1) absolute radiometric homogenization [26] and 2) geometric homogenization [27]. The former will allow the comparison, from the physical viewpoint, between images acquired by different Landsat sensors. While the later will allow the pixel-to-pixel correspondence. According to the spatial resolution offered by the Landsat series, and depending on the region to be study, a specific number of LC can be classified. A specific project, known as CORINE land cover (CLC), was created for the classification of Landsat and SPOT images of the European Union (EU), in 1987 [28]. The project aims at creating a European database at scale 1:100 000 of the LULC, through the interpretation of RS data. To this aim, a methodology to evaluate the changes on the territory was established. This methodology allows the classification of the LULC in between different levels and gives a guideline for the development of classification algorithms. The CLC methodology has been adapted by different countries outside the EU in order to define the LC to be classify over a certain area.

According to the CLC definition, and the LC known to be present in Andean areas, ten classes were defined for the classification: water bodies, snow, páramo, urban areas, permanent crops, transitory crops, bare soil, grassland, natural forest, and

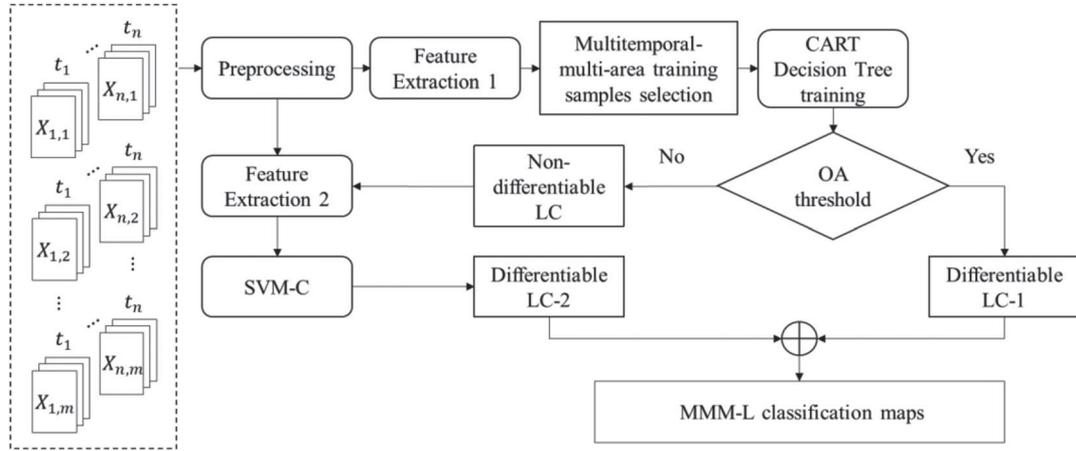


Fig. 1. General block scheme of the proposed approach for multitemporal multiregion classification.

planted forest. Two additional classes representing the clouds and their shadows were added to the classification process in order to avoid possible errors due to the nonexistence of a given LC under the cloud or the shadows. A final class, corresponding to the nonclassified pixels, is also created. In accordance with the 13 classes, the original spectral bands from each sensor may form part of the candidate features to be used for the classification. In order to keep the physical meaning of the features and in accordance with the state of the art, we propose the use of radiometric indices. The selection of a specific radiometric index is related to their capability to highlight certain LC, allowing its separation from the other LCs. Table I shows the list of radiometric indices selected for the classification of LC on the specific case of Andean areas. LCs can be divided into four main groups: 1) vegetation, 2) water, 3) urban area, and 4) bare land [29]. Based on it, most well-known indices allowing the separation of this kind of LCs were selected. All the indices take advantage of the spectral response of different LCs, allowing the separation of vegetation from urban areas, or water from bare soil. Other classes, such as the clouds and their shadows, can be distinguished as well by means of the original spectral bands or some radiometric indices. Such is the case of the automated cloud cover assessment (ACCA) algorithm [30] developed by the USGS to distinguish among different kind of clouds and used in their data center to evaluate the percentage of clouds present in each path/row image. Additional information provided by the digital elevation model (DEM) is also used. DEM plus indices in Table I correspond to the feature extraction 1, and the extracted ones are as input for the CART training. DEM is mainly used for the classification of LCs, such as the páramo, snow, planted forest, and transitory crops.

After the first classification with CART is applied, differentiable and nondifferentiable LCs are obtained. The nondifferentiable ones are used as a mask to extract new reduced features from the set of radiometric indices given in Table I. The reduced set is dependent on the nature of nondifferentiable LCs. In this research, it was found that nondifferentiable LCs usually correspond to vegetation kind ones. Therefore, feature extraction

TABLE I
LIST OF RADIOMETRIC INDICES USED AS FEATURES FOR
CLASSIFICATION OF ANDEAN REGIONS

Radiometric Index	Equation
Normalized difference Vegetation index [31]	$NDVI = \frac{NIR - R}{NIR + R}$
Enhanced vegetation index [32]	$EVI = \frac{2.5 * (NIR - R)}{NIR + 2.4 * R + 1}$
Normalized difference water index [33]	$NDWI = \frac{G - NIR}{G + NIR}$
Modified normalized Difference water index [34]	$MNDWI = \frac{SWIR1 - G}{SWIR1 + G}$
Moisture spectral index [35]	$MSI = \frac{SWIR1}{NIR}$
Soil adjusted vegetation index [36]	$SAVI = \frac{NIR - R}{(NIR + R + L)} * (1 + L)$
Simple ratio [37]	$SR = \frac{NIR}{R}$
Normalized difference snow index [38]	$NDSI = \frac{G - SWIR1}{G + SWIR1}$
Atmospheric resistance vegetation index [39]	$ARVI = \frac{NIR - (2 * R - B)}{NIR + (2 * R - B)}$
Structure insensitive pigment index [40]	$SIPI = \frac{NIR - B}{NIR - R}$
Carotenoid reflectance index [41]	$CRI1 = \frac{1}{B} - \frac{1}{G}$
Automated cloud cover assessment [42]	$ACCA : B3 > 0.8; B6 > 300; (\frac{B4}{B3}) > 2; (\frac{B4}{B5}) > 1; (1 - B3) * B6 < 225$

2 (see Fig. 1) are SR, MNDWI, EVI, NDWI, SIPI, and SAVI indices. This features are extracted from the original $X_{n,m}$ images throughout the mask obtained from the nondifferentiable LCs. Then, the SVM-C is trained in a semisupervised way.

B. MMM-L Training Samples Selection

In the proposed approach, the selection of the training samples depends on the classification algorithm that we are using. For the first stage of the proposed approach, CART algorithm is

used and then pixel samples can be considered for the training process. For the second stage of the MMM-L approach an SVM-C is used. In this case, larger areas can be considered as training samples. Let us first discuss about the problems and selection of training samples for CART and later on about the ones for SVM-C.

Selection of training samples is traditionally done by photointerpretation and by selecting a z number of samples for each of the classes to be classified in the study area and always over one single image. Given the orographic or the weather conditions, this task can become difficult for some classes than for others. For these reasons, the number of z samples can be lower for some of the classes. In some cases, there are that many clouds, that it is impossible to collect training samples for all the classes. When the size of the study area increases, the above mentioned problems become more and more critical. Such is the case of images acquired over the Andean regions, where the probability to get a free or partially cloud free image is really low. In order to increase the probability to have more cloud free images, the use of multitemporal images could be considered. On the other hand, adjacent regions or path/rows can also be considered. An additional solution can be considered by taking advantage of the location of basins and/or subbasins on the study area. In this specific case, the probability to find suitable Landsat images increases over the time, given the partial reduction of the study area. Traditionally, suitable images for classification and selection of training samples are filtered by the cloud coverage percentage. This reduces the number of images that can be used for classification and thus, for performing a temporal analysis. When subbasin areas are considered, it is possible to have more temporal images and it is also possible to have a greater coverage of a certain LC that allows us to extract enough training samples. The selection of training samples coming from different times is comparable to the variability found when samples are extracted along adjacent Path/Rows, plus additional changed pixels that might be found. Considering all these possibilities may allow the increase of the classifier's robustness, as well as the OA. Therefore, we proposed a solution in which training samples are extracted from multitemporal, multiregion areas. When working with Landsat series images, these multitemporal images are acquired usually by different Landsat satellites, i.e., by different sensors. Therefore, the proposed approach for the training sample selections is also multisensor. Classes to be classified are defined in previous section and training samples are selected from different adjacent path/rows and at different n times. Since most of the images were acquired many years ago, there might not necessarily be availability of *in situ* data. Therefore, most of the training samples are selected by photointerpretation. For the latest years, field visits can be done in order to get the *in situ* training samples.

Once the CART classification result is obtained, the non-differentiable LC are used to create a mask, such that only the pixels falling inside that area are considered for the second stage of the MMM-L classification. With this mask in mind, new reduced features are extracted such that it is possible to separate among the nondifferentiable LCs. All these pixels are used as the training samples for the SVM-C, which is then used as a

semisupervised classifier, allowing the separation of the second set of differentiable LCs.

C. Semisupervised Hybrid Classification Approach for MMM-L Data

The proposed approach to MMM-L classification is based on the use of physical meaningful features and machine learning techniques. Three main steps compose the proposed approach: 1) the features extraction, 2) training sample selection, and 3) classification step. As mentioned in Section II-A, the features extracted for the classification process are mainly radiometric index based. Many other classification models, based on these kind of features, identify the LCs types by using linear model [8]. Some problems for the proper classification are usually found, given the lack of pure pixels representing each LCs. On the other hand, the computational cost of linear models is much lower, compared to nonlinear models, and is, therefore, preferred for classification. When larger areas and multitemporal data are considered, automatic techniques are also preferred due to the training samples selection process. Both, supervised and unsupervised classifications present advantages and disadvantages. Hence, we propose the implementation of a hybrid classifier made up of a CART decision tree and an SVM-C. The selection of this hybrid technique corresponds to the selective search over different classification algorithms that could allow us to perform the classification in an efficient way from two perspectives: 1) the computational time and 2) under variable conditions with multiple-regions and multitemporal images. Under multiregion and multitemporal conditions, it is common to find several LCs making the problem even more complex. Supervised classifiers were mainly explored in order to have a more controlled solution of the problem. After analyzing single-by-single LC, it was possible to see that there were certain classes with a tendency to be classified worse than the others due to the lack of separability from the available features. These allow us to conclude that the best solution was to perform the classification by stages, helping in the first stage the classes that were better recognize by the CART and in the second stage the remaining classes. For the second stage, the remaining classes were evaluated with different supervised algorithms, but finally it was found that by using an unsupervised method, such as SVM-C, it was possible to separate them in a reasonable way.

A preliminary classification, done by the CART decision tree, is carried on over the whole training samples of the different classes. All the LC results are divided into two groups: 1) differentiable and 2) nondifferentiable LCs. The rule followed to separate the LCs is based on the single class accuracy assessment. A threshold T limit is set up such as that all the LCs with an accuracy assessment lower than T are grouped as nondifferentiable and the others as differentiable. A second classification is performed by using SVM-C, but this time only over the nondifferentiable LCs. The separation of these LCs is performed by applying a radiometric index filter base. The selection of the radiometric index is correlated to the nature of the nondifferentiable LCs. For SVM-C, the initial training samples are mapped to a different dimensional space in which a Gaussian

kernel is used to make the data more separable. In this space, the smallest spheres that can enclosed the resulting clusters are searched. With this spheres, the data is transformed to the original space and contours limiting the space of the different clusters are created. These limits are correlated to the location where the support vectors will be in the alternative space. The parameters used in SVM-C are the same as for the standard SVM and are defined as follows: 1) gamma factor determines if a sample belongs or not to a cluster according to its distance to the support vector, and 2) the cost function C makes the situation of outliers flexible by using soft margins, this allows the spheres not to enclose all the samples in the space and at the same time allows to work with overlapping clusters [18].

The improved classifications (LC-2) are added to the previous ones (LC-1) and the MMM-L classification result is obtained. The OA is expected to be improved by using the hybrid approach, instead of a single CART decision tree. On the other hand, the computational time required for the classification is expected to be reduced, while comparing to a classification that would have been done by using only a SVM or any other unsupervised technique.

IV. STUDY AREA AND EXPERIMENTAL RESULTS

A. Study Area and Dataset description

The study area selected for the training and application of the MMM-L classification approach is the upper basin of the Cauca River, located in the South-West part of Colombia. The upper basin of the Cauca River is comprised between $74^{\circ}59'55.9''$ – $77^{\circ}01'59.5''$ W and $1^{\circ}55'33.6''$ – $6^{\circ}45'14.3''$ N (UTM coordinates), with a total area of 25.000 km^2 . Location of the study area is represented in Fig. 2 by the dark blue enclosed area. The area hosts the 20% of Colombia population [43] and covers five Departments (Caldas, Cauca, Quindío, Risaralda, and Valle del Cauca) and 183 municipalities. This area is of strategic importance because it covers almost all the production chains prioritized by the agriculture ministry in Colombia, covering most of the technical and high value agriculture in the Country. We can find sugar and coffee industry, fruits, cassava, rice, cocoa, beans, bananas, corn, forest for commercial purposes and manufacturing production. Among its orographic characteristics, we can find the high variability of the altitude, between 950 m in the alluvial valley of Cauca, and 4650 m at the top of the Purace volcano (Department of Cauca). Annual precipitation ranges from 1500 to 3000 mm.

The proposed approach is developed for its application over images acquired by the Landsat series satellites [44] over an Andean region. The goal of this research is to have multitemporal classifications over all the upper basin of the Cauca River. Therefore, images acquired at different times are considered. Given the high variability on the altitude and precipitation, problems with high presence of clouds are to be expected. This generates two main problems: 1) lack of enough free of clouds images and 2) problems to extract enough training samples when considering images acquired by one single Landsat satellite. Therefore, the use of images acquired by the Landsat 4 (TM and MSS), Landsat 5 (TM and MSS), and Landsat 7 (ETM+) satellites

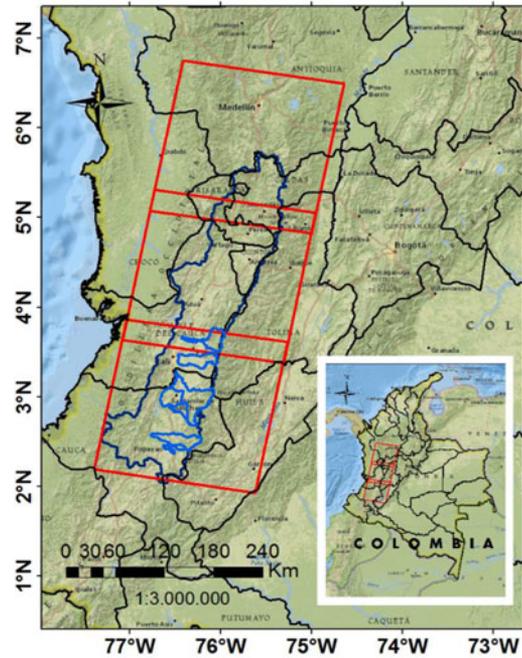


Fig. 2. Location of the study area.

TABLE II
LIST OF LANDSAT IMAGES USED FOR THE TRAINING AND CLASSIFICATION PROCESSES

Year	Path/Row		
	009/056	009/057	009/058
1	LT4/1989219	LT4/1989219	LT4/1989219
2	LT5/1996215	LT5/1999191	LT5/1999191
3	LE7/2001028	LE7/2004037	LE7/2008256
4	LE7/2002207	LE7/2008256	LE7/2010021
5	LE7/2009194	LE7/2010005	LE7/2015003
6	LE7/2015003	LE7/2015003	none

is necessary in order to increase the probability to get free of clouds images and the possibility to extract training samples. In order to cover the study area, images acquired over three consecutive path/rows are required (see red squares in Fig. 2). A deep search on the US Geological Survey (USGS) [45] was carried on in order to obtain the suitable MMM-L images. Nevertheless, and given the orography of the study area, the amount of clouds present in the different images was still too high to perform a proper classification analysis. Therefore, a different strategy is used to select the suitable images. The strategy consists on the selection of subbasins of the study area, throughout the use of vector files projected in the USGS database, allowing the construction of a more extensive database. Table II shows the detailed information about the images used for the training and classification processes, their path/rows, the satellite that acquired them and the acquisition date. Some of the subbasins used to build the MMM-L database are highlighted in light blue in Fig. 2.

Given the use of Landsat images acquired by different sensors, an absolute homogenization process, by transforming the Digital Numbers (DN) to TOA reflectance, was applied to all the images. The common spatial resolution for all the images was of 30 m, images acquired by MSS sensor were already resampled by the USGS in the moment of downloading them from their database. Additional corrections for the Scan Line Corrector-OFF problem of the ETM+ sensor was done by using the GDAL/FillNoData tool [46]. The smoothing option provided by this function was not considered in order to reduce the amount of undesirable artifacts.

B. Results for Multitemporal Classification

The final classification process is based on the combination of two classification algorithms: CART decision tree and SVM-C. Their implementations were done by means of functions from the OpenCV library [47]. In the case of SVM-C, the required parameters were estimated by cross-validation process. In order to train the CART decision, 15 200 samples were selected from the different subbasins mentioned in Table II. These samples represent 12 classes. The 63% of the samples were used for training while the others were used for validation process. A total of 11 radiometric indices, the DEM file, and the ACCA criteria were used as features for the training process. After the CART decision tree training, the classification of the study area was performed. By means of the accuracy assessment and by the use of a threshold over the single-class accuracy, it was possible to separate the differentiable and nondifferentiable LCs. The threshold was setup to 88% of accuracy and resulted in the selection of all the vegetative classes as the nondifferentiable ones. Even though the differentiation was not wrong for all the vegetation LCs, we considered that a better performance could be achieved for them while using another machine learning method. Therefore, an SVM-C was used. Given that nondifferentiable LCs correspond to vegetation areas, a NDVI filter base was used to select only vegetation areas for the training process. Only the SR, MNDWI, EVI, NDWI, SIPI, and SAVI indices were used as features for the SVM-C training process. The final step is to perform the classification of the nondifferentiable LC and build the final classification maps for each year and the whole study area. For the SVM-C, a total of four clusters (corresponding to the nondifferentiable LCs) were used as input, and a RBF kernel of grade 3 was used. The optimal parameters for the kernel were obtained by means of the grid search algorithm. The best OA for the four classes was of 87.83% for $C = 7.4$ and $\gamma = 0.0025$. Kernels with higher dimensionality were also tried, but similar results in terms of accuracy requiring a bigger classification time were obtained. It is important to note that the initial number of clusters are not affecting in a critical way the final accuracy of the SVM-C, not more than the same training samples. This is because the number of clusters are not a parameter of the SVM-C.

Given the size of the study area, only one of the subbasins, present in the upper basin of the Cauca River, is used to show the classification results over the years 1989, 1999, 2008, 2010, and 2015 (see Fig. 3), but classification results are available for

all the images mentioned in Table II. The subbasin is located in Path-009/Row-058 and known as Palacé subbasin, its area is of 642.93 km². The results show 11 out of the 12 initial classes. This was due to the difficulty for the system to separate the classes corresponding to water bodies and shadows from clouds, because of the presence of high density clouds.

In order to have a better idea of the classifications results and the improvements gained while implementing a semisupervised hybrid classifier, smaller areas extracted from the Valle del Cauca Department (Path-009/Row-057) and acquired in 2015, by Landsat 7 sensor, are shown in Fig. 4. The left column shows the 453 false color composition of each area. Whereas the central and right column shows the CART and CART+SVM-C classification results, respectively. A qualitative and quantitative comparison with other state of the art classifiers was also carried on. In addition to CART decision tree, random forest (RF), NN, and single SVM were also applied. Because of space constrain, the classification maps for the RF, NN, and SVM are not shown in Fig. 4, but the quantitative results are further analyzed in the next.

The real area of the images in Fig. 4 is different from one to the other but in order to take advantage of the space in the figure; they were resampled to the same width. Real areas are like this: Candelaria is 221.28 km², Pradera is 92.78 km², Roso is 150.66 km², and El Sonso Lagoon is 70.97 km². These areas were selected because of their high presence of crops and vegetation LCs, whose classification was improved by adding the SVM. The first thing we can see from the results is the visual improvement reached by the CART+SVM-C classifier when referring to crops LC. When using only CART, all the permanent crops are confused with the spectral response of the planted and natural forests. The addition of the SVM-C classifier allows, the proper differentiation of permanent crops and planted and natural forests. There are other areas with problems in the differentiation of grassland and transitory crops as well.

In the upper basin of the Cauca River, there is the presence of some important volcanos. The peak of one of them is covered with snow during the whole year. As common from the volcanos' areas, rocks can be also found in the surrounding areas. This class is not considered in the CLC methodology and, therefore, no training samples were selected from them to train the CART classifier. The results from this situation is the incorrect classification of that LC and its specific confusion with the snow LC. A specific example is shown in Fig. 5, where a small portion of the Nevado del Ruiz volcano, located in the Caldas Department - Path-009/Row-057, can be seen. In the 453 false color composition shown in Fig. 5(a), LC, corresponding to the rocks, can be seen in the dark grey area, whereas the snow is represented by the fuchsia color. An additional LC of importance for this case is the páramo one, which is shown as the emerald green color. Fig. 5(b) shows the classification results, where it is clear that snow class is given for the whole rock area, given the nonexistence of such a class. On the other hand, transitions from snow to páramo are difficult to handle for the algorithm confusing it with bare soil and grassland. Apart from these specific cases, visual inspection of the classification results over the MMM-L images are considered as satisfactory.

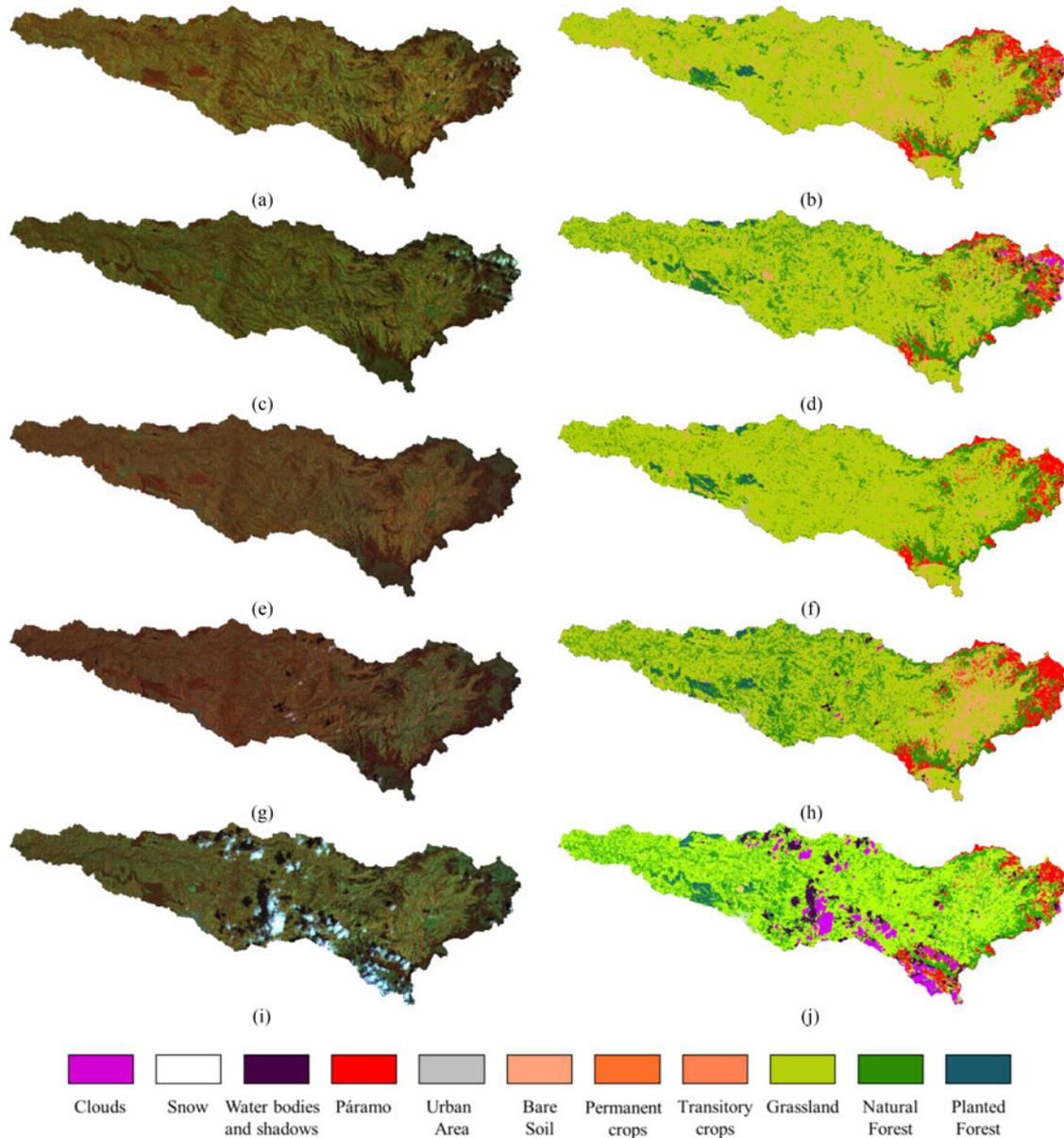


Fig. 3. Classification results with CART+SVM-C algorithm. Left column shows the 453 combination of the area and right side its corresponding classification.

A quantitative analysis was also carried on in order to evaluate the classification performance and to know the computational cost.

C. Performance Evaluation

In order to evaluate the performance of the proposed approach for the classification of MMM-L images over the Andes area, multitemporal reference data, collected by photointerpretation and field trips, were used. The evaluation is performed over the whole areas described in Table II. A total of 15 200 samples were collected over the MMM-L. From them, 5493 samples were used to evaluate the performance of the CART, RF, NN, SVM, and the CART+SVM classifiers. Three accuracy measurements were employed in the process of accuracy assessment: 1) Accuracy for each single class, 2) OA, and

3) kappa coefficient (κ). Table III shows the results for the performance evaluation of the CART+SVM-C algorithm. In order to make a comparison, performance results from the CART decision tree, before adding the SVM-C, are also shown. Now it becomes clear that the highest improvement is assessed for the permanent crops LC, which were wrongly classified while using only CART decision tree. On the other hand, improvements for grassland and natural and planted forest are also achieved. For the specific case of RF, it is possible to see that the accuracy assessment is in general higher for most the LCs, at least when compared to the CART algorithm. Nevertheless, it is important to recall that these improvements are not achieved for all the LCs, but just for some of them, which are not the vegetation areas ones. Therefore, we selected CART over RF. The different LC improvements result in the increase of the OA of about 9%

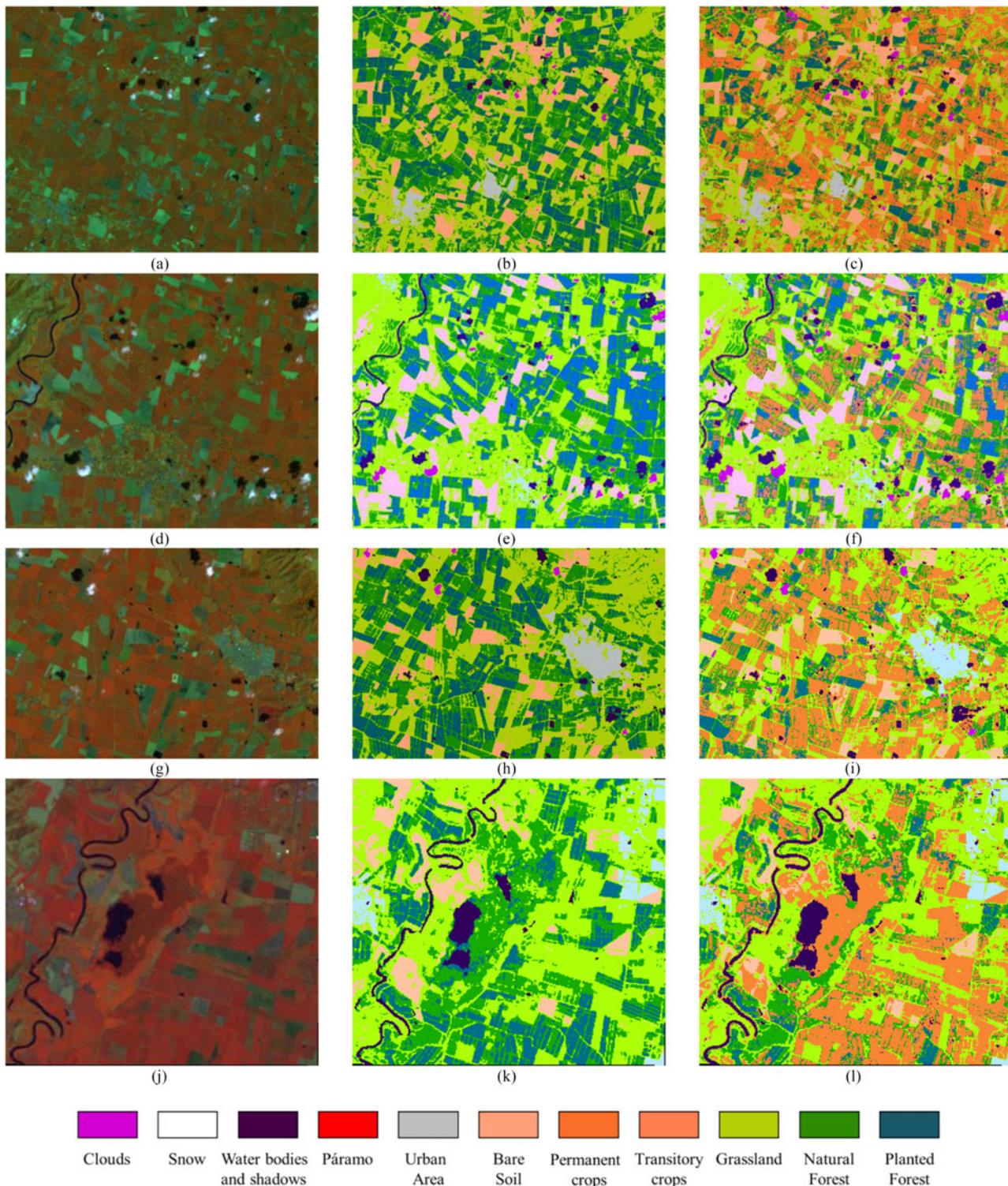


Fig. 4. Classification results obtained with the single CART decision tree (central column) and the joined CART+SVM-C (right column). (a), (b), and (c) Candelaria; (d), (e), and (f) Pradera; (g), (h), and (i) Roso; (j), (k), and (l) El Sonso Lagoon.

for the CART+SVM-C over CART, and of around 12%, 8%, and 11% for SVM, RF, and NN, respectively.

The required time to perform the classification by using the CART+SVM-C algorithm is correlated to the size of the study area. Areas with similar size to a single path/row from Landsat

take around 508 min to be processed. This time is obtained by using OpenCV 2.4 on a standard workstation. Hardware is Intel(R) Core(TM) i7-3630QM CPU @2.40 GHz, 8.00 GB Ram. The processing time decreases as the area to be classified decrease. Table IV shows the processing times for areas with

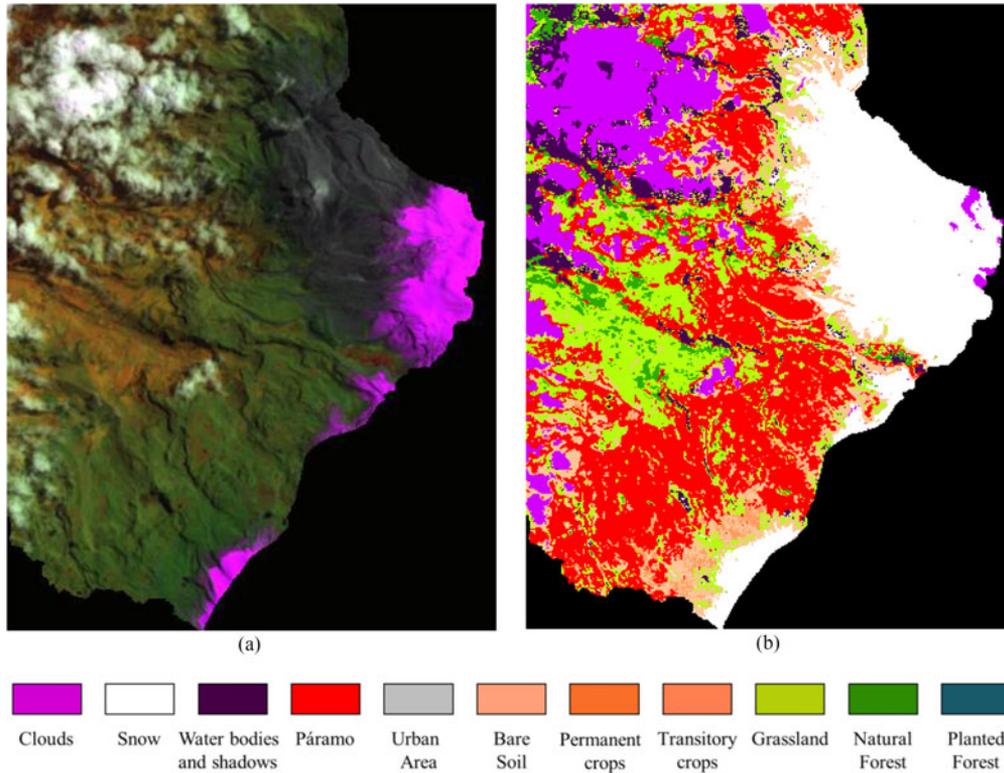


Fig. 5. Classification results obtained with the CART+SVM-C classifier for the Nevado del Ruiz volcano.

TABLE III
ACCURACY ASSESSMENT RESULTS FOR THE CLASSIFICATION PROCESS WITH SVM, RF, NN, CART, AND CART+SVM-C

Class	Evaluation accuracy (%)				
	SVM	RF	NN	CART	CART+SVM-C
Clouds	98.86	99.51	99.51	98.40	98.40
Water bodies and shadows	91.08	99.32	96.22	99.46	99.46
Snow	99.71	99.43	100.00	89.28	89.28
Páramo	81.86	91.15	93.07	95.22	95.22
Urban areas	87.33	87.47	82.35	94.52	94.52
Grassland	26.21	37.54	51.78	87.93	88.79
Bare Soil	84.08	79.42	55.73	90.40	90.40
Permanent crops	64.68	88.51	85.11	0	86.38
Transitory crops	78.17	78.17	83.80	79.72	79.72
Natural Forest	56.94	58.80	76.85	65.81	90.93
Planted Forest	84.62	86.32	66.84	60.75	85.21
κ	0.79	0.84	0.80	0.83	0.92
OA	81.45	85.44	82.84	84.84	93.03

TABLE IV
PROCESSING TIME REQUIRED BY THE CART+SVM ALGORITHM FOR THE CLASSIFICATION

Path/Row size	Processing time (minutes)
1	508
1/2	245
1/4	127
1/8	69
1/16	35
1/32	18

different sizes, as portions of a single path/row. The processing time for the single CART decision tree is 5.7 min. Nevertheless, the use of CART decision tree, as proved by results in Table III, does not allow a proper classification for some of the classes.

V. CONCLUSION

A semisupervised hybrid approach for the classification of MMM-L images has been proposed. For the hybrid approach, a CART decision tree and a SVM-C classifiers were used. Features used for the training process were selected based on the CORINE land cover methodology and the LC known to be present in the Andes region. The selection of training samples from smaller regions located over different path/rows and coming from multiple years, allows the proposed MMM-L classifier to be robust enough as to be able to classify images coming from different Landsat sensors and acquired over different years, without having to perform a new training process. Qualitative and quantitative analysis carried on over the final multitemporal classification maps lead to a final OA improvement of about 9% over the use of a single CART decision tree and around 8–12% for other state of the art classifiers. Time required for the classification process was found to be reasonable, especially if a parallelization process is considered in the future. Some problems regarding the proper classification of rock classes, and the transitions from snow to páramo areas, need to be considered in future works.

Future developments could consider the use of other features in order to achieve the proper separation of water bodies and

clouds' shadows. The analysis over the land cover evolution could be considered as well. The proposed approach was developed for the classification of Landsat images acquired by MSS, TM and ETM+ sensors, but OLI sensor was not considered for the lack of free of clouds images to extract training samples at the time of the experiments. A new search over the USGS database should be carried on in order to integrate new upcoming images and to increase the capability of the approach. Further experiments by using pansharpened Landsat data (15 m resolution) could be carried on in order to understand the robustness of the approach to the variability of the spatial resolution. This is in order to be able to integrate the proposed approach with the new Sentinel-2 images, which spatial resolution is slightly higher than the one of Landsat images.

REFERENCES

- [1] C. Onur and A. Tezer, "Ecosystem services based spatial planning decision making for adaptation to climate changes," *Habitat Int.*, vol. 47, pp. 267–278, Jun. 2015.
- [2] R. Espada Jr., A. Apan, and K. McDougall, "Spatial modelling of natural disaster risk reduction policies with Markov decision processes," *Appl. Geography*, vol. 53, pp. 284–298, Sep. 2014.
- [3] T. Ariti, J. van Vliet, and P. H. Verburg, "Land-use and land-cover changes in the Central Rift Valley of Ethiopia: Assessment of perception and adaptation of stakeholders," *Appl. Geography*, vol. 65, pp. 28–37, Dec. 2015.
- [4] S. Zakaria Aldwaik and R. Gilmore Pontius Jr., "Intensity analysis to unify measurements of size and stationarity of land changes by interval, category, and transition," *Landscape Urban Planning*, vol. 106, no. 1, pp. 103–114, May 2012.
- [5] J. S. Rawat and M. Kumar, "Monitoring land use/cover change using remote sensing and GIS techniques: A case study of Hawalbagh block, district Almora, Uttarakhand, India," *Egypt. J. Remote Sens. Space Sci.*, vol. 18, no. 1, pp. 77–84, Jun. 2015.
- [6] J. G. Masek, D. J. Hayes, M. J. Hughes, S. P. Healey, and D. P. Turner, "The role of remote sensing in process-scaling studies of managed forest ecosystems," *Forest Ecology Manage.*, vol. 355, pp. 109–123, Nov. 2015.
- [7] N. W. T. Quinn and O. Epshtein, "Seasonally-managed wetland footprint delineation using Landsat ETM+ satellite imagery," *Environ. Model. Softw.*, vol. 54, pp. 9–23, Apr. 2014.
- [8] D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, "Machine learning in geosciences and remote sensing," *Geosci. Frontiers*, vol. 7, no. 1, pp. 3–10, Jan. 2016.
- [9] S. Winters-Hilt and S. Merat, "SVM clustering," *BMC Bioinf.*, vol. 8, no. 7, pp. 1–12, 2007.
- [10] J. R. Otukei and T. Blaschke, "Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 12, no. 1, pp. S27–S31, Feb. 2010.
- [11] J. Al-Doski, S. Mansori, and H. Mohd, "Image classification in remote sensing," *J. Environ. Earth Sci.*, vol. 3, no. 10, pp. 141–148, 2013.
- [12] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, 2007.
- [13] M. Li, "A review of remote sensing image classification techniques: The role of spatio-contextual information," in *Proc. EuJRS Eur. J. Remote Sens.*, 2014, pp. 389–411.
- [14] X. Li, Q. Xing, and L. Kang, "Remote sensing image classification method based on evidence theory and decision tree," in *Proc SPIE Multispectral, Hyperspectral, Ultraspectral Remote Sens. Technol., Techn. Appl. III*, 2010, pp. 1–8.
- [15] K. Perumal and R. Bhaskaran, "Supervised classification performance of multispectral images," *J. Comput.*, vol. 2, no. 2, pp. 124–129, 2010.
- [16] L. Gómez-Chova, J. Muñoz-Marí, V. Laparra, J. Malo-López, and G. Camps-Valls, "A review of kernel methods in remote sensing data analysis," in *Proc. Opt. Remote Sens.*, 2011, pp. 171–206.
- [17] S. Prasad, T. Satya, and I. Murali, "Techniques in Image Classification: A Survey," *Global J. Res. Eng.*, vol. 15, no. 6, pp. 16–32, 2015.
- [18] A. Ben-Hur, D. Horn, H. Siegelmann, and V. Vapnik, "Support vector clustering," *J. Mach. Learn. Res.*, vol. 2, pp. 125–137, Dec. 2001.
- [19] K. Jong, E. Marchiori, and A. van der Vaart, "Finding clusters using support vector classifiers," in *Proc. 2003 11th Eur. Symp. Art. Neural Netw.*, Bruges, Belgium, Apr. 2003, pp. 223–228.
- [20] F. de Morsier, D. Tuia, M. Borgeaud, V. Gass, and J.-Ph. Thiran, "Unsupervised change detection via hierarchical support vector clustering," in *Proc. IEEE IAPR Workshop Pattern Recognit. Remote Sens.*, 2012, pp. 1–4.
- [21] R. Infascelli, S. Faugno, S. Pindozi, L. Boccia, and P. Merot, "Testing different topographic indexes to predict wetlands distribution," *Procedia Environ. Sci.*, vol. 19, pp. 733–746, 2013.
- [22] J. L. Zarate-Valdez, M. L. Whiting, B. D. Lampinen, S. Metcalf, S. L. Ustin, and P. H. Brown, "Prediction of leaf area index in almonds by vegetation indexes," *Comput. Electron. Agriculture*, vol. 85, pp. 24–32, Jul. 2012.
- [23] S. Serpico, M. D'Inca, F. Melgani, and G. Moser, "Comparison of feature reduction techniques for classification of hyperspectral remote sensing data," *Proc. SPIE 4885, Image Signal Process. Remote Sens. VIII*, vol. 347, pp. 347–358 2003.
- [24] V. Kumar and S. Minz, "Feature selection: A literature review," *Smart Comput. Rev.*, vol. 4, no. 3, pp. 211–229, Jun. 2014.
- [25] Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "Change detection in very high resolution multisensor optical images," in *Proc. SPIE Conf. Image Signal Process. Remote Sens. XX*, 2014, vol. 9244.
- [26] S. M. Vicente-Serrano, F. Pérez-Cabello, and T. Lasanta, "Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of Landsat images," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3916–3934, Oct. 2008.
- [27] S. Hantson and E. Chuvieco, "Evaluation of different topographic correction methods for Landsat imagery," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 13, no. 5, pp. 691–700, Oct. 2011.
- [28] J. Feranec, G. Hazeu, S. Christensen, and G. Jaffrain, "Corine land cover change detection in Europe (case studies of the Netherlands and Slovakia)," *Land Use Policy*, vol. 24, no. 1, pp. 234–247, Jan. 2007.
- [29] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 67, pp. 93–104, Jan. 2012.
- [30] G. R. Watmough, P. M. Atkinson, and C. W. Hutton, "A combined spectral and object-based approach to transparent cloud removal in an operational setting for Landsat ETM+," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 13, no. 2, pp. 220–227, Apr. 2011.
- [31] T. N. Carlson and D. A. Ripley, "On the relation between NDVI, fractional vegetation cover, and leaf area index," *Remote Sens. Environ.*, vol. 62 no. 3, pp. 241–252, 1997.
- [32] Z., Jiang, A.R. Huete, K. Didan, and R. Miura, "Development of a two-band enhanced vegetation index without a blue band," *Remote Sens. Environ.*, vol. 112, no. 10, pp.3833–3845, 2008.
- [33] B. C. Gao, "Normalized difference water index for remote sensing of vegetation liquid water from space," *Proc. SPIE*, vol. 2480, pp. 225–236, 1995.
- [34] K. Rokni, A. Ahmad, A. Selamat, and S. Hazini, "Water feature extraction and change detection using multitemporal Landsat imagery," *Remote Sens.*, vol. 6, no. 5, pp. 4173–4189, 2014.
- [35] E. Hunt Jr. and B. N. Rock, "Detection of changes in leaf water content using near- and middle-infrared reflectances," *Remote Sens. Environ.*, vol. 30, pp. 43–54, 1989.
- [36] R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote Sens. Environ.*, vol. 25, pp. 295–309, 1988.
- [37] P. J. Sellers, "Canopy reflectance, photosynthesis and transpiration," *Int. J. Remote Sens.*, vol. 6, pp. 1335–1372, 1985.
- [38] Y. Inoue, J. Peñuelas, A. Miyata, and M. Mano, "Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO₂ flux measurements in rice," *Remote Sens. Environ.*, vol. 112, no. 1, pp. 156–172, 2008.
- [39] Y. J. Kaufman and D. Tanre, "Strategy for direct and indirect methods for correcting the aerosol effect on remote sensing: From AVHRR to EOS-MODIS," *Remote Sens. Environ.*, vol. 55, pp. 65–79, 1996.
- [40] J. F. Penuelas, I. Baret, and I. Filella, "Semi-empirical indices to assess carotenoids/chlorophyll-a ratio from leaf spectral reflectance," *Photosynthetica*, vol. 31, pp. 221–230, 1995.
- [41] Y. Zur Gittelson, O. B. Chivkunova, and M. N. Merzlyak, "Assessing carotenoid content in plant leaves with reflectance spectroscopy," *J. Photochem. Photobiol.*, vol. 75, pp. 272–281, 2002.

- [42] H. Choi and R. Bindenschadler, "Cloud detection in Landsat imagery of ice sheets using shadow matching technique and automatic normalized difference snow index threshold value decision," *Remote Sens. Environ.*, vol. 91, no. 2, pp. 237–242, May 2004.
- [43] CDKN, "Análisis interinstitucional y multisectorial de vulnerabilidad y adaptación al cambio climático para el sector agrícola de la cuenca alta del Río Cauca impactando políticas de adaptación," *Climate Develop. Knowl. Netw.*, London, U.K., Tech. Rep. (Aug. 2013). [Online]. Available: <http://cdkn.org/wp-content/uploads/2014/02/Reporte-tecnico-final-AVA.pdf>
- [44] USGS, Landsat Project. (2016). [Online]. Available: <http://landsat.usgs.gov/>.
- [45] USGS, Global Visualization Viewer. (2016). [Online]. Available: <http://glovis.usgs.gov/>
- [46] N. Mueller *et al.*, "Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia," *Remote Sens. Environ.*, vol. 174 pp. 341–352, Mar. 2016.
- [47] OpenCV Developers Team. OpenCV Library. (2016). [Online]. Available: <http://opencv.org/>



Edgar Leonairo Pencue-Fierro (S'15) received the B.S. degree in physics engineering from the University of Cauca, Cauca, Colombia, in 2003. Finished the specialization in Project Management from the University of Cauca in 2009. He is currently working toward the Ph.D. degree in environmental sciences at the University of Cauca, Cauca, Colombia.

He is currently a Full Professor in the Physics Department of the University of Cauca, Colombia. He is a Researcher and member of the Optics and Laser Group and the Environmental Studies Group at the

University of Cauca where he is developing projects to integrate the information coming from remote sensors for the analysis of environmental conditions and trends in Colombia. He is currently a Researcher of RICCLISA program (Inter-institutional Network for Climate Change and Food Security) funded by COLCIENCIAS.



Yady Tatiana Solano-Correa (S'13) received the B.S. degree in physics engineering (honorable mention) from the University of Cauca, Cauca, Colombia, in 2011.

From 2009 to 2013, she worked as a Researcher and Teaching Assistant for the research groups: Optics and Laser Group (GOL), Environmental Studies Group (GEA), and CIAgua group. She is currently working toward the Ph.D. degree as member of the Remote Sensing Laboratory at the Department of Information and Communication Technologies, Uni-

versity of Trento and the Remote Sensing for Digital Earth Unit, Fondazione Bruno Kessler, Trento, Italy. Her research interests include remote sensing applications, change detection in multispectral and very high resolution images, multitemporal analysis, pattern recognition, image classification, and machine learning.



Juan Carlos Corrales-Muñoz received the B. S. and M.S. degrees in telematics engineering from the University of Cauca, Colombia, in 1999 and 2004, respectively. He received the Ph.D. degree in computer science from the University of Versailles Saint-Quentin-en-Yvelines, France, in 2008.

He is currently a Full Professor in the Electronics and Telecommunications Department, University of Cauca, Cauca, Colombia. He is the Director of the Telematics Engineering Group (GIT), University of Cauca. His research interests focus on service composition, data analysis and remote sensing. He promotes and supervises research on these topics within the frameworks of many local and national projects.



Apolinar Figueroa-Casas received the B.S degree in biology from the University of Cauca, Colombia, in 1982, the M.S. degree in ecology from the University of Barcelona, Spain, in 1986, and the Ph.D. in biological sciences from the University of Valencia, Spain, in 1999.

He has been a Visiting Professor at the National University of Rosario, Argentina, University of Valle, Colombia, Technological University of Pereira, Colombia, and Polytechnic University of Catalonia, Spain. He is currently a Full Professor at the Biology Department of the University of Cauca, Cauca, Colombia. He is the Founder and Director of the Environmental Studies Group, University of Cauca. He is also the Founder of the Ph.D. program in Environmental Studies of the University of Cauca. His research interest include Ecology of Andean systems, Ecosistemología, anthropic ecosystem processes, and Remote Sensing. He has led projects in the framework of sustainability environmental financed by the Colombian government and international cooperation. He is currently the Scientific Director of RICCLISA program (Inter-institutional Network for Climate Change and Food Security) funded by COLCIENCIAS.