Title: A Novel Approach to the Unsupervised Update of Land-Cover Maps by Classification of Time Series of Multispectral Images

This paper appears in: IEEE Transactions on Geoscience and Remote Sensing

Date of Publication: 2019

Author(s): Claudia Paris, Lorenzo Bruzzone, Diego Fernández-Prieto

Volume:

Page(s):

DOI: 10.1109/TGRS.2018.2890404

A Novel Approach to the Unsupervised Update of Land-Cover Maps by Classification of Time Series of Multispectral Images

Claudia Paris, *Member, IEEE*, Lorenzo Bruzzone, *Fellow, IEEE*, Diego Fernández-Prieto

***Abstract*—This paper presents an unsupervised approach that extracts reliable labeled units from outdated maps to update them using time series (TS) of recent multispectral (MS) images. The method assumes that: (1) the source of the map is unknown and may be different from remote sensing (RS) data; (2) no ground truth is available; (3) the map is provided at polygon level, where the polygon label represents the dominant class; and (4) the map legend can be converted into a set of classes discriminable with the TS of images (i.e., no land-use classes that require manual analysis are considered). First, the outdated map is adapted to the spatial and spectral properties of the MS images. Then, the method identifies reliable labeled units in an unsupervised way by a two-step procedure: (1) a clustering analysis performed at polygon level to detect samples correctly associated to their labels, and (2) a consistency analysis to discard polygons far from the distribution of the related land-cover class (i.e., having high probability of being mislabeled). Finally, the map is updated by classifying the recent TS of MS image with an ensemble of classifiers trained using only the reference data derived from the map. The experimental results obtained updating the 2012 Corine Land Cover (CLC) and the GlobLand30 in Trentino Alto Adige (Italy) achieved 93.2% and 93.3% overall accuracy (OA) on the validation data set. The method increased the OA up to 18% and 11.5% with respect to the reference methods on the 2012 CLC and the GlobLand30, respectively.**

***Index Terms*—Automatic classification, land-cover maps update, (RS) remote sensing, Sentinel 2, (TS) time series, unsupervised methods.**

I. Introduction

THE availability of remote sensing (RS) satellites optical images such as Sentinel 2 or Landsat 8 guarantee a constant way for monitoring the Earth' surface at high temporal resolution. This leads to dense time series (TS) of images at global level, which provide informative multitemporal data for long-term environmental analysis [1]. Moreover, several land-cover/ land-use maps are now available at national, continental or global level due to their importance for commercial and environmental monitoring and planning [2]–[4]. The capability to updating existing thematic maps using recently acquired optical images is extremely interesting from the operational view point. However, keeping cartographic products regularly updated is a complex and time consuming task, which typically requires ground reference data.

Supervised learning approaches represent the best solution to update land-cover maps due to the possibility of generating accurate classification results [5]–[8]. Recently, Valero *et al* [7] defined a supervised classification system to generate binary "annual-cropland/no-annual-cropland" maps produced for monitoring crop growing condition using TS of Sentinel 2 images. Due to the availability of informative ground truth, the method achieves 90% of accuracy at the end of the agricultural season. In [8], Belgiu *et al* proposed a system architecture for cropland mapping based on the availability of in-situ surveys. A time-weighted dynamic time warping method is employed to reconstruct the temporal spectral behaviour of the crop types from a TS of Sentinel 2 images to accurately distinguish different cultivations. The method tested in three different study areas (Romania, Italy and USA) achieves an Overall Accuracy (OA) ranging from 78.05% to 96.19%. However, these approaches need labeled data for the image to be classified, which are expensive and time-consuming to collect. Moreover, in RS application it is not feasible to assume ground truth available every time a land-cover map has to be updated.

To solve this problem, a large effort has been devoted to the development of semisupervised methods that rely on the availability of reference data for an image previously acquired over the same area (source image) [9]–[13]. The main idea of these approaches is to include in the classifier (pre-trained on the source image) unlabeled samples extracted from the target image, in order to adjust the discriminant function to the new data. In [10] Zanotta *et al* take advantage of the repetitive nature of recurrent changes to adapt the change detection rule computed for a given pair of multitemporal images (source domain) to a new multitemporal pair of images (target domain) expected to have the same kind of change (e.g., deforestation). The method achieves a 91.5% detection rate having a 7.5% of false alarms in the worst case. In [14], the temporal correlation between multispectral (MS) images belonging to the same TS is employed to update the parameters of the maximum likelihood classifier trained on the reference image. Experimental results conducted on Landsat 5 images demonstrate the effectiveness of the method that achieved an OA of 90.97%. This approach is further generalized in [15] in the framework of the Bayesian rule for cascade classification considering maximum likelihood and neural-network classifiers. The method, tested on Landsat 5 data, generates an updated land-cover map characterized by an OA of 95.54%. In [16], Bahirat *et al* address the case of multitemporal acquisitions characterized by different sets of classes. Graph-based methods are used to connect labeled and unlabeled samples according to their similarity [17]–[21]. The topological structure of the graph naturally injects unlabeled samples in land-cover classes under the assumption of consis-

tency (i.e., nearby points should belong to the same class) [17]. Then, the labels of the unchanged reference samples are propagated to generate a training set for the image to classify. The method achieves an OA higher than 85% both on a multitemporal medium resolution data set (Landsat 5) and a multitemporal high resolution data set (Quickbird).

Although these strategies can be effective in enlarging small training data set, to correct initial conditions is necessary to guarantee accurate classification results. Semisupervised methods are indeed very sensitive to the initial model assumptions since bad matching of pattern structure can lead in degradation of classifier performances. To solve this problem, few methods in the literature propose strategies to check the reliability of the unlabeled samples extracted from the target domain [11], [22]. In [11], Bruzzone *et al* propose a circular validation strategy which is employed to assess the reliability of the classification result obtained. The method achieves on a multitemporal Landsat 5 data set an OA of almost 91%. In [22], Demir *et al* propose a change detection based transfer learning method for generating a reference data set without any labeled data. First, the method detects changes occurred between the MS image to be classified and the one where ground truth is available. Then, the labels of the unchanged reference samples are propagated to generate a training set for the image to classify. The method, tested on three multitemporal data sets including Landsat 5, Landsat 7 and Quickbird MS images, reaches OA higher than 85% by classifying the second image with the unchanged reference samples. Besides the possibility of increasing the reliability of the semisupervised classification result, all these studies assume that: (i) there are reliable reference data available for at least one image of the TS, and (ii) the same type of RS data are used to perform the update. Although these assumptions are reasonable in many real applications, they are no more affordable when updating land-cover maps at national or global scale.

To relax the TS constraint, in [23] Inglada *et al* propose an operational work-flow to generate an automatic training set in an unsupervised way by combining several cartographic products to classify long TS of Landsat 8 images. Since many mislabeled samples may be included in the training set, a classifier tolerant to noise [24] is selected to generate the land-cover map. Although the proposed work-flow allows a fast delivery of maps at country level, the quality of the training samples is not sufficient to obtain high classification accuracy for all the classes. Classes such as woody moorlands, beaches and vineyards obtain a classification accuracy lower than 50% due to the availability of few training samples. The impact of training label noise on classification perfomances for land-cover mapping is evaluated in [25]. The authors compare the results obtained with Support Vector Machine (SVM) and Random Forest classifiers on a TS of Landsat 8 and Satellite Pour l'Observation de la Terre (SPOT) 4 images acquired during one year. The results show that when the number of wrongly labeled samples is lower than 30% both classifiers are little influenced by the training label noise (i.e., the OA% decreases less than 10%). In contrast, their performances drop down for higher noise levels.

To reduce the amount of class label noise, few works introduce approaches to automatically check the labeled samples extracted from existing thematic maps [26], [27]. In [27], Radoux *et al* randomly select training samples from existing land-cover maps to classify 300 m Medium Resolution Imaging Spectrometer (MERIS) data acquired from 2008 to 2012. A spatial filtering analysis is performed to discard samples on the boundaries between different land-cover areas since they are more often incorrectly labeled due to inaccurate geolocation. Then, a "cleaning" process excludes the outliers from the distribution of the spectral signatures using a probabilistic iterative trimming [28]. The same procedure is applied in [29], where Matton *et al* present a method for mapping "annual-cropland/no-annual-cropland" areas based on the normalized vegetation index temporal behavior of the pixels. To rely on a dense TS of images SPOT 4 images are integrated to Landsat 8 data. The OA obtained for the eight sites distributed throughout the world ranges from 71% to 99%. A similar spectral filtering analysis is presented in [26] where an editing process is proposed to remove pixels identified as outliers in the spectral domain for each class. Silva *et al* extract the labeled samples from the 2006 Corine Land Cover (CLC) map to train a classifier with MERIS data. Although the OA is 89%, some classes characterized by very poor accuracy (lower than 60%). Thus, even though these outlier removal strategies increase the probability of selecting reliable samples from the map, their main drawback is the risk of removing informative samples [24] that strongly affecting the generalization capability of the classifier. Moreover, the method updates an existing land-cover map using MERIS images, characterized by a spatial resolution of 300 m. When increasing the spatial resolution at 10 m (e.g., Sentinel 2 data), the classifier is even more sensitive to the quality of the training set [30].

A. Motivation

From this brief analysis of the literature, it turns out that a large effort has been devoted to develop methods that reuse existing reference data to classify recent MS images. In contrast, little effort has been done to extract reference data directly from the map that needs to be updated. However, the source of many thematic products is not necessarily a satellite image belonging to a TS. Most of the existing thematic products have been generated by multiple sources (e.g., photo-interpretation, ancillary cartographic data) [3], [31], [32]. Moreover, the recent launch of optical RS satellites such as Sentinel 2 guarantees a constant way for monitoring the Earth' surface at high temporal resolution (5 days upon complete constellation). Sentinel 2 increases the potential of land-cover maps updating rate establishing a detailed analysis of dynamic phenomena also in areas that are typically affected by cloud cover [1]. In this framework it is necessary to develop methods being able to extract reliable reference data from the available thematic maps to regularly update them in an unsupervised way (i.e., assuming that no new labeled data are available). However, many challenges arise when extracting labeled samples from existing thematic products. Because of changes occurred on the ground and possible classification errors, wrong labels may be inherited from the map. Moreover, land-cover maps are

typically provided at polygon level, where the polygon label is assigned through a majority rule. Therefore, the samples belonging to the minority classes present in the polygon are associated with a wrong semantic label. Finally, often the map legend of is not optimized for the spectral properties of the MS image used to perform the update [33], [34].

B. Contribution

In this paper we propose an unsupervised method that aims to automatically extract reliable training samples from an available thematic map (hereafter referred as "source map"), using the multitemporal spectral information provided by TS of recent MS images. The method is based on three main components: (1) a preprocessing phase, where the source map is adapted to the peculiarities of the available MS data, (2) an unsupervised training set identification phase, where samples having high probability to be still associated to correct labels are detected, and (3) a land-cover map update phase, where a pool of SVMs is trained with the extracted labeled units to classify recent TS of MS images. Differently from the literature, no spectral filtering at pixel level is used to remove the outliers, since we aim at representing the distribution of land-cover classes. To this end, first a data-driven clustering analysis is performed at polygon level to automatically detect samples associated to the dominant land-cover class. Then, a consistency analysis of polygons belonging to the same class is performed to discard polygons far from the distribution of the considered class (i.e., having high probability to be mislabeled).

The proposed method extends and enhances the work described in [35] by: (1) removing the consistency check analysis at pixel level which may discard informative training samples, (2) introducing an ensemble-based classification step that mitigates the problem of including possible mislabeled samples in the reference data extracted from the map, and (3) significantly extending the experimental analysis. The main novelties of the proposed method are: (1) the preprocessing step that aims at reducing the discrepancy between the thematic product and the MS images, and (2) the identification of reliable and informative map samples by an automatic procedure tailored to the map properties.

Note that in the considered method we exploit a TS of MS images to enhance the capability of discriminating complex land-cover classes such as the agricultural ones [36]. However, the proposed method is flexible to handle any RS data (e.g., synthetic aperture radar polarimetry data which are suitable for land-use/land-cover classification [37]–[40]) under the assumption that the considered data allow the discrimination of the set of classes present in the map legend.

C. Outline

The rest of the paper is organized as follow. Section II introduces the notation and the proposed system architecture by illustrating all the phases of the method in detail. Section III describes the data set in terms of thematic products and optical satellite images employed to perform the update. Section IV discusses the experimental results obtained. Finally, Section V draws the conclusion of the paper and presents possible future developments.

II. Proposed Land-Cover Map Updating Method

A. Problem Formulation and Notation

The aim of the proposed approach is to exploit the information provided by outdated land-cover maps to extract reliable reference data. Due to the lack of ground truth and knowledge on the source of the thematic map, we are addressing an extremely ill-posed problem. In such a complex scenario, the method is based on the following constrained assumptions: i) the semantic of the classes of the thematic product can be converted in an exhaustive set of classes discriminable with the spectral content of the available TS of MS images; and ii) the available land-cover map is provided at polygon level labeled according to a majority rule criterion (i.e., the associated label represent the dominant class of the polygon). Note that the first assumption clarifies that no land-use classes are included in the map legend, since they require intensive photo-interpretation analysis and cannot be discriminated by using only the spectral properties of the TS of images.

Let us define the notation used in this paper. MS images are usually represented as 3-D arrays (e.g., tensor flow or data cube). For notation convenience, the representation followed in this paper considers the MS data as a 2-D matrix, where each column corresponds to a spectral band, containing the lexicographically ordered pixels of that band. Bold lowercase denotes vectors (e.g., $\mathbf{x}$), while bold uppercase denotes matrices (e.g., $\mathbf{X}$).
Let $\mathbf{X}_q \in \mathbb{R}^{B\times d}$ be the MS image acquired at time t_q and made up of d spectral channels and B pixels. Accordingly, the TS of co-registered MS images acquired on the same area at different times can be represented as the horizontal concatenation of the corresponding 2-D matrices. Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_Q)$ be the TS of MS images, where $\mathbf{X} \in \mathbb{R}^{B\times n}$ has B pixels and $n = d \cdot Q$ spectral channels. Note that the TS of images can be also represented as vertical concatenation of spectral vectors $\mathbf{X} = (\mathbf{x}_1; \mathbf{x}_2; \cdots; \mathbf{x}_B)$, where each row $\mathbf{x}_b \in \mathbb{R}^{1\times n}$ represents the multitemporal spectral values associated to the bth pixel (i.e., elements from 1 to d are the spectral values of $\mathbf{X}_1$, elements from $d+1$ to $2d$ are the spectral values of $\mathbf{X}_2$, etc.).
Let $\mathbf{M} = \{\Omega, \mathcal{P}\}$ be the source map composed of a set of land-cover classes $\Omega = \{\omega_u\}_{u=1}^{U}$ and a set of polygons $\mathcal{P} = \{\lambda_j\}_j$, where each polygon is associated to a unique label.

Fig. 1 shows the system architecture of the proposed land-cover map updating method. The approach is separated into three main phases: i) preprocessing; ii) automatic identification of the *pseudo training sets*; and iii) update of the land-cover map by classification of recent TS of MS images. In the following, we describe in detail each phase of the proposed method.

B. Preprocessing

The first phase of the proposed method, *preprocessing*, seeks to prepare the TS of MS images, and adapt the thematic map to the properties of the considered MS images. First,

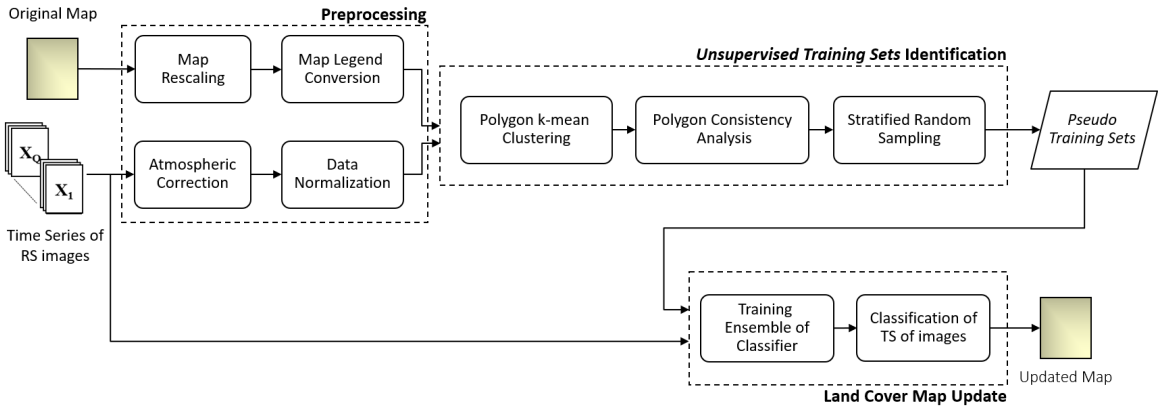


Fig. 1: System architecture of the proposed automatic land-cover map updating approach. The approach is separated into three main phases: (i) preprocessing, (ii) *unsupervised training sets* identification, and (iii) land-cover map update.

the images are atmospherically corrected and the clouds are masked, thus leading to surface reflectance products. In the experimental analysis of this paper Sentinel 2 images are employed to update the thematic products since these data are particularly suited to frequently update of land-cover maps due to the combination of high geometric resolution, novel spectral capabilities, large swath width and short revisit time [1]. However, any other MS optical data can be employed (e.g., Landsat 8). The atmospheric correction and cloud detection was performed using the Sen2cor tool provided by the European Space Agency (ESA) [41], while cloud gaps were filled according to [42]. Due to the peculiar multi-resolution property of S2 images [1], a cubic interpolation is used to match the spatial resolution of the 20m bands to 10 m for the entire TS. Finally, a radiometric normalization is eventually applied to the interpolated images so that each spectral band is rescaled between zero and one.

In the land-cover map preprocessing, first we rescale the thematic product at the spatial resolution of the highest resolution channel of Sentinel 2. Then, the semantic of the legend is converted into a set of classes discriminable considering the spectral content of the image TS. To this end, we referred to the Land Cover Classification System (LCCS) [43], which is a standard common land-cover language for translating and comparing existing legends. In the literature several works make use of legend harmonization approaches due to the need of comparing existing thematic products and integrating multi-spatial/spectral and multi-semantic data sets [43]–[47]. LCCS is a comprehensive, a priori world-wide reference system for land-cover created for mapping. In order to have a high level of flexibility (describe land-cover features at any scale or level of detail), LCCS standardizes terminology and the attributes used to define thematic classes instead of defining a single universal land-cover legend. In particular, the system uses a set of independent diagnostic criteria that allow correlation with existing classifications and legends independent of the scale or means used to map. Let $\Omega_c = \{\omega_u\}_{u=1}^{U_c}$ be the subset of $U_c \leq U$ classes converted from the original map legend. Let $\mathbf{l} = (l_1, l_2, \cdots, l_B)^T \in \mathbb{R}^{B\times 1}$ and $\mathbf{p} = (p_1, p_2, \cdots, p_B)^T \in \mathbb{R}^{B\times 1}$ be the vectors representing the land-cover classes and the polygons associated to the multitemporal vectors of the TS, respectively, with $\mathbf{l} \in \Omega_c$ and $\mathbf{p} \in \mathcal{P}$. Therefore, l_b and p_b represent the label and the polygon associated to the bth spectral vector $\mathbf{x}_b$, respectively.

C. Unsupervised Training Sets Identification

The second phase of the proposed method, *unsupervised training sets identification*, aims to automatically extract from the map samples having high probability to be still associated with valid labels. When extracting reference data from existing thematic products, many wrong labels may be inherited due to: i) changes occurred on the ground, ii) classification errors present in the map, and iii) mislabeled samples due to the polygon aggregation. To address all these issues, the proposed method aims to first analyze each polygon to detect the samples belonging to the dominant class (i.e., the one having the highest probability to be correctly associated to the polygon label). Then, it analyzes the land-cover distribution in order to keep only the most reliable units. We assume that most of areas in the map remains unchanged, which is a typical scenario when updating thematic products representing widespread land-cover classes such as Urban, Crops or Forest. In contrast, the update of specific thematic products such as the agricultural ones cannot be updated with the proposed method, since most of the pixels would be associated to wrong labels. Moreover, we assume that the multitemporal spectral signatures of samples belonging to the same land-cover class are similar. Note that even though the method can be applied on individual RS data, the valuable information provided by TS of images allows the accurate characterization of the land-cover classes from the spectral view point. Finally, we train an ensemble of SVM classifiers (however any classification technique can be used). Each SVM is trained independently using the randomly chosen training samples via a bootstrap technique. The final decision is defined as the result of a simple majority voting combination rule.

For sake of simplicity, let us focus the attention on the jth polygon $\mathbf{P}_j = \mathbf{X}_{(p_i=\lambda_j,:)} = (\mathbf{x}_1; \mathbf{x}_2; \cdots, \mathbf{x}_{B_j}) \in \mathbb{R}^{B_j\times n}$ associated to B_j pixels and characterized by the n spectral channels of the TS. Let us assume that the polygon label is

(a) (b) (c) (d)

Fig. 2: Qualitative example of polygon k-means clustering result: (a) original polygon associated to the "crop" label, (b) dominant land-cover class detected $\mathbf{C}_1$, (c) first minor class detected $\mathbf{C}_2$ (road), and (d) second minor class detected $\mathbf{C}_3$ (river).

ω_u, i.e., $\mathbf{l}_{(p_i=\lambda_j,1)} = \omega_u$. To identify the multitemporal vectors belonging to $\mathbf{P}_j$ correctly associated to ω_u (i.e., the dominant class of the polygon), we partition the polygon samples into K_j clusters $\{\mathbf{C}_j^1, \mathbf{C}_j^2, \ldots, \mathbf{C}_j^{K_j}\}$ according to their multitemporal spectral similarity, where $\mathbf{C}_j^k = (\mathbf{x}_1; \mathbf{x}_2; \cdots, \mathbf{x}_{N_j^k})$. To this end we considered for simplicity a k-means clustering algorithm, which is efficient from the computational view point [48]. However, any other clustering technique can be used. Starting from a random initial set of centroids, the k-means clustering algorithm associates each multitemporal vector $\mathbf{x}_b \in \mathbf{P}_j$ to the closest centroid based on the euclidean distance metric [49]. By progressively adjusting the centroid positions, the global minimum is reached by minimizing:

$$\sum_{k=1}^{K_j} \sum_{\mathbf{x}_b \in \mathbf{C}_j^k} ||\mathbf{x}_b - \mathbf{m}_k||^2 \tag{1}$$

where $\mathbf{m}_k$ is the centroid of the cluster $\mathbf{C}_j^k$. Since a different number of classes may be present in different polygons, the proposed approach automatically estimates the number of K_j classes present in the jth polygon $\mathbf{P}_j$. This data-driven analysis allows us to adapt the cluster detection to each polygon, thus increasing the probability of accurately detecting the dominant cluster of the polygon. To identify K_j, the proposed method uses the Calinski Harabasz (CH) Index, which is widely used in the literature to detect the number of clusters in an unsupervised way [50]–[52]. The number of clusters K_j is estimated by considering the average between- and within-cluster sum of squares as follows:

$$\text{CH} = \left[\frac{\sum_{k=1}^{K_j} N_j^k ||\mathbf{m}_k - \mathbf{m}||^2}{K_j - 1} \right] \Big/ \left[\frac{\sum_{k=1}^{K_j} \sum_{b=1}^{B_j} ||\mathbf{x}_b - \mathbf{m}_k||^2}{B_j - K_j} \right] \tag{2}$$

where $\mathbf{m}$ is the centroid of $\mathbf{P}_j$. Due to the spectral similarity of samples belonging to the same class, the algorithm automatically assigns pixels belonging to the dominant class of the polygon to the same cluster. Based on the majority decision rule, it is reasonable to assume that the cluster having the highest number of samples represents the dominant polygon class. Let $\mathbf{C}_j^{max}$ be the dominant cluster of $\mathbf{P}_j$. Fig. 2 shows a qualitative example of the clustering result obtained on a polygon associated to the "crop" label where $K_j = 3$. Due to the automatic detection of the number of natural classes present in the polygon, the method distinguishes the dominant class "crop" $\mathbf{C}_1$, from the minor classes "road" $\mathbf{C}_2$ and "river" $\mathbf{C}_3$.

Although the cluster analysis allows us to solve the aggregation problem of the polygon, we cannot rely completely on the considered thematic product since wrong labeled polygons may be present in the map (i.e., polygons where changes happened to the ground or associated to wrong labels). To address this issue, instead of considering an outlier filtering approach at sample level, the proposed method performs a polygon consistency analysis, to remove those polygons which are far from the class distribution. This condition allows us to discard samples associated to wrong labels without removing informative samples. Let us focus the attention on the cluster associated uth land-cover class ω_u. Let N_u be the number of polygons associated to ω_u, i.e., $\{\mathbf{P}_1, \mathbf{P}_2, \cdots, \mathbf{P}_{N_u}\}$. At the end of the clustering procedure, each polygon is represented by its dominant cluster $\{\mathbf{C}_1^{max}, \mathbf{C}_2^{max}, \cdots, \mathbf{C}_{N_u}^{max}\}$. For each cluster $\mathbf{C}_j^{max}$, we compute its Bhattacharyya distance from the whole set of clusters associated to ω_u as follows [53]:

$$B_{ju} = \frac{1}{8}(\boldsymbol{\mu}_u - \boldsymbol{\mu}_j)^T \left(\frac{\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_j}{2} \right)^{-1} (\boldsymbol{\mu}_u - \boldsymbol{\mu}_j) + \\ + \frac{1}{2} ln \left(\frac{|\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_j|/2}{\sqrt{|\boldsymbol{\Sigma}_u||\boldsymbol{\Sigma}_j|}} \right) \tag{3}$$

where $\boldsymbol{\mu}_u$ and $\boldsymbol{\Sigma}_u$ represent the mean and the covariance matrix of the set of multitemporal spectral vector $\mathbf{x}_b \in \{\mathbf{C}_1^{max}, \mathbf{C}_2^{max}, \cdots, \mathbf{C}_{N_u}^{max}\}$, while $\boldsymbol{\mu}_j$ and $\boldsymbol{\Sigma}_j$ are the mean and the covariance matrix of cluster $\mathbf{C}_j^{max}$. Only the most reliable clusters, i.e., those having distance from the cluster distribution smaller than the 65th percentile of the cluster distances, are used to represent ω_u. This condition allows us to extract from the thematic map a pool of reliable and informative samples to be used for generating a set of *pseudo*

training sets $\{T_1, T_2, \cdots, T_S\}$, where $T_s = \{(\mathbf{x}_b, y_b)\}_b$ is the sth training set having $\mathbf{x}_b \in \mathbb{R}^{1\times n}$ and $y_b \in \Omega_c$. Note that we called them *pseudo*, since they are not completely reliable as the ones made up of ground reference data.The *pseudo training sets* are generated without replacement (i.e., each sample belongs only to one *pseudo training set*) with the aim of training statistically independent classifiers which are expected to generate uncorrelated classification errors. A stratified random sampling design is considered to generate balanced training sets proportionate to the original prior probabilities of the land-cover classes $\Omega_c = \{\omega_u\}_{u=1}^{U_c}$ according to the thematic product. The number of pixels per land-cover class present in the thematic map is used as strata to determine the number of samples to select per class [54].

D. Land-Cover Map Update

In the last phase of the proposed method, we perform the update to the land-cover map by classifying the TS of MS images with the set of the *pseudo training sets* generated in the previous phase. First, a feature selection step is performed to detect the feature subspace where the land-cover classes are more discriminable. In the considered implementation, we exploit a Sequential Forward Floating Selection (SFFS) method based on the Jeffreys-Matusita distance as separability criterion [55], [56]. Then, an ensemble of SVMs classifiers is used to classify the TS of images. SVMs with Gaussian Radial Basis Function (RBF) kernels is considered [57] because of their capability of dealing with noisy samples in a robust way and to produce sparse solutions [58]. Please note that any other feature selection method or classifier can be used in the considered system. A bagging technique [59] is adopted to construct the ensemble, where several SVMs are trained independently via a bootstrap method. Since the classifiers are statistically uncorrelated (ideally independent), their errors are expected to be uncorrelated. Accordingly, the result of combination is expected to improve the reliability of the classification results. Note that also sophisticated techniques can be used to define many combination rules of the ensemble of classifiers. Here, to illustrate the effectiveness of the proposed approach we use a very simple majority voting decision rule. Let $\{f_1, f_2, \cdots, f_S\}$ be the decision functions of the ensemble of SVM classifiers trained using the $\{T_1, T_2, \cdots, T_S\}$ and let $\#\{f_s(\mathbf{x}_b) = \omega_u\}$ be the number of SVMs whose decision is the class ω_u for the pixel $\mathbf{x}_b$. The majority voting decision of the ensemble of SVMs for $\mathbf{x}_b$ is given by:

$$f_{mv} = \underset{u\in[1,\cdots,U_c]}{\operatorname{argmax}} \left(\#\{f_m(\mathbf{x}_b) = \omega_u\}\right) \tag{4}$$

Since the proposed method is completely unsupervised, the SVM classifiers trained are considered as weak learners. Thus, some wrongly labeled samples may be included in the *pseudo training sets*. However, the ensemble of classifiers allows us to mitigate this problem, thus leading to more reliable classification result with respect to those obtained by using one classifier [60].

III. Data set Description

To assess the effectiveness of the proposed method, we perform the update of both 2012 CLC map and 2010 GlobLand30. The 2012 CLC map produced at continental level (25 European Union member states) by the European Environment Agency [2], is provided with a spatial resolution of 100 m for the raster version and of 30 m of the vector version (used in the experiments). The map was generated with a minimum mapping unit of 25 ha, thus large polygons characterized this thematic product. The classification scheme of the CLC map presents a hierarchical structure of 44 classes. The first level is composed of 5 classes which correspond to the main categories of the land-cover/land-use (artificial areas, agricultural land, forests and semi-natural areas, wetlands, water surfaces). The second level (15 classes) covers physical and physiognomic entities at a higher level of detail (urban zones, forests, lakes, wetlands, crops, etc.), while in the third level presents all the detailed 44 classes. The production of this map was based on visual interpretation of optical/near-infrared satellite images (i.e., mainly Landsat imagery) and ancillary data such as aerial photograph, topographic maps or forestry maps [61].

The second thematic product is the 2010 GlobLand30 provided by the National Geomatics Center of China (NGCC) in the framework of the "Global Land Cover Mapping at Finer Resolution" project, one of the first global land-cover data sets provided at 30 m geometric resolution. To generate the map the collection and classification of more than 10,000 scenes was performed. The scenes were mainly acquired by the Landsat Thematic Mapper (TM), Landsat Thematic Enhanced TM plus (ETM+) and Chinese Environmental and Disaster (HJ-1) satellites [3]. GlobeLand30 is available in raster format at 30 m spatial resolution. The legend is based on 10 land-cover classes.

The satellite optical data used to perform the update of the maps are Sentinel 2 MS images. Although any MS data can be employed, Sentinel 2 data have a huge potential for this task, due to their spatial resolution (10 m in the visible range), their enhanced spectral capabilities (i.e., three bands in the "red edge"), a swath width of 290 km and their increased revisit time with respect to the past (5 days upon complete constellation). The considered study area, which extends for 1549 km^2 over the T32TPS Sentinel 2 granule, is located in the Trentino Alto Adige Region, Italy. To perform the land-cover map updating we considered a TS of four Sentinel 2 images acquired on the 27^{th} May 2017, 26^{th} June 2017, 18^{th} July 2017 and 21^{st} September 2017. The Sentinel 2 spectral bands considered are the ones acquired at 10m and 20m spatial resolution. Thus, the method considers 40 features (10 features per image) per pixel. To quantitatively evaluate the accuracy of the updated land-cover maps, we used a reference data set made up of 35678 samples manually labeled by photo-interpretation and distributed all over the region. The spatial distribution of the reference data is represented in Fig. 3, while Tab. I shows the number of samples divided per class.

IV. Experimental Results

In this section we present the experimental results obtained by updating both the 2012 CLC and the GlobLand30 maps

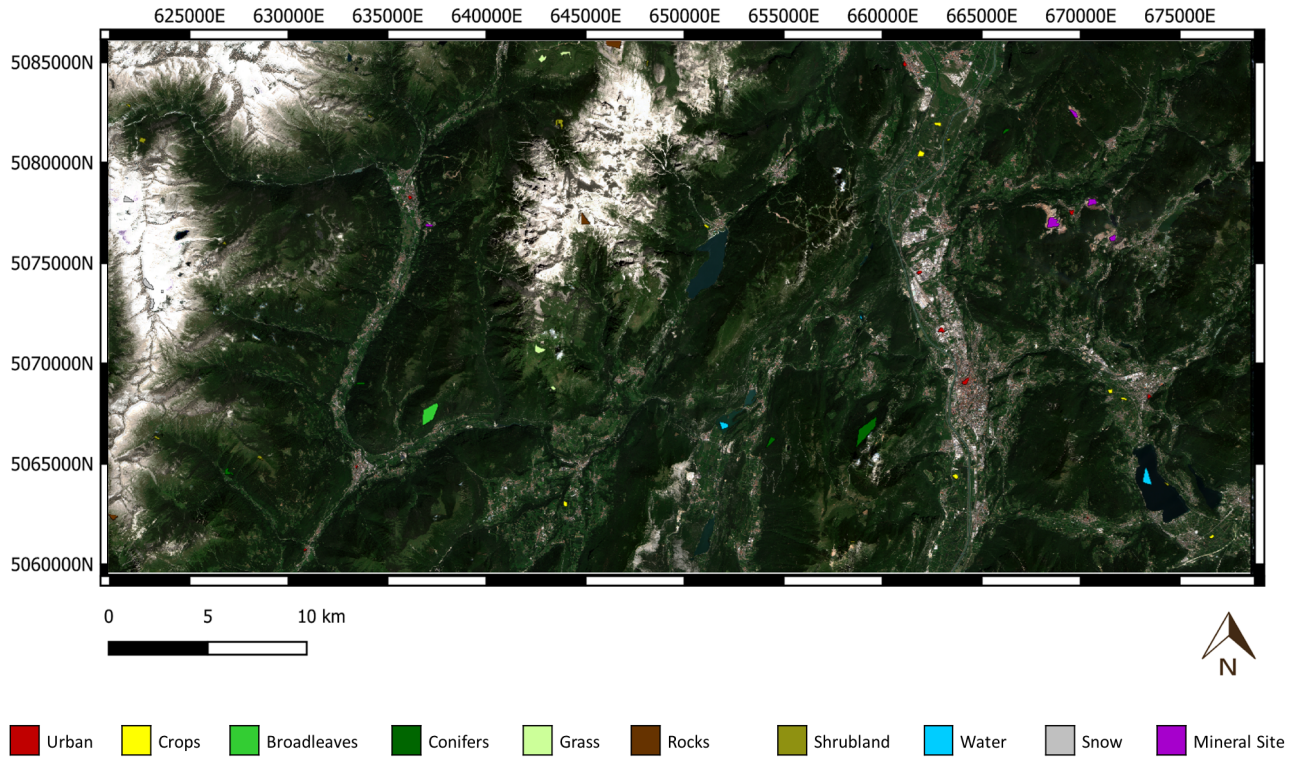


Fig. 3: True color representation of Sentinel 2 image acquired on the 18^{th} July 2017 with ground reference data (coordinates are reported in the UTM WGS84 32N system).

TABLE I: Ground reference data used for validating the results divided per class.

ID	Class	# Validation Samples
ω_1	Urban	2216
ω_2	Crops	2821
ω_3	Broadleaved	4741
ω_4	Conifers	7584
ω_5	Grass	2954
ω_6	Water	2858
ω_7	Rocks	4299
ω_8	Snow	2945
ω_9	Shrubland	1744
ω_{10}	Mineral Site	3516

using the proposed method. First, we present the experimental setup, by introducing the baseline methods used for comparison. Then, we describe in detail the legend conversion performed considering the spectral properties of Sentinel 2. We move to the analysis of the quality of the *pseudo training sets* extracted from the map by presenting both the clustering results obtained at polygon level and the statistical consistency analysis. Finally, the updated land-cover maps will be presented and discussed both from the qualitative and the quantitative view point.

A. Experimental Setup

The considered thematic products were generated by using multiple sources (several RS data, photo-interpretation, ancillary data, etc.). Thus, they cannot be updated by using any state-of-the-art methods which requires the availability of the image used to generate the map. To prove the effectiveness of the proposed method, we compared it with a tolerant noise classifier [25] and a standard outlier filtering approach [27]. In [25], Pelletier *et al* suggest to use Random Forest classifier, which does not require a demanding parameter tuning, when noise is present in the training set. In the considered implementation we follow the parameter setting suggested by the authors:

- number of trees to build equal to 200;
- number of input features randomly selected by each node equal to the square root of total number of features (in our case $\sqrt{40}$);
- maximum depth of the tree growth equal to 25;
- minimum number of instances in the node equal to 10.

In [27], Radoux *et al* first perform a spatial filtering analysis to discard samples on the boundaries between different land-cover classes. Then, a spectral filtering step excludes the outliers from the distribution of the spectral signatures using a probabilistic iterative trimming [28]. In the considered implementation, the polygon boundaries were discarded by performing a morphological erosion. The structural element employed is a disk having size 2 (i.e., removing a boundary of 20m around the polygon). To perform the spectral filtering step, the authors suggest to tune $\alpha \in [0.05, 0.1, 0.2]$. The results reported are the ones that achieve the best classification accuracy, obtained with $\alpha = 0.05$. The extracted training set is

used to classify the TS of Sentinel 2 images. The classification is performed using SVM classifier with RBF kernel. To detain a fair comparison with the proposed method, the SFFS is applied to remove redundant features, thus increasing the classification accuracy.
The proposed method classifies the TS of Sentinel 2 images with an ensemble of five SVMs with RBF kernel. For both the proposed method and [27], the optimal kernel parameters (i.e., the regularization parameter C and the spread of the kernel γ) were selected by a 3-fold cross-validation. In the proposed method case, the model selection is carried out for each SVM of the ensemble. The SFFS algorithm selected 20 features among the original 40 spectral channels for both the proposed and the baseline methods [27].

B. Legend Conversion

Tab. II shows the legend conversion used for the 2012 CLC map. The 2012 CLC legend was simplified to depict the widespread land-cover classes: "Urban", "Mineral Site","Crops", "Broadleaved", "Conifers", "Grass", "Shrubland", "Rock", "Snow" and "Water" according to the LCCS standard. The legend conversion allows us to exclude land-use classes such as "Airports", "Dump Sites" or "Sport and facilities" that cannot be discriminated by using the spectral properties of Sentinel 2. Moreover, we removed classes mixed from the semantic view point, such as "Sparsely Vegetated Area" or "Mixed Forest" which were defined in the CLC legend due to the need of assigning labels to aggregated polygons (i.e., minimum mapping unit of 25 ha). However, the updated map is generated at 10 m resolution, thus, the classification result obtained at pixels level strongly enhances the geometric details of the classification map (i.e., "Sparsely Vegetated Area" or "Mixed Forest" are replaced with "Broadleaved", "Conifers" or ' Grass") and does not require these mixed classes anymore. Note that since we are facing a complex and ill-posed problem, in the considered legend we are not including detailed land-cover classes such as "Vineyards" even though they can be discriminated by using the spectral properties of Sentinel 2. Thus, to generate a reliable updated product, we focused the attention on a simpler but enough informative class legend.

Tab. III shows the legend conversion used for the GlobLand30 map. All the land-cover classes present in the GlobLand30 legend can be discriminated by the spectral properties of Sentinel 2. Note that, the GlobLand30 legend includes also the land-cover classes "Wetland" and "Tundra". However, these classes are not present in the considered study area and thus, they are not reported in Tab. III. For a qualitative evaluation of the results obtained, the same colors used for the converted CLC legend are employed for the GlobLand30. Differently from the 2012 CLC, the polygons map is not provided together with the classification map. To overcome this issue, we automatically extract the polygons from the classification map by labeling the 4-connected components associated to the same label having at least 100 pixels.

TABLE II: Translation of the 2012 CLC legend into the proposed target legend for the considered test area (Trentino Alto Adige Region).

	CLC Class		**Target Class**
	Continuous urban fabric		Urban
	Discontinuous urban fabric		-
	Industrial or commercial units		Urban
	Airports		Urban
	Mineral extraction sites		Mineral Site
	Sport and leisure facilities		-
	Non-irrigated arable land		Crops
	Vineyards		Crops
	Fruit trees and berry plantations		Crops
	Olive groves		Crops
	Annual permanent crops		Crops
	Complex cultivation patterns		Crops
	Land with agriculture and vegetation		-
	Broadleaved forest		Broadleaved
	Coniferous forest		Conifers
	Mixed forest		-
	Natural grasslands		Grass
	Moors and heathland		Shrubland
	Transitional woodland-shrub		-
	Bare rocks		Rocks
	Sparsely vegetated areas		-
	Glaciers and perpetual snow		Snow
	Water courses		Water
	Water bodies		Water

TABLE III: Translation of the GlobLand30 legend into the proposed target legend for the considered test area (Trentino Alto Adige Region).

	GlobLand30 Class		**Target Class**
	Artificial Surfaces		Urban
	Cultivated Land		Crops
	Forest		Forest
	Shrubland		Shrubland
	Bare land		Bare land
	Grassland		Grass
	Permanent Snow and Ice		Snow
	Water		Water

C. Training Set Analysis

Fig. 4 shows examples of polygons identified as reliable by the proposed method (i.e., having a Bhattacharyya distance smaller than the 65th percentile of the cluster distances from the land-cover class distribution). Moreover, the automatic clustering results are reported per polygon. As one can notice, the data-driven clustering analysis is able to accurately detect the dominant land-cover class present in the polygon. While in the "Urban" polygon (Fig. 4a and Fig. 4g) only the samples belonging to artificial surfaces (such as buildings or roads) are selected, for the "Crop" polygon (Fig. 4b and Fig. 4g)

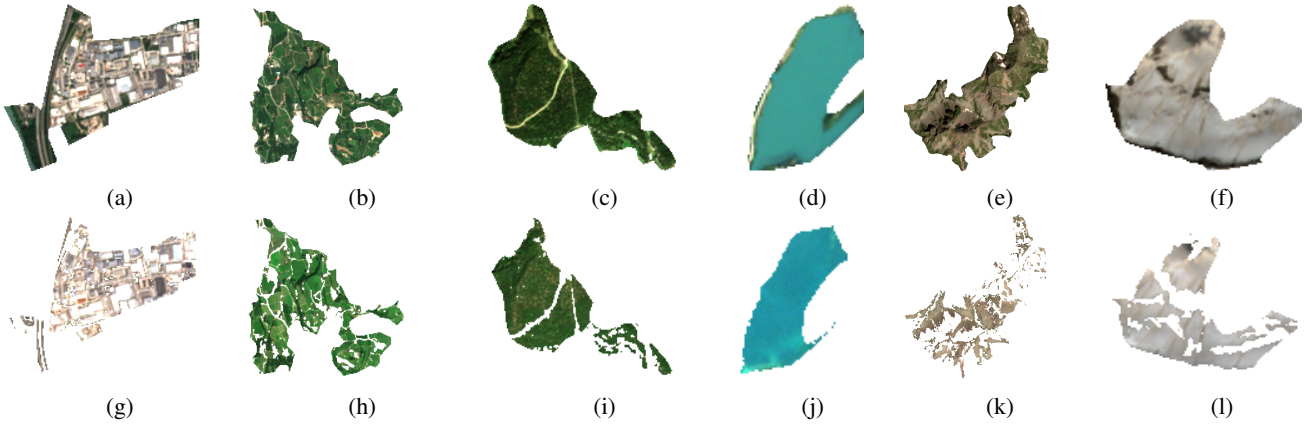


Fig. 4: Polygon consistency analysis: (a)-(e) examples of polygon selected to represent the land-cover class (Bhattacharyya distance $\leq$ 65th percentile of the cluster distances). (g)-(l) Clustering results obtained on the considered polygons (i.e., samples considered to generate the land-cover class distribution). The represented land-cover classes are: (a),(g) "Urban"; (b),(h) "Crops"; (c),(i) "Broadleaved"; (d),(j) "Water"; (e),(k) "Rock"; and (f),(l) "Snow".

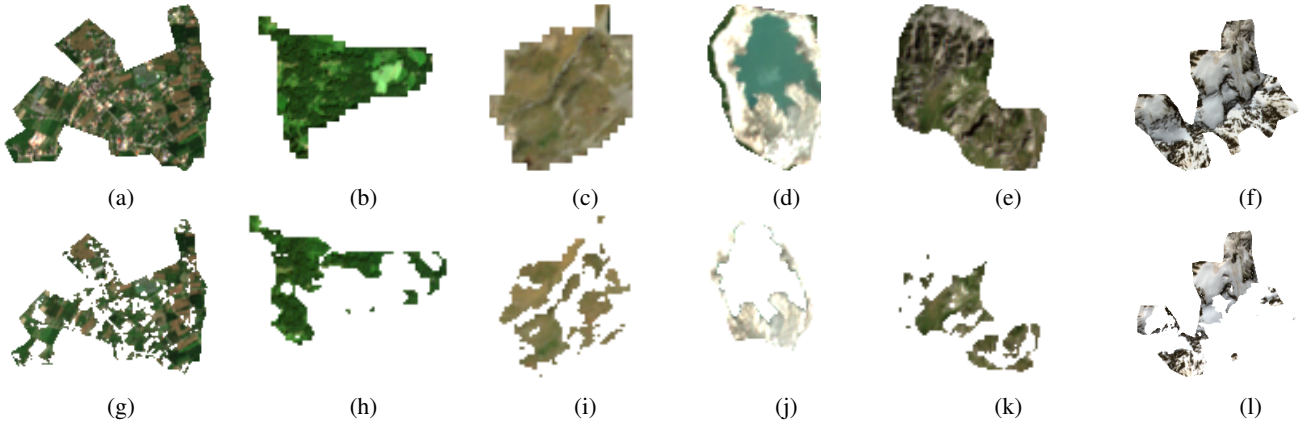


Fig. 5: Polygon consistency analysis. (a)-(e) Examples of discarded polygons (Bhattacharyya distance $>$ 65th percentile of the cluster distances). (g)-(l) Clustering results obtained on the considered polygons. According to the initial thematic product, the polygon are associated to the following land-cover classes: (a),(g) "Urban"; (b),(h) "Crops"; (c),(i) "Broadleaved"; (d),(j) "Water"; (e),(k) "Rock"; and (f),(l) "Snow".

the samples of rocks and roads are discarded. Similarly, in the polygon associated to the "Rock" label (Fig. 4e and Fig. 4k) the samples belonging to the grass are removed. Due to the majority rule criterion, when no changes occurred on the ground, the polygon label is correctly associated to the dominant class present in the scene. By accurately removing the samples that are not associated to the dominant class of the polygon, we strongly improve the capability of the method of selecting reliable units to generate accurate land-cover class distributions. However, results show that some polygons are not correctly associated to their labels.

Fig. 5 shows some examples of polygon discarded in the polygon consistency analysis step. In the "Urban" and "Water" polygon the dominant land-cover class is no more represented by artificial surfaces (Fig. 5a and Fig. 5g) or by the lake (Fig. 5d and Fig. 5j), respectively. Moreover, due to deforestation, the "Broadleaved" polygon is now characterized by the presence of bare land (Fig. 5c and Fig. 5i). Finally, one of the most critical land-cover class to deal with is the "Snow" due to its seasonal variability. Although in the considered thematic products the "Snow" class is defined as permanent snow, polygons associated to this label may show the presence of the land-cover usually covered by the snow, thus affecting the majority rule criterion. However, the spectral response of all these polygons is far from their land-cover class distribution and thus, they can be discarded during the polygon consistency analysis. Note that, differently from the spectral filtering strategies performed at sample level, the considered analysis does not remove informative samples since we are working at polygon level. Moreover, the adaptive tuning of the threshold to each land-cover class distribution (i.e., the 65th percentile) automatically excludes the polygons far from the distribution according to the spectral properties of the class.

D. Results of Land-Cover Map Update

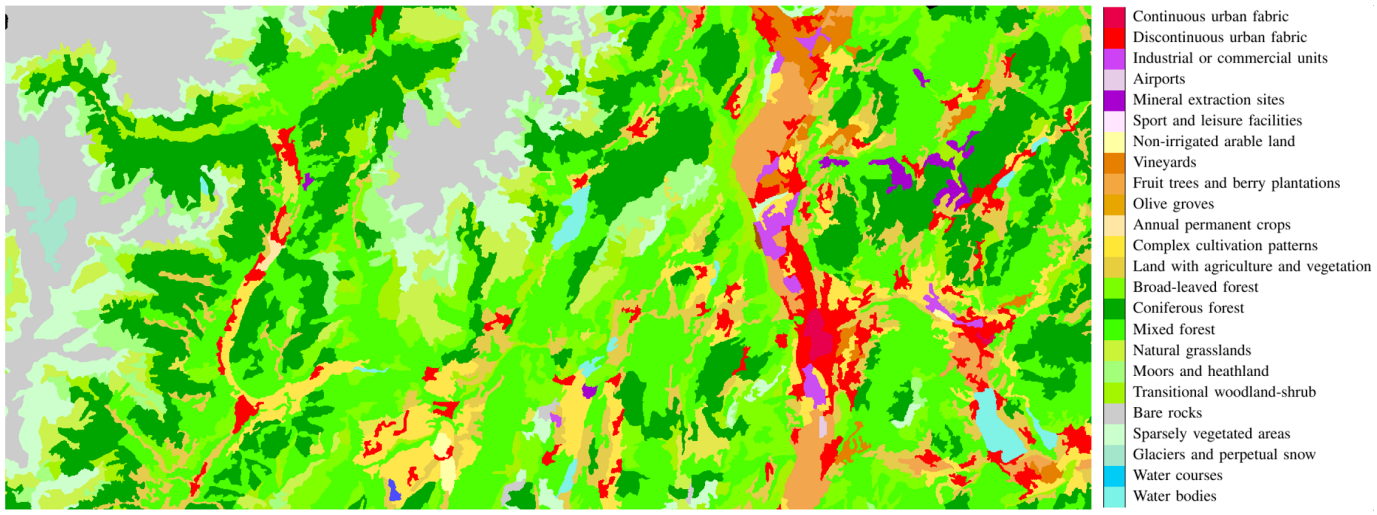
Tab. IV shows the comparison between the accuracies obtained when updating the 2012 CLC by the proposed method and the baseline methods [25] and [27], later defined as

TABLE IV: Land-cover map update result of the 2012 CLC map. The Overall Accuracy (OA%), User Accuracy (UA%), Producer Accuracy (PA%) and Fscore (F1%) are reported for: (1) the supervised Random Forest tolerant noise classifier (Baseline) [25]; (2) the supervised SVM classifier based on the outlier filtering approach (Baseline) [27]; (3) the proposed unsupervised land-cover map update approach.

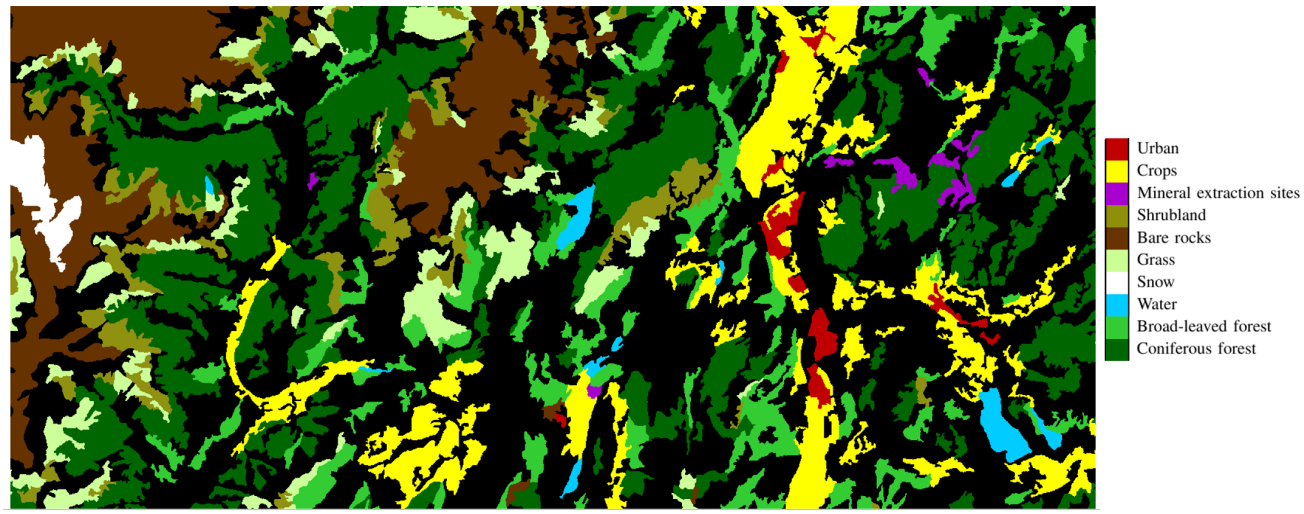
Classes	**Baseline Methods**						**Proposed Method**		
	Random Forest (RM1) [25]			Outlier Filtering (RM2) [27]					
	PA%	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**
Urban	35.96	90.93	51.54	89.44	58.28	70.57	91.11	77.53	83.78
Crops	64.38	59.98	62.10	80.33	89.99	84.88	95.68	90.75	93.15
Broadleaved	78.21	83.65	80.84	98.44	90.34	94.22	99.33	89.87	94.36
Conifers	85.58	91.27	88.34	93.20	94.89	94.03	93.60	94.78	94.19
Grass	75.99	97.49	85.41	98.48	88.04	92.97	99.63	92.78	96.08
Water	99.17	99.86	99.51	100	99.83	99.91	99.93	99.90	99.91
Rocks	75.50	91.05	82.55	93.13	78.23	85.03	98.70	94.96	96.80
Snow	100	86.15	92.56	82.28	100	90.28	100	99.09	99.54
Shrubland	84.87	42.14	56.32	61.30	93.53	74.06	70.47	86.06	77.49
Mineral Site	100	0.02	0.05	72.10	97.64	82.95	75.96	99.14	86.02
OA %	75.08			88.61			93.16		

Reference Method 1 (RM1) and Reference Method 2 (RM2), respectively. The proposed approach obtained classification results in terms of Fscore (F1%) equal or higher than the baselines methods for all the classes. In particular, the F1% achieved by the proposed method ranges from a 77.49% (for the "Shrubland" class) to 99.91% (for the "Water" class), whereas the RM1 ranges from 0.05% (for the "Mineral Site" class) to 99.51% (for the "Water" class) and the RM2 ranges from 70.57% (for the "Urban" class) to 99.91% (for the "Water" class). Note that the proposed method is able to reach high accuracy regardless of the land-cover class, whereas this is not true for the baseline techniques. For instance the RM1 obtained an F1% of 0.05% and 56.32% on "Mineral Site" and "Shrubland", respectively, compared to 86.02% and 77.49% of the proposed method. This is due to the fact that Random Forest alone is not able to capture the models of the classes having a low number of training samples [23]. By removing the outliers using a spectral filtering technique, it is possible to improve the classification accuracy, but the RM2 fails in modelling complex land-cover classes. This is clearly visible in the "Urban" class that reaches a User Accuracy (UA%) of 58.28 compared to the 77.53% of the proposed method. The spectral filtering approach discards all the samples far from the core of the land-cover class distribution. However, the "Urban" includes red and white building roofs, characterized by different spectral behaviour. As most of the samples belongs to the red roof category, the outlier spectral filtering discards all the white ones (identified as outliers). The two-step procedure performed at polygon level allows us to remove most of the noisy samples, while keeping the informative ones. The effectiveness of the proposed approach is also confirmed by the OA of 93.16% compared to the 75.08% and 88.61% achieved by the RM1 and the RM2, respectively.

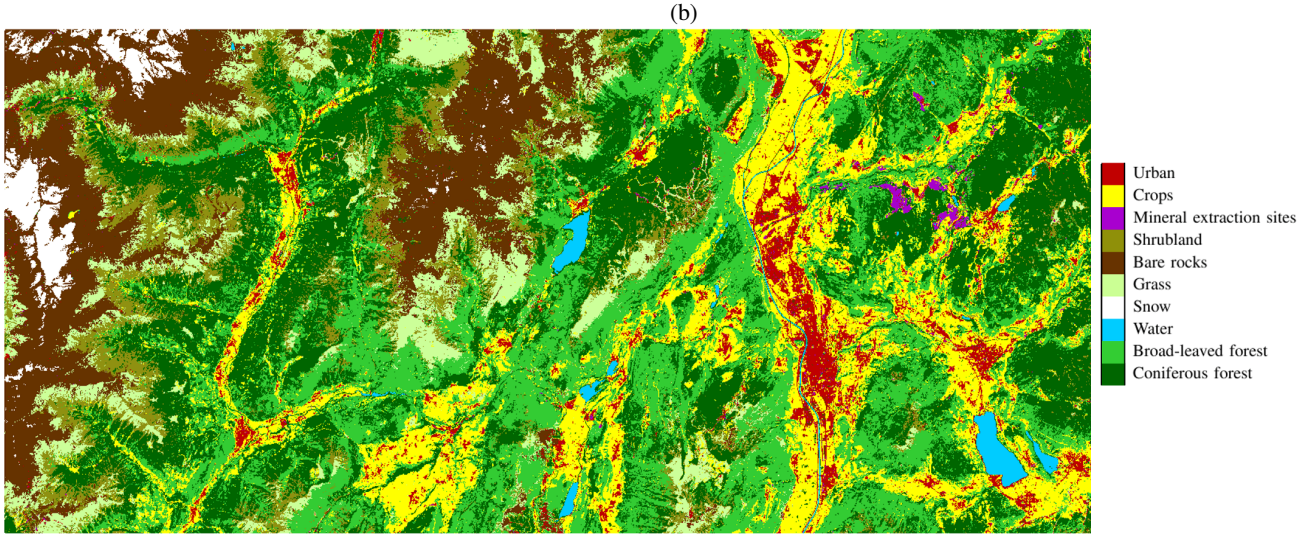
Fig. 6 shows the original 2012 CLC map, the map after the legend conversion and the updated map for the whole study area. To perform a qualitative analysis Fig. 7 shows different portions of the updated 2012 CLC map, by presenting the original map (Fig. 7a, e, i, m, q), the converted map (Fig. 7b, f, j, n, r), one of the Sentinel 2 image used for the updating task (Fig. 7c, g, k, o, s) and the updated map (Fig. 7d, h, l, p, t). Due to the minimum mapping polygon unit of 25 ha, the CLC map provides a high level description of the land-cover by removing important geographical details such as the presence of rivers or narrow roads (Fig. 7q). In contrast, the high spatial resolution of Sentinel 2 images, allows us to produce an updated map that represents the ground cover with a high geometric detail (see Fig. 7t). The ensemble of SVMs accurately classifies pixels belonging to the river and correctly delineates the geometrical structure of all the artificial surfaces (both in terms of buildings and roads). Similar results are depicted in Fig. 7p, which shows the capability of the proposed method to accurately classify land-cover classes similar from the spectral view point such as "Mineral Site" and "Rock" due to the extraction of reliable and informative samples from the map. Moreover, the proposed method allows the semantic decomposition of polygons associated to classes such as "Mixed Forest", by classifying the polygon pixels as "Broadleaved" or "Conifers" (see Fig. 7d 7h 7l). Tab. V shows the quantitative evaluation of the update performed on GlobLand30 by the proposed method and the baseline methods RM1 and RM2. Similarly to the CLC map case, the proposed method achieves accurate classification results. The F1% achieved by the proposed method ranges between a minimum of 79.34% (for the "Shrubland" class) and a maximum of 100% (for the "Snow" class), whereas the RM1 ranges from a minimum of 36.08% (for the "Shrubland" class) to a maximum of 99.55% (for the "Water" class) and the RM2 ranges from a minimum of 58.73% (for the "Shrubland" class) to a maximum of 99.77% (for the "Water" class). Also in this case, the proposed approach outperforms the baselines on the minority classes.



(a)



(b)



(c)

Fig. 6: 2012 CLC Map: (a) initial map, (b) map converted according to the LCCS standard, and (c) map updated with the proposed method.

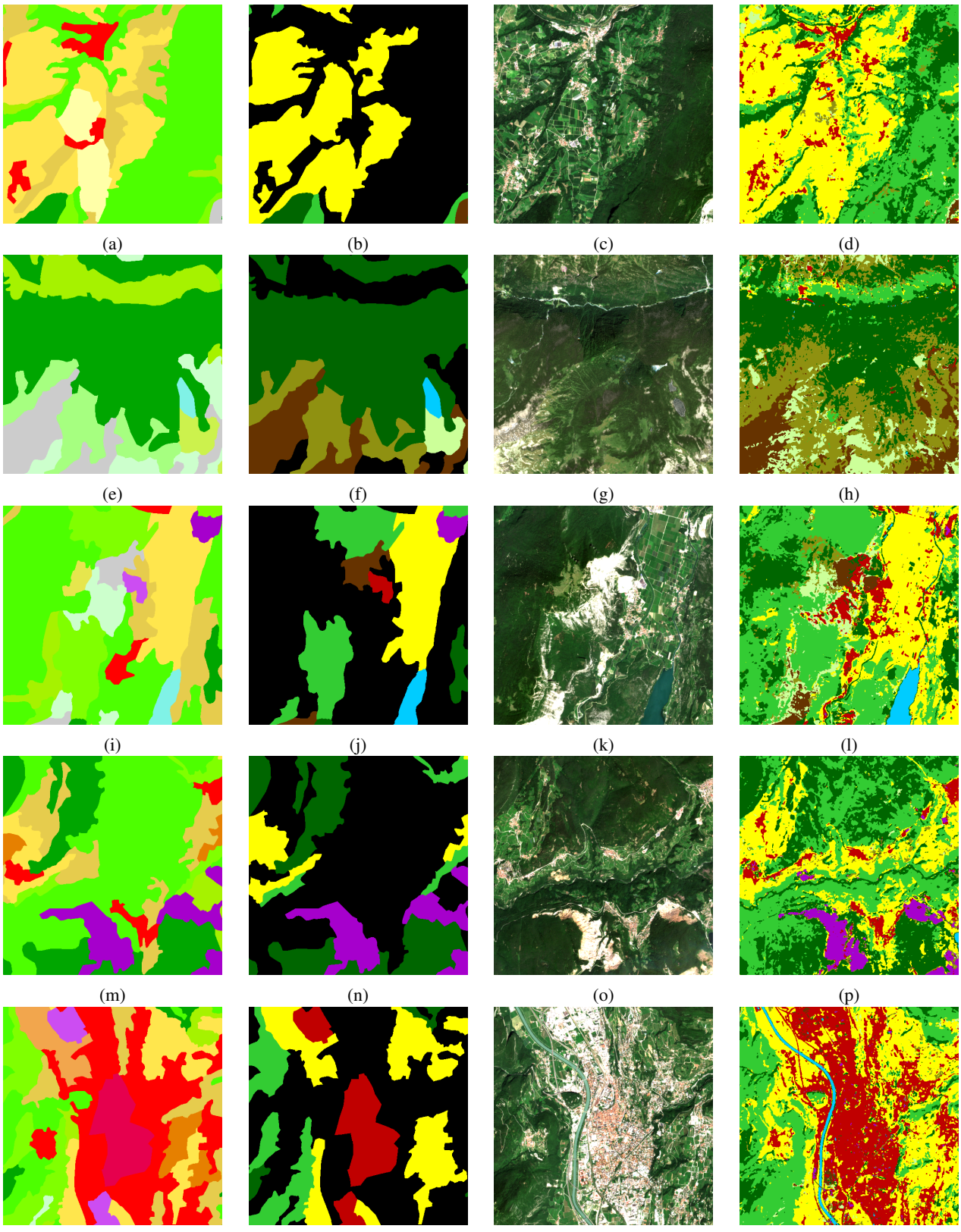


Fig. 7: 2012 CLC map updating results on small portions of the considered test area: (a),(e),(i),(m),(q) original CLC map; (b),(f),(j),(n),(r) converted CLC map; (c),(g),(k),(o),(s) one of Sentinel 2 images used to perform the update; and (d),(h),(l),(p),(t) updated map.

For instance, RM1 and RM2 obtained on the "Shrubland" class a F1% of 36.08% and 58.73%, respectively, while the proposed approach achieved 79.34%. Similarly, on the "Crop" class the proposed method obtained a F1% of 92.58% compared to 59.98% and 78.08% of the RM1 and the RM2, respectively. Moreover, the proposed method obtained comparable classification results on the update of both the maps, by achieving an OA% of 93.29 and of 93.16% starting from GlobLand30 and 2012 CLC maps, respectively. Note that results are provided at pixel level and no post-processing has been applied to the updated maps.

Fig. 8 shows the original GlobLand30 map, the map after the legend conversion and the updated map for the whole study area. Although the 2012 CLC map has been produced with a different geometric resolution and 2 years later than GlobLand30, there is a good agreement on the land-cover classes present in the scene (see Fig. 6a and Fig. 8a). Fig. 9 shows some qualitative examples of the updated GlobLand30 map, by presenting the original GlobLand30 map (Fig. 9a, e, i, m, q), the converted map (Fig. 9b, f, j, n, r), one Sentinel 2 image of the TS (Fig. 9c, g, k, o, s) and the updated map (Fig. 9d, h, l, p, t) for different portions of the considered study area. Differently from the CLC map, this map is provided at a higher spatial resolution (30m), even though from the semantic view point no distinction are made for "Broadleaved" and "Conifer" (both included in the "Forest" class), while "Mineral Site" is included into the "Urban" semantic class. However, the proposed approach further increases the geometric detail to spatial resolution of 10 m while accurately updating the land-cover map. This is also due to the fact that no aggregation is performed on the updated map differently form the considered thematic product. For instance, Fig. 9t shows the higher level of detail provided by the updated map and the accurate detection of the building present in the scene.

Tab. VI and VII report the computation times taken by an Intel Core i7-7700 CPU running at 3.60 GHz, with 32 GB of RAM using an implementation in MATLAB to update the land-cover map on the considered study area (2680×5780 pixels representing an area of 1549 km^2) starting from the 2012 CLC and the GlobLand30 maps, respectively. The model selection and the SVM classification have been parallelized over 5 cores for both the proposed method and the RM2, while the RM1 was not parallelized due to its low computational complexity. As expected the proposed method takes a higher computational time with respect to the reference methods due to the ensemble of 5 SVMs and the two step procedure performed at polygon level. However, this increased time results in a significant improvement of both the classification accuracy and the reliability of the obtained land-cover map. Note that the implementation of the proposed method can be optimized by parallelizing the polygon analysis to extract the reliable samples. Moreover, the classification time required from the ensemble of SVMs can be decreased by increasing the number of parallel tasks.

V. Conclusion and Discussion

In this paper, we presented a novel approach to the automatic update of existing land-cover maps by classification of recent TS of satellite MS images. The method is completely unsupervised and does not assume to know the source of the thematic product, which can be different from a RS data. The main assumptions of the method are that: (1) the source map is aggregated at polygon level according to a majority rule approach, and (2) the legend can be converted into an exhaustive set of classes discriminable using the spectral information provided by the MS data used to perform the update. Under these assumptions, it is possible to automatically extract a reliable training set which can be used to update the thematic product without any labor intensive manual analysis.

From the results obtained one can observe that the legend conversion allows us to deal with a classification scheme that represents the more widespread land-cover categories that can be discriminated by using the spectral properties of the MS data. The proposed system architecture allows us to extract from the land-cover map a reliable pool of labeled samples due to the two-step procedure performed at polygon level, i.e., polygon k-means clustering and polygon consistency analysis. The update is performed using a TS made up of four Sentinel 2 images acquired between May and September, which allows us to discriminate the set of land-cover classes present in the scene. By sampling the temporal signature of the considered classes in May, June, July and September, we increase the classification accuracy that can be obtained by using a single date acquisition. For instance, in May pixels belonging to "Crop" may present a similar spectral behavior to those belonging to "Grass" due to analogous vegetation state. However, the use of the September acquisition may solve the ambiguity as "Crop" pixels after harvesting have a completely different spectral signature from "Grass". Note that denser TS should be used to deal with a set of classes characterized by high temporal variability during the year (i.e., rapid rate of change), while longer TS of images can be employed to consider more complex classification schemes (i.e., distinguish different cultivations). In the first step the data-driven clustering approach discards the samples belonging to the minority classes present in the polygon (i.e., classes that are not associated to the polygon label), while in the second step only the polygons having the highest probability to be correctly labeled are considered. This polygon analysis allows us to remove unreliable samples without discarding informative samples useful to represent the land-cover class distribution.

The pool of reliable units detected, by the proposed approach is used to train a multiple classifiers system that results in more accurate and robust maps with respect to the presence of noisy samples in the *pseudo training sets* generated. The quantitative and qualitative analysis performed on the obtained thematic products confirm the effectiveness of the proposed approach, thus leading to reliable updated thematic products. Moreover, due to the 10 m spatial resolution of the Sentinel 2 data used to perform the update, the method sharply increases to 10 m the geometric details of the outdated maps (i.e., 100m for 2012 CLC and 30m for the GlobLand30). For the 2012 CLC map, this spatial improvement leads to a more accurate description of the thematic products from the semantic view point by replacing polygons having mixed

TABLE V: Land-cover map update result of the GlobLand30 map. The Overall Accuracy (OA%), User Accuracy (UA%), Producer Accuracy (PA%) and Fscore (F1%) are reported for: (1) the supervised Random Forest tolerant noise classifier (Baseline) [25]; (2) the supervised SVM classifier based on the outlier filtering approach (Baseline) [27]; (3) the proposed unsupervised land-cover map update approach.

Classes	**Baseline Methods**						**Proposed Method**		
	Random Forest (RM1) [25]			Outlier Filtering (RM2) [27]					
	PA%	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**
Urban	96.50	59.86	73.89	82.60	90.52	86.38	79.32	99.68	88.34
Crops	61.08	58.92	59.98	76.07	80.19	78.08	94.22	91.00	92.58
Forest	84.71	92.95	88.64	98.24	94.79	96.49	99.32	92.50	95.79
Grass	83.33	96.78	89.55	98.54	97.10	97.82	98.95	93.93	96.37
Water	99.10	100	99.55	100	99.55	99.77	100	99.86	99.93
Rocks	59.15	98.07	73.79	99.80	81.34	89.63	91.56	83.12	87.14
Snow	100	86.72	92.89	99.93	99.59	99.76	100	100	100
Shrubland	86.15	22.82	36.08	47.25	77.59	58.73	74.43	84.95	79.34
OA %	81.78			91.66			93.29		

TABLE VI: Run times obtained on the considered study area (2680 $\times$ 5780 pixels representing an area of 1549 km^2) for updating the 2012 CLC map.

	Time		
	Random Forest (RM1) [25]	Outlier Filtering (RM2) [27]	Proposed Method
Extract Map Labels	2.20 s	6.53 s	823.25 s
Feature Selection	-	47.81 s	47.81 s
Train the Classifier	14.83 s	259.38 s	259.38 s
Classify the Study Area	2543.50 s	4968.86 s	24844.30 s

TABLE VII: Run times obtained on the considered study area (2680 $\times$ 5780 pixels representing an area of 1549 km^2) for updating the GlobLand30 map.

	Time		
	Random Forest (RM1) [25]	Outlier Filtering (RM2) [27]	Proposed Method
Extract Map Labels	2.14 s	6.30 s	598.48 s
Feature Selection	-	41.42 s	41.42 s
Classifier Train	16.46 s	126.56 s	126.56 s
Classify the Study Area	2053.50 s	2262.22 s	11311.10 s

labels such as "Mixed Forest" with more pure pixels associated to "Broadleaved" or "Conifer".

We would like to remark that the proposed method is completely scalable since it extracts labeled samples all over the map. Since the experimental analysis was performed at local level, the ensemble of *pseudo training sets* used to train the classification rule was applied to the entire study area. However, when extending the analysis at national or continental level, the possibility of generating several ensemble of local *pseudo training sets* is fundamental to: (i) face the problem of the variability of the spectral signature, (ii) face the need of modeling different climate regions, and (iii) automatically generate locally balanced training sets, proportionate to the original prior probability of the land-cover classes in the considered area.

As future developments, we aim to design a system architecture based on the proposed approach to produce frequently updated maps at large scale (e.g, national level). To this end, we plan to define a pre-processing strategy that automatically generate cloud free composites that can be used to update the existing thematic products at seasonal level. The system would allow a constant land-cover monitoring performed at high spatial resolution in an unsupervised way. In this context, we aim to define a post-processing strategy to generate the map at object level (thus removing noisy pixels).

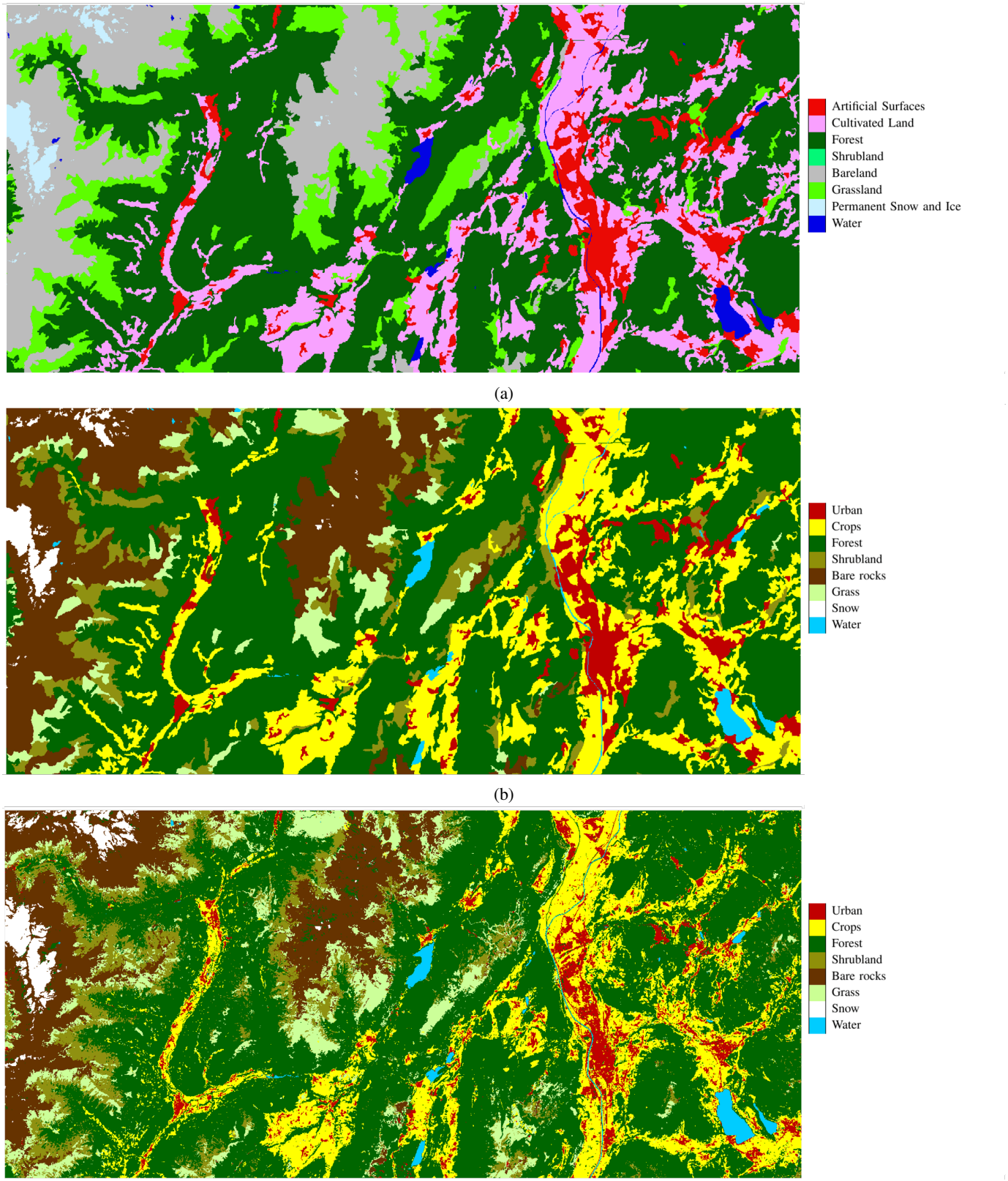


Fig. 8: GlobLand30 Map: (a) initial map, (b) map converted according to the LCCS standard, and (c) map updated with the proposed method.

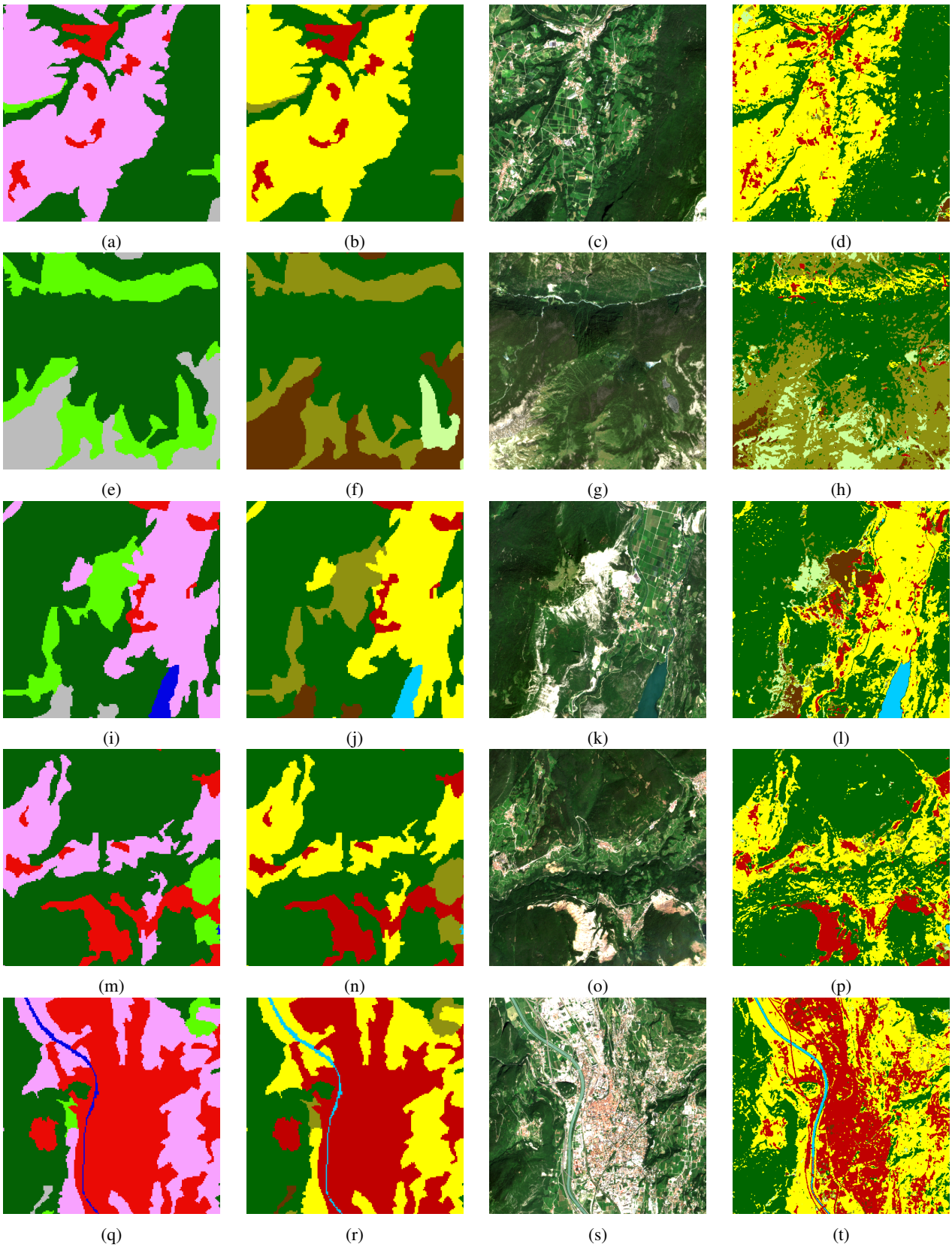


Fig. 9: GlobLand30 land-cover map updating results on small portions of the considered test area: (a),(e),(i),(m),(q) original GlobLand30 map; (b),(f),(j),(n),(r) converted GlobLand30 map; (c),(g),(k),(o),(s) one of Sentinel 2 images used to perform the update; and (d),(h),(l),(p),(t) updated map.

VI. Acknowledgment

This work was supported by the ESA in the framework of the "S2-4Sci Land and Water - Multitemporal Analysis" project.

References

[1] M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort *et al.*, "Sentinel-2: Esa's optical high-resolution mission for gmes operational services," *Remote Sensing of Environment*, vol. 120, pp. 25–36, 2012.

[2] J. Feranec, T. Soukup, G. Hazeu, and G. Jaffrain, *European landscape dynamics: CORINE land cover data*. CRC Press, 2016.

[3] J. Chen, J. Chen, A. Liao, X. Cao, L. Chen, X. Chen, C. He, G. Han, S. Peng, M. Lu, W. Zhang, X. Tong, and J. Mills, "Global land cover mapping at 30m resolution: A pok-based operational approach," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 103, pp. 7 – 27, 2015.

[4] C. Giri and J. Long, "Land cover characterization and mapping of south america for the year 2010 using landsat 30 m satellite data," *Remote Sensing*, vol. 6, no. 10, pp. 9494–9510, 2014.

[5] B. Demir, L. Minello, and L. Bruzzone, "Definition of effective training sets for supervised classification of remote sensing images by a novel cost-sensitive active learning method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 2, pp. 1272–1284, 2014.

[6] L. Bruzzone and B. Demir, "A review of modern approaches to classification of remote sensing data," *Remote Sensing and Digital Image Processing*, vol. 18, pp. 127–143, 2014.

[7] S. Valero, D. Morin, J. Inglada, G. Sepulcre, M. Arias, O. Hagolle, G. Dedieu, S. Bontemps, P. Defourny, and B. Koetz, "Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions," *Remote Sensing*, vol. 8, no. 1, p. 55, 2016.

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

[9] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive svm for semisupervised classification of remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3363 –3373, 2006.

[10] D. C. Zanotta, L. Bruzzone, F. Bovolo, and Y. E. Shimabukuro, "An adaptive semisupervised approach to the detection of user-defined recurrent changes in image time series," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3707–3719, 2015.

[11] L. Bruzzone and M. Marconcini, "Domain adaptation problems: A dasvm classification technique and a circular validation strategy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 770 –787, 2010.

[12] L. Bruzzone and D. F. Prieto, "A partially unsupervised cascade classifier for the analysis of multitemporal remote-sensing images," *Pattern Recognition Letters*, vol. 23, no. 9, pp. 1063–1071, 2002.

[13] L. Bruzzone and M. Marconcini, "Toward the automatic updating of land-cover maps by a domain-adaptation svm classifier and a circular validation strategy," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 4, pp. 1108–1122, 2009.

[14] L. Bruzzone and D. F. Prieto, "Unsupervised retraining of a maximum likelihood classifier for the analysis of multitemporal remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 2, pp. 456–460, 2001.

[15] R. Cossu, S. Chaudhuri, and L. Bruzzone, "A context-sensitive bayesian technique for the partially supervised classification of multitemporal images," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 3, pp. 352 – 356, 2005.

[16] K. Bahirat, F. Bovolo, L. Bruzzone, and S. Chaudhuri, "A novel domain adaptation bayesian classifier for updating land-cover maps with class differences in source and target domains," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 7, pp. 2810–2826, 2012.

[17] B. Banerjee, F. Bovolo, A. Bhattacharya, L. Bruzzone, S. Chaudhuri, and K. M. Buddhiraju, "A novel graph-matching-based approach for domain adaptation in classification of remote sensing image pair," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 4045–4062, 2015.

[18] G. Camps-Valls, T. Bandos Marsheva, and D. Zhou, "Semi-supervised graph-based hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3044–3054, 2007.

[19] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples," *The Journal of Machine Learning Research*, vol. 7, pp. 2399–2434, 2006.

[20] L. Gómez-Chova, G. Camps-Valls, J. Munoz-Mari, and J. Calpe, "Semisupervised image classification with laplacian support vector machines," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 3, pp. 336–340, 2008.

[21] T. S. Caetano, J. J. McAuley, L. Cheng, Q. V. Le, and A. J. Smola, "Learning graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 6, pp. 1048–1058, 2009.

[22] B. Demir, F. Bovolo, and L. Bruzzone, "Classification of time series of multispectral images with limited training data," *IEEE Transactions on Image Processing*, vol. 22, no. 8, pp. 3219–3233, 2013.

[23] J. Inglada, A. Vincent, M. Arias, B. Tardy, D. Morin, and I. Rodes, "Operational high resolution land cover map production at the country scale using satellite image time series," *Remote Sensing*, vol. 9, no. 1, 2017.

[24] B. Frénay and M. Verleysen, "Classification in the presence of label noise: a survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 845–869, 2014.

[25] C. Pelletier, S. Valero, J. Inglada, N. Champion, C. Marais Sicre, and G. Dedieu, "Effect of training class label noise on classification performances for land cover mapping with satellite image time series," *Remote Sensing*, vol. 9, no. 2, p. 173, 2017.

[26] J. Silva, F. Bacao, G. Foody, and M. Caetano, "Automatic selection of training areas using existing land cover maps," in *ESA Special Publication*, vol. 722, 2013, p. 184.

[27] J. Radoux, C. Lamarche, E. Van Bogaert, S. Bontemps, C. Brockmann, and P. Defourny, "Automated training sample extraction for global land cover mapping," *Remote Sensing*, vol. 6, no. 5, pp. 3965–3987, 2014.

[28] J. Radoux and P. Defourny, "Automated image-to-map discrepancy detection using iterative trimming," *Photogrammetric Engineering & Remote Sensing*, vol. 76, no. 2, pp. 173–181, 2010.

[29] N. Matton, G. S. Canto, F. Waldner, S. Valero, D. Morin, J. Inglada, M. Arias, S. Bontemps, B. Koetz, and P. Defourny, "An automated method for annual cropland mapping along the season for various globally-distributed agrosystems using high spatial and temporal resolution time series," *Remote Sensing*, vol. 7, no. 10, pp. 13 208–13 232, 2015.

[30] D. A. Landgrebe, *"Signal theory methods in multispectral remote sensing"*. John Wiley & Sons, 2005, vol. 29.

[31] M. Hansen, R. DeFries, J. R. Townshend, and R. Sohlberg, "Global land cover classification at 1 km spatial resolution using a classification tree approach," *International Journal of Remote Sensing*, vol. 21, no. 6-7, pp. 1331–1364, 2000.

[32] S. Faroux, A. K. Tchuenté, J.-L. Roujean, V. Masson, E. Martin, and P. Le Moigne, "Ecoclimap-ii/europe: A twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models," *Geoscientific Model Development*, vol. 6, no. 2, p. 563, 2013.

[33] R. G. Congalton, J. Gu, K. Yadav, P. Thenkabail, and M. Ozdogan, "Global land cover mapping: A review and uncertainty analysis," *Remote Sensing*, vol. 6, no. 12, pp. 12 070–12 093, 2014.

[34] C. Gómez, J. C. White, and M. A. Wulder, "Optical remotely sensed time series data for land cover classification: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, pp. 55–72, 2016.

[35] C. Paris, L. Bruzzone, and D. Fernandez-Prieto, "A novel automatic approach to the update of land-cover maps by unsupervised classification of remote sensing images," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 2207–2210.

[36] S. E. Franklin, O. S. Ahmed, M. A. Wulder, J. C. White, T. Hermosilla, and N. C. Coops, "Large area mapping of annual land cover dynamics using multitemporal change detection and classification of landsat time series data," *Canadian Journal of Remote Sensing*, vol. 41, no. 4, pp. 293–314, 2015.

[37] A. Masjedi, M. J. V. Zoej, and Y. Maghsoudi, "Classification of polarimetric sar images based on modeling contextual information and using texture features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 932–943, 2016.

[38] A. Buono, F. Nunziata, M. Migliaccio, X. Yang, and X. Li, "Classification of the yellow river delta area using fully polarimetric sar measurements," *International Journal of Remote Sensing*, vol. 38, no. 23, pp. 6714–6734, 2017.

[39] X. Niu and Y. Ban, "Multi-temporal radarsat-2 polarimetric sar data for urban land-cover classification using an object-based support vector

machine and a rule-based approach," *International Journal of Remote Sensing*, vol. 34, no. 1, pp. 1–26, 2013.
[40] Z. Qi, A. G.-O. Yeh, X. Li, and Z. Lin, "A novel algorithm for land use and land cover classification using radarsat-2 polarimetric sar data," *Remote Sensing of Environment*, vol. 118, pp. 21–39, 2012.
[41] Y. Li, J. Chen, Q. Ma, H. K. Zhang, and J. Liu, "Evaluation of sentinel-2a surface reflectance derived using sen2cor in north america," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 6, pp. 1997–2021, 2018.
[42] J. Chen, X. Zhu, J. E. Vogelmann, F. Gao, and S. Jin, "A simple and effective method for filling gaps in landsat etm+ slc-off images," *Remote Sensing of Environment*, vol. 115, no. 4, pp. 1053–1064, 2011.
[43] M. Herold, C. E. Woodcock, A. Di Gregorio, P. Mayaux, A. S. Belward, J. Latham, and C. C. Schmullius, "A joint initiative for harmonization and validation of land cover datasets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, pp. 1719–1727, 2006.
[44] M. Herold, P. Mayaux, C. Woodcock, A. Baccini, and C. Schmullius, "Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets," *Remote Sensing of Environment*, vol. 112, no. 5, pp. 2538–2556, 2008.
[45] T. R. Loveland, B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Yang, and J. W. Merchant, "Development of a global land cover characteristics database and igbp discover from 1 km avhrr data," *International Journal of Remote Sensing*, vol. 21, no. 6-7, pp. 1303–1330, 2000.
[46] A. T. K. Tchuenté, J. L. Roujean, and S. M. De Jong, "Comparison and relative quality assessment of the glc2000, globcover, modis and ecoclimap land cover data sets at the african continental scale," *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, no. 2, pp. 207–219, 2011.
[47] C. Giri, Z. Zhu, and B. Reed, "A comparative analysis of the global land cover 2000 and modis land cover data sets," *Remote Sensing of Environment*, vol. 94, no. 1, pp. 123–132, 2005.
[48] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using k-means clustering algorithm and subtractive clustering algorithm," *Procedia Computer Science*, vol. 54, pp. 764–771, 2015.
[49] A. K. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651 – 666, 2010.
[50] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1650–1654, 2002.
[51] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1–27, 1974.
[52] Y. Liu, Z. Li, H. Xiong, X. Gao, J. Wu, and S. Wu, "Understanding and enhancement of internal clustering validation measures," *IEEE Transactions on Cybernetics*, vol. 43, no. 3, pp. 982–994, 2013.
[53] C. Persello and L. Bruzzone, "Active learning for domain adaptation in the supervised classification of remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4468–4483, 2012.
[54] S. V. Stehman and R. L. Czaplewski, "Design and analysis for thematic map accuracy assessment: fundamental principles," *Remote Sensing of Environment*, vol. 64, no. 3, pp. 331–344, 1998.
[55] S. B. Serpico and L. Bruzzone, "A new search algorithm for feature selection in hyperspectral remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 7, pp. 1360–1367, 2001.
[56] L. Bruzzone, F. Roli, and S. B. Serpico, "An extension of the jeffreys-matusita distance to multiclass cases for feature selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 6, pp. 1318–1321, 1995.
[57] A. Harikumar, F. Bovolo, and L. Bruzzone, "An internal crown geometric model for conifer species classification with high-density lidar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2924–2940, 2017.
[58] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1351–1362, 2005.
[59] L. Breiman, "Bagging predictors," *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.
[60] J. C.-W. Chan, C. Huang, and R. DeFries, "Enhanced algorithm performance for land cover classification from remotely sensed data using bagging and boosting," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 693–695, 2001.
[61] B. Kosztra, G. Büttner, G. Hazeu, and S. Arnold, "Updated clc illustrated nomenclature guidelines," *Nomenclature Guidelines. Final Report by European Environmental Agency.*, 2017.

APPENDIX

The mathematical notation used in this paper is listed in the following table. Symbols are reported in the order of appearance.

Symbol	**Description**
B	number of pixels of the MS images
d	spectral channels of the MS images
$\mathbf{X}_q$	qth MS image acquired at time t_q, where $\mathbf{X}_q \in \mathbb{R}^{B\times d}$
Q	number of MS images of the TS
n	number of spectral channels of the TS of images, where $n = d \cdot Q$
$\mathbf{X}$	TS of MS images, where $\mathbf{X} \in \mathbb{R}^{B\times n}$
$\mathbf{x}_b$	multitemporal spectral vector of the TS, where $\mathbf{x}_b \in \mathbb{R}^{1\times n}$
$\mathcal{M}$	Source Thematic Product
ω_u	uth land-cover class
U	number of land-cover classes of the source map legend
Ω	Set of land-cover classes of the source thematic product, where $\Omega = \{\omega_u\}_{u=1}^{U}$
λ_j	jth polygon of the source thematic map
$\mathcal{P}$	Set of polygons of the source thematic product, where $\mathcal{P} = \{\lambda_j\}_j$
U_c	number of land-cover classes converted from the source map legend
Ω_c	Set of converted land-cover classes, where $\Omega_c = \{\omega_u\}_{u=1}^{U_c}$
$\mathbf{l}$	labels associated to the multitemporal vectors of the TS, where $\mathbf{l} \in \mathbb{R}^{B\times 1}$
$\mathbf{p}$	polygons associated to the multitemporal vectors of the TS, where $\mathbf{l} \in \mathbb{R}^{B\times 1}$
B_j	number of multitemporal vectors associated to the jth polygon
$\mathbf{P}_j = \mathbf{X}_{(p_i=\lambda_j,:)}$	set of multitemporal vectors associated to the jth polygon, with $\mathbf{P}_j \in \mathbb{R}^{B_j\times n}$
K_j	number of clusters automatically detected in the polygon $\mathbf{P}_j$
$\{\mathbf{C}_j^1, \mathbf{C}_j^2, \ldots, \mathbf{C}_j^{K_j}\}$	set of clusters derived by partitioning $\mathbf{P}_j$
N_j^k	number of multitemporal vectors associated to the kth cluster $\mathbf{C}_j^k$
$\mathbf{m}_k$	centroid of $\mathbf{C}_j^k$
$\mathbf{m}$	centroid of $\mathbf{P}_j$
N_u	number of clusters associated to the land-cover class ω_u
$\boldsymbol{\mu}_u$	mean value of ω_u
$\boldsymbol{\Sigma}_{\boldsymbol{u}}$	covariance matrix of ω_u
$\boldsymbol{\mu}_j$	mean value of $\mathbf{C}_j$
$\boldsymbol{\Sigma}_{\boldsymbol{j}}$	covariance matrix of $\mathbf{C}_j$
$\{T_1, T_2, \cdots, T_M\}$	set of M *pseudo training sets*
$\{f_1, f_2, \cdots, f_M\}$	set of decision functions of the ensemble of SVM classifiers