A Novel Approach to the Unsupervised Extraction of Reliable Training Samples from Thematic Products

Claudia Paris, *Member, IEEE*, Lorenzo Bruzzone, *Fellow, IEEE*

***Abstract*—Supervised classification algorithms require a sufficiently large set of representative training samples to generate accurate land-cover maps. Collecting reference data is difficult, expensive and unfeasible at large scale. To solve this problem, this paper introduces a novel approach which aims to extract reliable labeled data from existing thematic products. Although these products represent a potentially useful information source, their use is not straightforward. They are not completely reliable since they may present classification errors. They are typically aggregated at polygon level, where polygons do not necessarily correspond to homogeneous areas. Finally, usually there is a semantic gap between map legends and Remote Sensing (RS) data. In this context, we propose an approach which aims to: (i) perform a domain understanding to detect the discrepancies between the thematic map domain and the RS data domain, (ii) use RS data contemporary to the map to decompose the thematic product from the semantic and spatial view point, and (iii) extract a database of informative and reliable training samples. The database of weak labeled units is used for training an ensemble of classifiers on recent data, which results are then combined in a majority voting rule. Two sets of experimental results obtained on MS images by extracting training samples from a crop type map and the 2018 Corine Land Cover (CLC) map, respectively, confirm the effectiveness of the proposed approach.**

***Index Terms*—Weak learning classification, remote sensing (RS), unsupervised methods, land-cover map update.**

I. Introduction

THE major bottleneck of supervised Remote Sensing (RS) data classification is the availability of an adequately large set of representative training samples (i.e., reference data). At operational level this is a crucial issue, since it is impossible to obtain a large amount of either ground reference data or annotated data by photo-interpretation. Besides the amount of training samples, it is also necessary to have a set of informative labeled units being able to represent the behavior of the classes in different portions of a scene. This is particularly evident when classifying multispectral (MS) or hyperspectral optical images, because of the spatial variability of the spectral signatures of the land-cover classes [1]. Different ground conditions strongly affect the spectral response of the same land-cover class, which should be properly characterized to guarantee accurate classification results (i.e., training samples collected all over the scene). Moreover, if the number of labeled units is relatively small compared to the number of features, the system architecture may fail in estimating accurately the classifier parameters and lead to classifier with poor generalization capabilities [2]–[4].

To tackle these problems, in the last years many semi-supervised approaches have been proposed [2]–[7]. These methods aim to enlarge the set of labeled data by using the unlabeled data to better model the distributions of the classes, thus increasing the classification accuracy. Typically, iterative procedures gradually include unlabeled units in the training set to progressively adjust the classification function [3], [5], or graph-based methods are used to connect labeled and unlabeled units according to their similarity [8]–[13]. When the graph is established, unlabeled units can be naturally associated with their land-cover classes under the assumption of consistency (i.e., nearby points should belong to the same class) [14]. Although these strategies can be effective in enlarging small training datasets, often results of semisupervised methods are affected by the initial model assumptions, i.e., inaccurate matching of pattern structure may lead to a degradation of classifier performances. Thus, the possible use of semisupervised techniques requires the choice of strategy robust to initial conditions.

To ensure a reliable transfer of labeled units, several works exploit the multitemporal correlation of Time Series (TS) of RS images. When ground truth is available for at least one image of the TS, it is possible to transfer the labeled units to more recent images in a reliable way [15]–[17]. In [17], Yang *et al* present a domain adaptation framework for multitemporal hyperspectral data. By assuming that local geometries between multitemporal data are similar, two manifold alignment strategies are defined for classifying the hyperspectral images in a common manifold space. In [15], Demir *et al* first detect unchanged areas between the image to be classified and the one where training samples are available. Then, the labels of the unchanged reference areas are used to classify the more recent image. Although these approaches are effective at local level, at country or continental scale most of these methods do not guarantee robust solutions to generate training sets representative of the whole study area. Due to the high spatial variability of the spectral signatures of classes, different portions of the scene present different spectral behavior for the same land-cover classes because of physical factors (e.g., soil moisture, vegetation), and atmospheric conditions [18]. Thus, by extracting samples from small local areas, there is no sufficient information for modelling this variability. Moreover, samples taken from the same region usually have high correlation, thus violating the required assumption of independence [18].

The need of large sets of training samples is even more evident at operational level, when the goal is to generate/update land-cover maps at country, continental or global level. In the last decades, a lot of effort has been devoted to develop thematic/cartographic products due to their valuable contribution to a wide range of applications (e.g., climate change models, monitoring of natural resources, spatial distribution of ecosystems and landscapes, etc). At global level, various thematic products are available [19]–[22]. However, they present many discrepancies when harmonized and compared [23], [24]. This is mainly due to the fact that these land-cover maps were generate by using different data sources, classification schemes, and methodologies. At European level, the Corine Land Cover (CLC) map [25] is one of the most accurate cartography [26], with its detailed classification scheme composed of 44 classes (mixed land-cover and land-use classes). Nevertheless, the minimum mapping unit of 25 Ha does not allow the direct extraction of training samples from the map. At such coarse scale, many pixels aggregated within the same polygon are not correctly associated to their labels. Including them in the training set leads to poor classification accuracies [27].

To generate reliable thematic products, some methods propose to fuse different maps [26], [28], [29]. In [28], Lesiv *et al* generate a hybrid forest map by fusing several well-known cartographic products (e.g., GLC2000, GlobCover 2005, etc) with crowd-sourced data on forest cover collected through the Geo-Wiki project [30]. A Crowd-sourced thematic product is also used in [31], where the authors extract training samples from OpenStreetMap to classify a TS of MS images. A noise tolerant classifier is used to handle the mislabeled units present in the extracted training set due to the inaccurate matching between the polygon boundaries and the real land-cover class. In [26], Pérez-Hoyos *et al* generate a hybrid land-cover map at European level by combining the GLC2000, the MODIS GLC, the GlobCover and the CLC Map. All the maps are re-projected and co-registered into the GLC2000 grid (1km spatial resolution) and the legends of the existing products are linked using semantic rules based on affinity scores. Although mixing different products can be effective, the result strongly depends on the diversity and the initial accuracy of the fused thematic maps. While diversity ensures that the dataset make uncorrelated errors, the initial accuracy is necessary to avoid poor classifications when combining the maps.

Similar results are obtained in [32], [33], where different cartographic products are merged to extract large databases of training samples in an unsupervised way. To deal with the considerable amount of mislabeled units present in the resulting training set, the authors exploit a tolerant to noise classifier [34]. Although the selected classifier can tolerate more than 15% of mislabelled units in the training step, due to the difficult heterogeneous landscape the obtained land-cover map contains numerous classification errors. In [32], better classification results are obtained since the authors merge databases provided at national level (more accurate and updated) and ground data collected during fieldwork campaigns. In particular, the French National Land Cover database produced by the French mapping agency at 1 m spatial resolution is used together with the French Land Parcel Information System database (which maps annually the French crop fields). However, from an operational view point it is not feasible to assume such updated and high resolution cartographic products available at large scale.

Few works introduced approaches to reduce the class noise (i.e., pixels with wrong class assignments) present in the extracted training set [27], [35]. Since thematic products are usually provided at polygon level, within the same polygon not all the pixels belong to the polygon label. To increase the probability of selecting pixels correctly associated to their labels, typically pixels on the polygon boundary are discarded via a simple erosion performed along the edges of the polygon [27], [35]. Moreover, a spectral analysis of the labeled units extracted from the map associated to the same class can be performed to remove the outliers from the distribution (i.e., pixels associated to wrong labels) [36]. Although these outliers removal strategies increase the probability of selecting reliable units from the map, their main drawback is the risk of removing diverse but informative training samples [34], thus strongly affecting the generalization capability of the classifier. In [37] Lin *et al* propose a transfer learning approach to frequently update land-cover maps of rapidly urbanizing regions. First, a rule-based approach based on prior knowledge is used to extract labeled units from the 2010 GlobeLand30 map available at global level. Then, a relational knowledge transfer technique is applied to transfer the labels to a recent RS image and update the map.

Besides their large uncertainty, leverage on existing thematic products seems to be a promising way to generate large databases of labeled units. Thematic/cartographic products represent an extremely interesting source of information to generate reference data at large scale. However, their use is not straightforward. As emerged from the literature overview, these products are not completely reliable since they may present misclassified units. They are typically aggregated at polygon level, where the polygon label represents the predominant class, i.e., most of the units belonging to the polygon are correctly associated to the polygon label but not all of them. Moreover, the polygon boundaries do not perfectly match the grid of pixels of the RS data, thus leading to spurious pixels associated to a single label. Besides the spatial component, it is also necessary to accurately manage the semantic gap between the map legend and the RS data. Most of these products have been generated by multiple sources (e.g., photo-interpretation, ancillary data, crowd sourcing assessment), thus leading to a map legend which does not necessarily correspond to classes discriminable using RS data. In addition, frequently map legends present semantic classes which aggregate natural classes discriminable through the information provided by the RS data, i.e., the land-cover classes. In this context, it is necessary to accurately model the discrepancy between the map domain and the RS domain to extract reliable information from existing thematic products.

This paper presents a novel approach for the extraction of labeled units from existing thematic maps. The approach is based on four main components: (i) source domain understanding, (ii) source domain decomposition, (iii) design the training database and, (iv) land-cover map production. The properties of the thematic product are analyzed to point

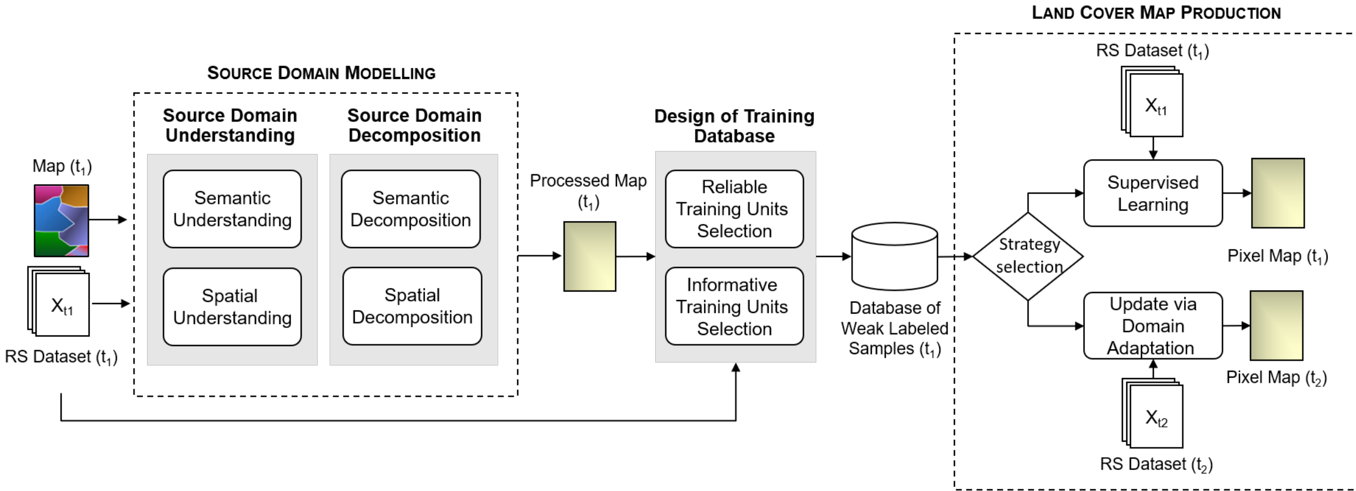


Fig. 1: Work flow of the proposed approach for the automatic extraction of reliable training samples from existing thematic products for the classification of recent RS data.

out its main discrepancy with respect to the RS data. In particular, we analyze the relationship between the spatial properties of the RS data and the map (i.e., map projection, spatial resolution and minimum mapping unit), as well as the semantic gap between the map legend and the set of classes discriminable with the RS data. Then, the approach performs a spatial and a semantic decomposition of the map to facilitate the detection of pure spectral pixels correctly associated to their labels. The training database is designed by selecting informative and reliable labeled units. Finally, the obtained database of weak labeled units is used to produce a high resolution land-cover product provided at pixel level. Due to the complex ill-posed problem faced, the method is based on the following assumptions: (1) RS data contemporary to the map are available, (2) the vector map has been converted into raster and accurately co-registered to the RS data, and (3) the map legend has been converted into an exhaustive set of classes discriminable with the considered RS data.

The rest of the manuscript is organized into nine sections. Section II gives an overview of the proposed approach. Section III describes the source domain understanding component providing a taxonomy of the semantic and spatial properties of the existing thematic products. Section IV focuses on the source domain decomposition component, while Section V explains the design of the training database. In Section VI the production of the land-cover map is presented. Section VII reports the employed dataset in terms of thematic products and RS data images employed, while Section VIII discusses the experimental results obtained. Finally, Section IX draws the conclusion of the paper and presents possible future developments.

II. Proposed Approach to the Extraction of Reliable Training Samples from Existing Thematic Products

Fig. 1 shows the work flow of the proposed approach for the design of systems which extract reliable labeled units from existing cartographic products. Once the discrepancies between the RS data and the thematic product are understood, the elements of the system architecture can be implemented with data analysis techniques that handle the inconsistencies between the selected thematic map and the RS data. The proposed approach is based on the following four components:

1) *Understand the source domain properties*. The thematic map is analyzed from the spatial and semantic view point to detect its discrepancy with respect to the considered RS data. This requires an a priori understanding of the set of land-cover classes that can be recognized using the spectral information provided by the MS data.
2) *Decompose the source domain*. The systems is designed to generate a map decomposed from the semantic and spatial view point, which guarantees the extraction of training samples having the highest probability of being correctly associated to their labels.
3) *Design the training database*. This is the phase in which the pixels having the highest probability of being reliable and informative are extracted from the decomposed map. The database is designed in order to model the prior probabilities of the land-cover classes present in the scene.
4) *Land-cover map production*. The database of reliable labeled units is used to generate a pixel-level classification map. A supervised learning approach is applied to high spatial resolution RS data contemporary to the map to obtain a new updated map characterized by better geometric details than the initial one. If RS data more recent than the map are used, a standard domain adaptation technique should be employed to produce the high spatial resolution updated map.

The proposed approach is conceived for MS optical images since these data are typically used to generate and update land-cover maps with many classes. However, it is flexible and its general concept can be applied to any RS data (e.g., polarimetric synthetic aperture radar data [38], [39]) under the assumption that the considered data allow the discrimination of the set of classes present in the map legend. It is worth noting that once the setup and the design of the architecture are over, the system automatically extracts the training samples from the thematic product in an unsupervised way without any labor intensive manual analysis. To the best of the author's knowledge, current research on the extraction of training sets from existing maps focuses on the removal of mislabeled units at the end of the extraction procedure. There is no work in the literature addressing the spatial and semantic decomposition of the thematic map to increase the probability of detecting reliable and informative samples during the selection process.

III. Source Domain Understanding

Many land-cover products are now available at regional, national, continental and global level. At local scale, very high spatial resolution RS data are typically used to detect detailed spatial patterns. When moving to large scales, coarse spatial resolution RS images become a primary data source to map the extent and the distribution of the major land-cover classes. In this context, it is necessary to understand the properties of the considered thematic product to extract reliable knowledge from it. Fig. 2 reports a categorization of the spatial and semantic properties of existing thematic products.

A. Semantic Understanding

First, it is necessary to analyze and understand the nomenclature of the thematic map. The main goal of this step is to identify the type of classes present in the legend. Indeed, cartographic products usually present semantic classes that do not correspond to land-cover classes that can be discriminated by using the MS information. At the highest level, we can distinguish among four main types of semantic in thematic products: 1) land-use classes (Ω_{Use}), 2) land-cover classes (Ω_{Cov}), 3) spatially aggregated classes (Ω_{Spa}), and 4) semantically aggregated classes (Ω_{Sem}). Each category is detailed as follows.

Land-Cover Classes (Ω_{Cov})*:* Natural classes which can be discriminated with the spectral information provided by the MS image. These classes represent different physical and biological cover of the Earth' surface, which are thus characterized by different spectral signatures (e.g., "Grass", "Water", etc.).

Land-Use Classes (Ω_{Use})*:* classes that describe the socio-economic purpose of the territory assigned by photo-interpretation but not discriminable using the spectral information provided by the MS data. For instance, at pixel level, the "Industrial Units" class is not characterized by a pure spectral signature but can include different natural classes [23].

Spatially Aggregated Classes (Ω_{Spa})*:* The definition of the thematic product is constrained by the minimum mapping unit, even though the corresponding natural classes are present in the map legend. For instance, even though the land-cover classes "Broad-leaves" and "Conifers" are represented, the "Mixed forest" class has to be assigned to areas where both "Broad-leaves" and "Conifers" are present in the scene with an extension smaller that the minimum unit (e.g., minimum mapping unit of 5 ha).

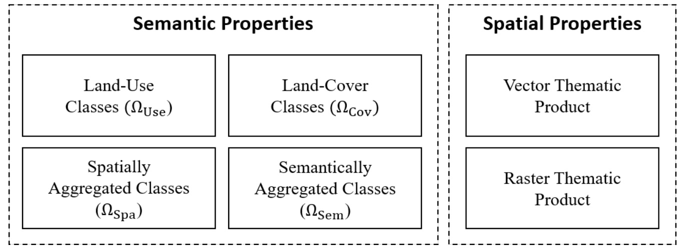


Fig. 2: Taxonomy of the semantic and spatial properties of existing thematic products.

Semantically Aggregated Classes (Ω_{Sem})*:* natural classes that have been semantically aggregated in the map, since their labels are not present in the map legend. This typically occurs in thematic products provided at large scale. The larger is the map scale, the higher is the level of abstraction. A clear example is the agricultural case. At large scale, it is not possible to include in the map legend all the different cultivations present in the scene. While at continental level, typically thematic products present classes such as "Winter crops" or "Summer crops", at continental or global scale they may be categorized simply as "Crops".

B. Spatial Understanding

In the second step of this component, we analyze the spatial properties of the thematic products. From the spatial view points, the cartographic products can be categorized according to the data structure used to encode the spatial information: 1) vector thematic product, and 2) raster thematic product. Vectors have been widely employed for surveying and map-making due to their capability of capturing topological information difficult to achieve with the raster model. However, raster maps are particularly useful to easily perform spatial analysis and comparison [40].

Vector Thematic Products*:* Databases made up of geo-referenced polygons where each element is associated to a thematic attribute. Due to the predefined minimum mapping unit, some polygons may include different land-cover classes even though they are associated to a single label. Typically, the majority rule approach is employed to assign the label to the polygon, i.e., the dominant class is the polygon label. Since the polygon boundaries do not perfectly match the pixel grid of the optical data, when re-sampling the map on the pixel grid of a MS image, several pixels may fall across vector boundaries.

Raster Thematic Products*:* Maps sampled on a geo-referenced grid according to a predefined ground sampling distance (GSD). The need of projecting the land-cover areas on a predefined grid penalizes the naturally fuzzy boundaries between classes as well as the topological details of complex geometric structures. Typically raster products generated at large scale (continental of global) are provided at coarse spatial resolution. Note that, if the MS data used are characterized by a different map projection and spatial resolution, the map has to be re-sampled to match the grid of the MS data. These maps can be provided at polygon or pixel level.

In both cases, there are a one-to-many and a many-to-one relations between the label assigned to the minimum mapping unit (i.e., polygon or pixel) and the ones correctly associated to the pixels of the MS data since: (i) the minimum mapping unit may include different classes, and (ii) re-sampling the thematic product on the MS image pixel grid leads to spurious pixels associated to partially correct labels.

IV. Source Domain Decomposition

Fig. 3 summarizes the source domain properties that should be accurately modeled to extract reliable knowledge from the considered thematic product. The goal of this component is to convert the initial thematic product into a map which is: 1) spatially decomposed, and 2) semantically decomposed into an exhaustive set of land-cover classes. According to the taxonomy presented in Section III-A, Ω may be partitioned into the following categories $\{\Omega_{\text{Cov}}, \Omega_{\text{Use}}, \Omega_{\text{Sem}}, \Omega_{\text{Spa}}\}$. While Ω_{Cov} can be directly inherited, the Ω_{Use} should be converted into land-cover labels according to the Land-Cover Classification System (LCCS), which is the standard common land-cover language for translating and comparing existing legends [41]. For instance, the "Industrial Units" class, which is a land-use definition that can be assigned by photo interpretation, should be converted into "Artificial Surfaces" since at pixel level no pure spectral signature can be unambiguously associated to the "Industrial Units" definition [23], [24]. The spatially aggregated classes Ω_{Spa} can be neglected since the land-cover classes included in Ω_{Spa} are already present in the legend. Thus, the pixels belonging to these classes will be replaced by the corresponding land-cover classes if correctly classified. In contrast, Ω_{Sem} should be decomposed. Thus, first the thematic map is converted in order to have only classes $\Omega_1 = \{\Omega_{\text{Cov}}, \Omega_{\text{Sem}}\}$. Then, the spatial and semantic decomposition is performed.

Let $\mathbf{X}^{t_1}$ be the MS image acquired at time t_1 and $\mathbf{M}^{t_1}_{\Omega_1}$ the contemporary thematic product co-registered and re-sampled at the same spatial resolution of $\mathbf{X}^{t_1}$. The MS image is made up of $N \times M$ pixels and characterized by B spectral channels, i.e., $\mathbf{X}^{t_1} \in \mathbb{R}^{N\times M\times B}$. The considered map $\mathbf{M}^{t_1}_{\Omega_1}$ is characterized by a set of K classes $\Omega_1 = \{\omega_k\}_{k=1}^{K}$ and a set of J polygons $\mathcal{P} = \{\mathbf{P}_j\}_{j=1}^{J}$. The number of polygons is expected to be different from the number of classes since many polygons can be associated to the same label (i.e., $J \gg K$). Therefore, the ith pixel $\mathbf{x}_i \in \mathbf{X}^{t_1}$ is a B-dimensional spectral vector $\mathbf{x}_i \in \mathbb{R}^B$, with $i \in [1, \cdots, N\times M]$, associated to a unique label $\omega_k \in \Omega_1$ and a unique polygon $\mathbf{P}_j \in \mathcal{P}$.

A. Spatial Decomposition

According to the spatial analysis presented in Section III-B, the approach has to deal with: the possible presence of more than one natural class in each polygon (i.e., minimum mapping unit decomposition), and (ii) spectrally spurious pixels associated to unique labels (i.e., pixel decomposition). Note that the map is assumed to be characterized by a coarser spatial resolution with respect to the MS data used. In this context, it is necessary to spatially decompose the map into a pixel map having the same spatial resolution of the considered MS data.

Let $\mathbf{P}_j = (\mathbf{x}^j_1; \mathbf{x}^j_2; \cdots, \mathbf{x}^j_{n_j}) \in \mathbb{R}^{n_j\times B}$ be the jth polygon composed of n_j pixels and characterized by the B spectral channels of $\mathbf{X}^{t_1}$. Let us assume that the polygon label is ω_k. The proposed system aims to exploit the MS information to detect the pixels belonging to $\mathbf{P}_j$ that are correctly associated to ω_k. To this end, the polygons are partitioned into V_j clusters according to their spectral similarity. The number of clusters V_j is automatically detected by using the Calinski Harabasz (CH) Index [42], which is widely employed for determining the optimal number of clusters in a data set. This index is computed as the ratio between the overall within-cluster variance and the overall between-cluster variance, as follows:

$$V_j = \underset{V_j\in[2,L]}{\text{argmax}} \left\{ \frac{[traceB_j/(V_j-1)]}{[traceW_j/(n_j-V_j)]} \right\} \tag{1}$$

where B_j and W_j are the between and within cluster scatter matrices computed for $\mathbf{P}_j$, respectively, and V_j is the optimal clustering value among the L tested. Due to the spectral similarity of the labeled units belonging to the same class, the algorithm automatically detects homogeneous clusters belonging to different land-cover classes. Here, for simplicity we use the standard K-means clustering algorithm, but any other clustering technique can be employed. At each iteration, the method adjusts the centroid position with respect to the cluster centers by minimizing the intra-cluster variance in the feature space, i.e.,:

$$\sum_{q=1}^{n_j}\sum_{v=1}^{V_j} ||\mathbf{x}^j_q - \mathbf{m}_v||^2 \tag{2}$$

where $\mathbf{m}_v$ is the centroid of cluster v. For the land-cover classes Ω_{Cov}, it is reasonable to assume that the cluster having the highest number of labeled units represents the dominant polygon class. For the semantically aggregated classes Ω_{Sem}, which may include several land-cover classes, the method removes the cluster having the smallest number of labeled units which has the highest probability to be wrongly associated to its polygon label.

B. Semantic Decomposition

The spatial decomposition step allows us to discard most of the pixels having the highest probability of being associated to wrong labels. Then, the main goal of the semantic decomposition step is to ensure that all the land-cover classes aggregated under the same semantic label are identified. Let us focus on the generic semantic class $\omega_k \in \Omega_{\text{Sem}}$. In the considered implementation, we assume to know the number of land-cover classes of ω_k. First, we fit a multivariate Gaussian Distribution to the labeled units belonging to the semantic class by considering its number of modes (i.e., number of land-cover classes). Then, for each pixel $\mathbf{x}_i$ (still associated to the ω_k label after the spatial decomposition step), we calculate the vector of Mahalanobis distances from each Gaussian mode as follows:

$$\mathbf{D}_M(\mathbf{x}_i) = \sqrt{(\mathbf{x}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k)} \tag{3}$$

where $\boldsymbol{\mu_k}$ and $\boldsymbol{\Sigma}_k$ are the mean vector and the covariance matrix of the multivariate Gaussian distribution representing ω_k. The unit $\mathbf{x}_i$ is associated to the nearest natural class (i.e., Gaussian mode) from the spectral view point (i.e., the

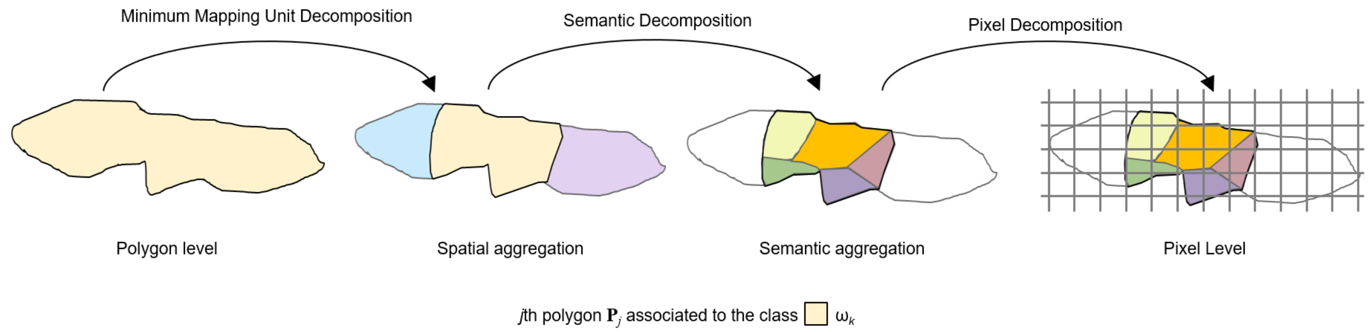


Fig. 3: Qualitative representation of the source domain properties in a vector polygon map. (i) Each polygon may present a spatial aggregation of parcels (homogeneous spectral areas) due to the minimum mapping unit of the map, (ii) each parcel may present a semantic aggregation of land-cover classes, and (iii) each polygon/parcel has to be mapped onto the geo-referenced pixel grid of the MS images thus leading to spurious pixels associated to partially correct labels.

class having minimum Mahalanobis distance). At the end of this step we have the decomposed thematic product $\mathbf{M}_{\Omega_2}^{t_1}$ characterized by a set of G land-cover classes $\Omega_2 = \{c_k\}_{k=1}^{G}$ where the pixels having the highest probability to be wrongly associated to their labels are neglected.

V. Design of Training Database

Due to the large availability of labeled units extracted from the map, we are in the condition of selecting the ones that will be use to generate a training database. To extract reliable and informative training samples from existing thematic products it is necessary to: 1) accurately represent the land-cover classes present in the scene from the spectral view point, 2) define a strategy for identifying pure spectral pixels associated to a valid label. Thus, even though the spatial decomposition strongly increases the probability of selecting labeled units correctly associated to their labels, we need to take into account that: (i) the cluster analysis may fail in detecting the pixels correctly associated to their label, and (ii) some polygons may be wrongly associated to their labels.

Under the reasonable assumption that the classes are Gaussian-distributed, we extract from each natural class present in the decomposed map $\mathbf{M}_{\Omega_2}^{t_1}$, the labeled units closer to the core of the distribution. Hence, it is reasonable to assume that these units have the highest probability of being correctly associated to their labels. Moreover, due to the semantic decomposition performed in the previous phase, we are in the condition of generating informative databases since we guarantee the selection of units belonging to all the land-cover classes present in Ω_{Sem}.

The number of labeled units per class is defined according to a stratified random sampling strategy, by taking advantage from the information provided by the thematic product in terms prior probabilities of the land-cover classes. Thus, the amount of pixels per class present in the original map is used as reference to define the number of units per class [43].

VI. Land-Cover Map Production

The last component of the proposed approach generates the high resolution land-cover map at pixel level. If MS data contemporary to the map are employed, the approach generates a thematic product characterized by a better geometric detail with respect to the initial one (i.e., supervised learning case). If recent MS data are considered, a standard domain adaptation technique is employed to produce an updated map (i.e., domain adaptation case). In the following details are given.

A. Supervised Learning

The main advantage of the proposed approach is the possibility of including a huge amount of units in the database of weak labeled pixels extracted from the map. Thus, the database can be sampled without replacement in order to generate a set of S statistically independent weak training sets $\{T_1, T_2, \cdots, T_S\}$. These weak training sets are then used to train an ensemble of classifiers combined with a majority voting rule. In this paper we use the Support Vector Machine (SVM) classifier but any classification technique can be used with the proposed approach.This classifier has been widely employed in the RS literature since it does not require an estimation of the statistical distributions of classes to perform the classification task [44]. Moreover, SVM is intrinsically effective compared to traditional classifiers due to the structural risk minimization principle, which leads to accurate classification results and good generalization capabilities [44]. Let $\{f_s\}_{s=1}^{S}$ be the decision functions of the ensemble of S classifiers trained using the S training sets extracted from the weak database of labeled units. The majority voting decision of the ensemble of SVMs for $\mathbf{x}_i$ is given by:

$$\mathbf{x}_i \in c_k \text{ if } \quad c_k = \underset{c_k \in \Omega_2}{\operatorname{argmax}}(\#\{f_s(\mathbf{x}_i) = c_k\}),\ s \in [1, S] \tag{4}$$

where $\#\{f_s(\mathbf{x}_i) = c_k\}$ is the number of SVMs whose decision for the pixel $\mathbf{x}_i$ is the class c_k.

B. Domain Adaptation

If the considered thematic product is outdated, the database of weak labeled units can be employed to classify a more recent MS image. Let $\mathbf{X}^{t_2}$ be the MS image acquired at time t_2 and used to perform the update. The multitemporal correlation between the MS images is employed to transfer the knowledge in a reliable but effective way. As we are considering a multitemporal dataset, we assume to deal with a *covariate shift* problem, where the prior probabilities of the classes in t_1 and t_2 are different [i.e., $P_{t_1}(\mathbf{x}) \neq P_{t_2}(\mathbf{x})$], while the conditional probabilities are almost the same [i.e., $P_{t_1}(c_k|\mathbf{x}) \approx P_{t_2}(c_k|\mathbf{x})$ with $c_k \in \Omega_2$].

In the considered implementation, we exploit the semisupervised LapSVM [10] to maintain consistency with the supervised learning step. LapSVM has been extensively applied to RS domain adaptation problems [9], [10] since it models the data distribution by using both the labeled pixels and the information provided by the high number of available unlabeled pixels. LapSVM formulation takes advantage from both the kernel function of the SVM and the graph Laplacian for manifold regularization. The data are first projected into a high dimensional feature space by means of the SVM kernel function, thus increasing the separability of the labeled units. Then, the intrinsic geometry of marginal distribution of data is captured by a graph in which nodes are both labeled and unlabeled units connected by weights [45]. The weights are calculated by minimizing the regularized function representing the graph in the kernel space, thus improving the estimate of the marginal distribution of the considered land-cover classes. We refer to [10] for more details on LapSVM. Although LapSVM allows us to face the *covariate shift* problem, more sophisticated domain adaptation method can be employed [46]. Similar to the supervised classification step, $\mathbf{X}^{t_2}$ is classified by an ensemble of LapSVM classifiers using the weak database of labeled units $\{T_1, T_2, \cdots, T_S\}$ derived from the decomposed map.

VII. Dataset Description

A. Dataset 1: Czech Republic

To assess the effectiveness of the proposed system in updating outdated thematic products, we considered a crop type vector map of Czech Republic generated in the framework of the Sen2Agri project [47]. The data used to generate this map are Sentinel 1A, Sentinel 2A, Landsat 7 (L7), Landsat 8 (L8) images, the Crop Parcel Dataset [Czech Land Parcel Identification System (LPIS)], in situ crop data, IACS (crop declaration data) and IACS (OTCS results - ground truth data) [48]. The RS data were acquired from November 2014 to September 2015 to characterize the main annual cultivations. The map is characterized by 7 classes, where four of them present semantic aggregation (see Tab. I). In greater detail, “winter cereals”, “spring cereals” and “fodder crops” present three land-cover classes, while “annual crops” includes five land-cover classes. The map has been aggregated at polygon level according to the GIS-tabase Czech LPIS [49]. Almost 20% of the polygons of the full Czech LPIS dataset present more cultivations in a single polygon. The crop label has been assigned following the majority rule criterion.

For the experimental analysis, we considered a portion of the whole thematic product (5129 km^2). The coordinates of the central point of the study area are 50.272588 latitutde, 14.354876 longitude (see Fig. 4). In situ data acquired on 2016 were used to quantitatively evaluate the obtained updated LC map. The spatial distribution of the reference data is

TABLE I: Semantic properties of the crop type map. (Czech Republic dataset)

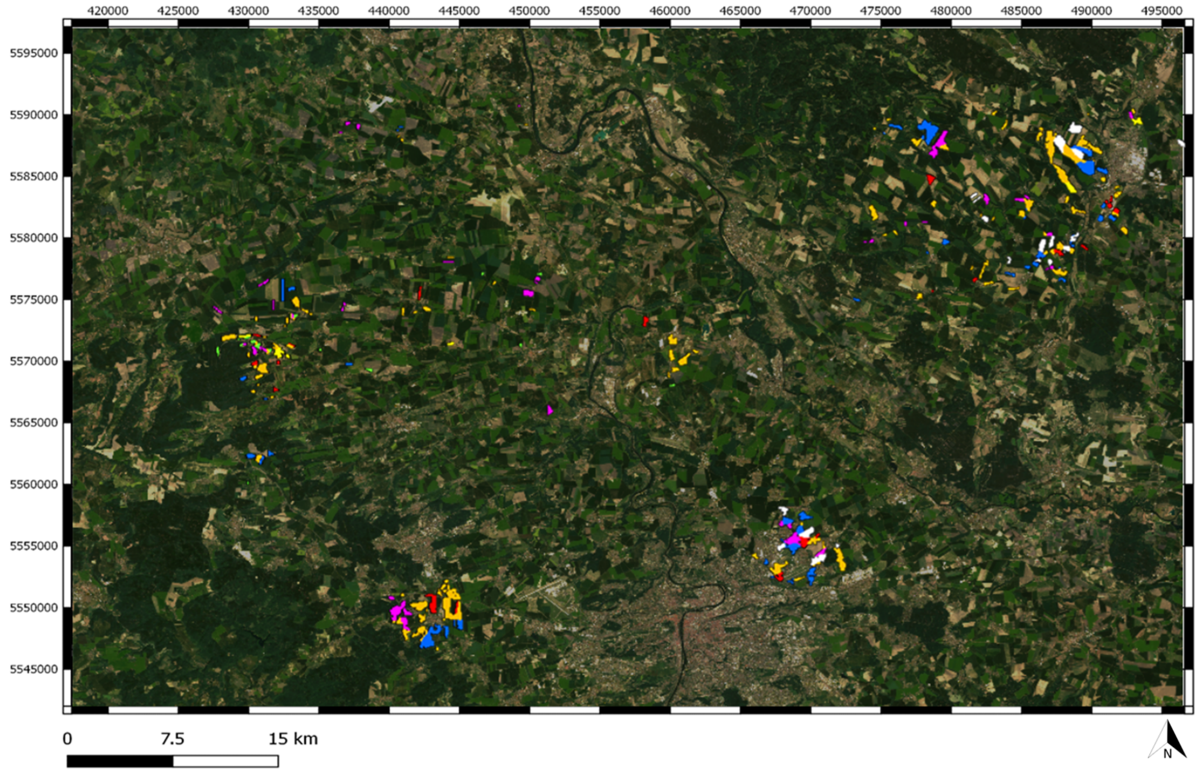
Map Legend	Class Type	Land-Cover Classes
Rapeseed	Ω_{Cov}	-
Winter Cereals	Ω_{Sem}	winter wheat winter triticale winter barley
Spring Cereals	Ω_{Sem}	spring barley oat spring wheat
Sugar Beet	Ω_{Cov}	-
Maize	Ω_{Cov}	-
Fodder Crops	Ω_{Sem}	alfalfa grass trefoil
Annual Crops	Ω_{Sem}	Soy Peas Poppy Mustard Wheat

TABLE II: Reference data collected by field survey in 2016 divided per class. The data have been used to validate the results obtained when classifying the 2016 TS of L8 images with the training set extracted from the 2015 crop type map. (Czech Republic dataset)

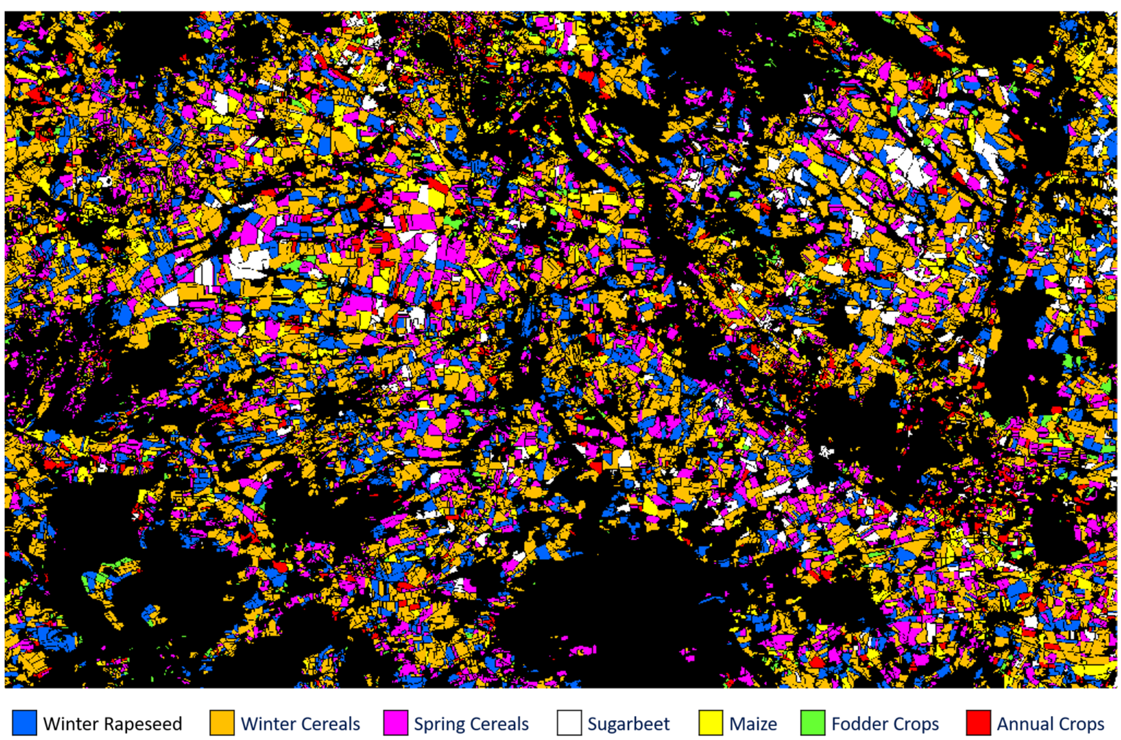
ID	Class	# Validation Units
ω_1	Rapeseed	4932
ω_2	Winter Cereals	9177
ω_3	Spring Cereals	2259
ω_4	Sugar beet	2855
ω_5	Maize	437
ω_6	Fodder Crops	119
ω_7	Annual Crops	1200

TABLE III: L8 images used in the experiments. The TS acquired at time t_1 (contemporary to the map) was used to perform the spatial, the semantic decomposition and to generate weak training set. The TS acquired at time t_2 was classified to generate the updated land-cover map.

TS of L8 images (t_1)	TS of L8 images (t_2)
13/01/2015	31/12/2015
18/03/2015	04/03/2016
19/04/2015	21/04/2016
06/06/2015	24/06/2016
09/08/2015	27/08/2016
12/10/2015	28/09/2016



(a)



(b)

Fig. 4: Czech Republic dataset: (a) 2016 validation dataset superimposed on the true color composite of the L8 image acquired on the 06^{th} June 2015, (b) outdated thematic product representing the 2015 crops. Coordinates are reported in the UTM WGS84 33N system.

represented in Fig. 4, while Tab. II shows the number of labeled units divided per class. Please note that the considered study area is complex due to the crop rotation practice which leads to many land-cover changes on the ground. An accurate extraction of reliable and informative labeled units from the initial map is thus fundamental to generate an accurate land-cover product.

The satellite optical data considered are L8 images, due to the availability of these data in 2015 (i.e., contemporary to the considered thematic product). The L8 spectral channels considered are the seven bands acquired at 30m spatial resolution. Thus, each pixel is characterized by 42 features. To perform the source domain modeling and the domain adaptation step, we considered a TS of six L8 images acquired in 2015 and 2016, respectively (see Tab. III). The acquisition dates of the considered TS allow us to model the phenological cycle of the crops present in the study area in both years. Clouds were detected considering the Fmask algorithm [50] and removed according to [51].

B. Dataset 2: France

To assess the capability of the proposed approach to increase the spatial resolution of existing thematic products, we considered the 2018 Corine Land Cover (CLC) generated by the European Environment Agency. The classification scheme is composed of 44 classes (mixed land-cover and land-use classes) with 25 ha minimum mapping unit. This map is generated and updated at national level by means of visual interpretation of satellite images. This data set is located in France and is characterized by a spatial extent of 1840 km^2. The coordinates of the central point of the study area are 45.687477 latitutde, 4.625595 longitude. The complex legend of the thematic product is suitable to test the capability of the proposed approach to extract a reliable and informative training set. In particular, in the considered study area there are seven Ω_{Cov} classes, seven Ω_{Use} classes, four Ω_{Spa} classes and two Ω_{Sem} classes (see Tab. IV).

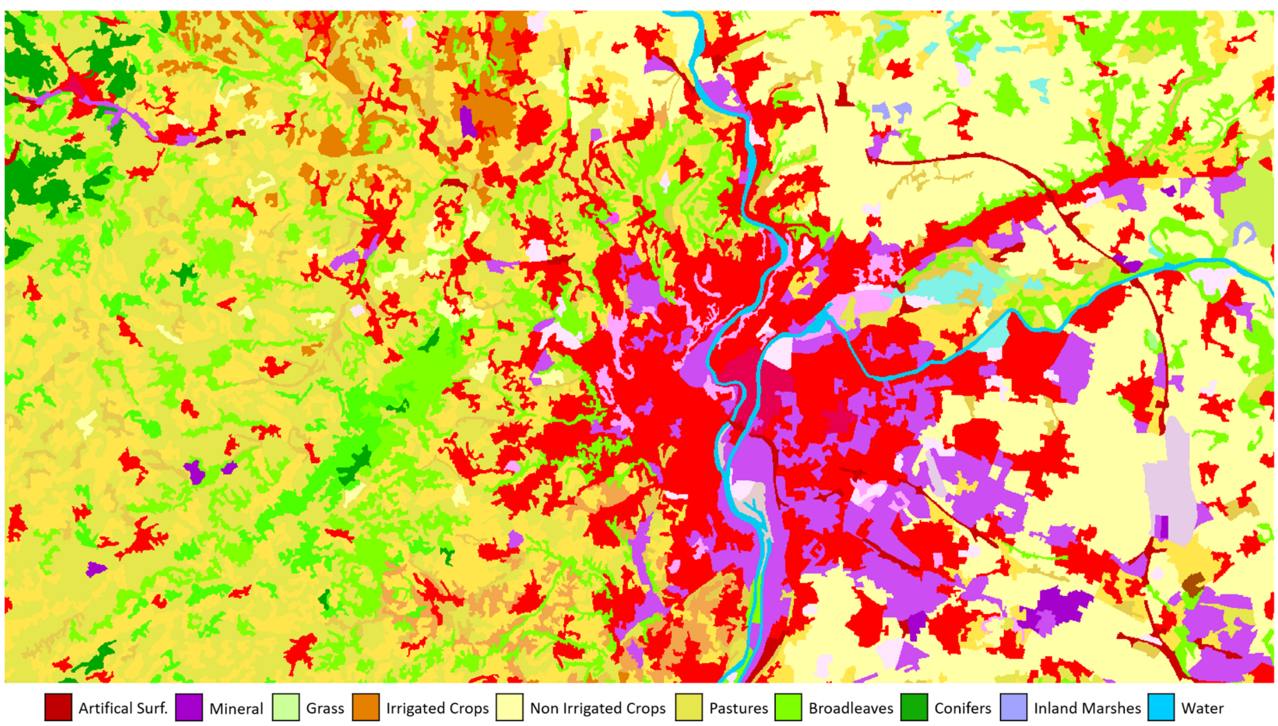
The satellite optical data considered are Sentinel 2 images contemporary to the map. In particular, we considered a TS of four cloud-free Sentinel 2 images (see Tab. V for the acquisition dates). The Sentinel 2 spectral channels considered are the four bands acquired at 10m and the six bands acquired at 20m spatial resolution. This leads to a feature vector of 40 spectral channels. The data were downloaded atmospherically corrected directly from the ESA's Sentinel 2 Scientific Data Hub [52].

To quantitatively evaluate the accuracy of the updated land-cover maps, we employed a reference dataset made up of 1023 pixels manually labeled by photo-interpretation and distributed all over the region. First, the prior probabilities of the classes were estimated by considering the information provided by the CLC Map. Then, a stratified random sampling strategy was applied to establish the validation samples locations. Finally, the label of each sample was defined by photo-interpretation by visually checking both Sentinel 2 data and ESRI ArcGIS Online World high-resolution aerial optical images. The spatial distribution of the reference data is represented in Fig. 5, where the scale of the samples is exaggerated to improve their visibility. The number of labeled units divided per class is reported in Tab. VI.

TABLE IV: Semantic properties of the 2018 CLC map for the considered study area. (France dataset)

CLC Class	Type	Land-Cover Classes
Continuous urban fabric	Ω_{Cov}	Artificial Surfaces
Discontinuous urban fabric	Ω_{Spa}	Artificial Surfaces Bare Soil Vegetated Areas
Industrial Units	Ω_{Use}	Artificial Surfaces
Road and rail networks	Ω_{Use}	Artificial Surfaces
Port areas	Ω_{Use}	Artificial Surfaces Bare Soil
Airports	Ω_{Use}	Artificial Surfaces Bare Soil Vegetated Areas
Mineral extraction sites	Ω_{Cov}	Mineral Site
Green urban areas	Ω_{Use}	Parks in the Cities Trees in the Cities
Sport facilities	Ω_{Use}	Artificial Surfaces Grass
Non-irrigated arable land	Ω_{Sem}	Cereals Legumes Fodder crops
Permanently irrigated land	Ω_{Sem}	Arable crops Non-permanent grass Greenhouses Crops
Pastures	Ω_{Cov}	Dense grass cover
Complex cultivations	Ω_{Spa}	Annual crops Pasture Permanent crops
Agriculture and vegetation	Ω_{Spa}	Agriculture Grass
Broadleaved forest	Ω_{Cov}	Broadleaved
Coniferous forest	Ω_{Cov}	Conifers
Mixed forest	Ω_{Spa}	Broadleaved Conifers
Natural grasslands	Ω_{Cov}	Grass
Inland Marshes	Ω_{Cov}	Inland Marshes
Water courses	Ω_{Use}	Water
Water bodies	Ω_{Use}	Water

(a)



(b)

Fig. 5: France dataset: (a) reference data superimposed on the true color composite of the Sentinel 2 image acquired on the 26^{th} August 2018, (b) original thematic product. Coordinates are reported in the UTM WGS84 31N system. The scale of the validation units is exaggerated to improve their visibility.

TABLE V: Sentinel 2 images dataset classified by using the training set extracted from the 2018 CLC map (France dataset).

TS of Sentinel 2 images (t_1)
18/04/2018
27/06/2018
26/08/2018
05/10/2018

TABLE VI: 2018 reference data used to validate the classification results obtained on the 2018 Sentinel 2 images (France dataset).

ID	Class	# Validation Units
ω_1	Artificial Surfaces	110
ω_2	Mineral Site	28
ω_3	Grass	41
ω_4	Non Irrigated Crops	334
ω_5	Irrigated Crops	45
ω_6	Pastures	164
ω_7	Broadleaved	149
ω_8	Conifers	57
ω_9	Inland Marshes	26
ω_{10}	Water	69

VIII. Experimental Results

In this section, first we present the experimental setup, introducing the baseline methods used for comparison and defining the parameter setting used in the work. Then, the obtained decomposed maps are analyzed from the qualitative view point, whereas the results obtained in terms of updated land-cover products are quantitatively evaluated. Finally, an analysis on the quality of the extracted training set is carried out for the 2^{nd} dataset (France).

A. Experimental Setup

To prove the effectiveness of the proposed approach, we compared the results obtained with the tolerant noise Random Forest classifier [53] and a standard outlier filtering approach [35] used in the literature to extract labeled units from existing thematic products. When performing the domain adaptation, the proposed system was compared also with the standard LapSVM [10], while for the supervised learning analysis we considered the standard SVM with Radial Basis Function (RBF) kernel functions [54]. The parameters of the Random Forest classifier are tuned according to [53], where Pelletier *et al* suggest to use Random Forest classifier when dealing with noisy training sets (such as the one extracted from the thematic products) by setting the number of trees to build equal to 200, the number of input features per node equal to the square root of the total number of features, the maximum depth of the tree growth equal to 25 and the minimum number of instances in the node equal to 10.

To perform the spectral filtering step, in [35], Radoux *et al* suggest to tune the probabilistic iterative trimming considering $\alpha \in [0.05, 0.1, 0.2]$. In the following, we reported the best results that were achieved with $\alpha = 0.05$. For the supervised learning analysis, the proposed system employed an ensemble of five SVMs with RBF kernels. For the proposed system, the standard RBF SVM and [35], the optimal kernel parameters (i.e., the regularization parameter C and the spread of the kernel γ) were selected by a 5-fold cross-validation. For the domain adaptation analysis, we need to tune two regularization parameters of the LapSVM, namely γ_M and γ_L. While γ_M controls the complexity of the classifier decision function in the geometry of the marginal data distribution, γ_L controls its complexity in the associated Hilbert space. According to [9], [10] γ_M was set equal to 0.5 for both the baseline and the proposed methods, while γ_L was set equal to $\gamma_M/(u+l)^2$, where u and l are the numbers of unlabeled and labeled units, respectively.

B. Results: Source Domain Modelling

Fig. 6 reports some examples of the obtained map decomposition results by showing the original crop type maps (Fig. 6a, 6f, 6k, 6p, 6u), the spatially decomposed maps (Fig. 6b, 6g, 6l, 6q, 6v), the semantically decomposed maps (Fig. 6c, 6h, 6m, 6r, 6w), the false color representations of the NDVI derived from three L8 images of the considered TS (Fig. 6d, 6i, 6n, 6s, 6x) and the true color compositions of the L8 image acquired in April 2015 (Fig. 6e, 6j, 6o, 6t, 6y). The false color composition of the NDVI was stretched for visual enhancement to emphasize the different cultivations present in the scene.

From the results obtained it turned out that even though the units of the LPIS polygon database represent agricultural parcels managed by single farmers [49], more cultivations may be present in the same polygons. This is mainly due to the multiple cropping practice (growing two or more crops in the same piece of land in same growing seasons) or can be related to possible outdated information present in the database. However, the TS of images contemporary to the map allows the accurate discrimination of different crops present in the same polygon. For instance, in Fig. 6u, the two largest polygons associated to the “sugar beet” label include different cultivations (parcels characterized by different spectral behaviors) clearly visible in the false color composition of the NDVI (see Fig. 6x). In contrast the smallest “sugar beet” polygon is associated with an homogeneous area from the spectral view point and similar to the ones selected by the proposed system. The spatial decomposition step accurately removes the labeled units belonging to the minor clusters, thus increasing the probability of selecting units correctly associated to the “sugar beet” label (Fig. 6v). Similarly, in Fig. 6p the largest crop labeled as “maize” includes a parcel having spectral behavior similar to the “spring cereal” cultivation (see Fig. 6s), which is discarded by the spatial decomposition step. Note that no post-processing was performed on the decomposed maps and the results are presented at pixel level.

Due to the semantic aggregation of the map legend, it is necessary to guarantee the selection of labeled units belonging

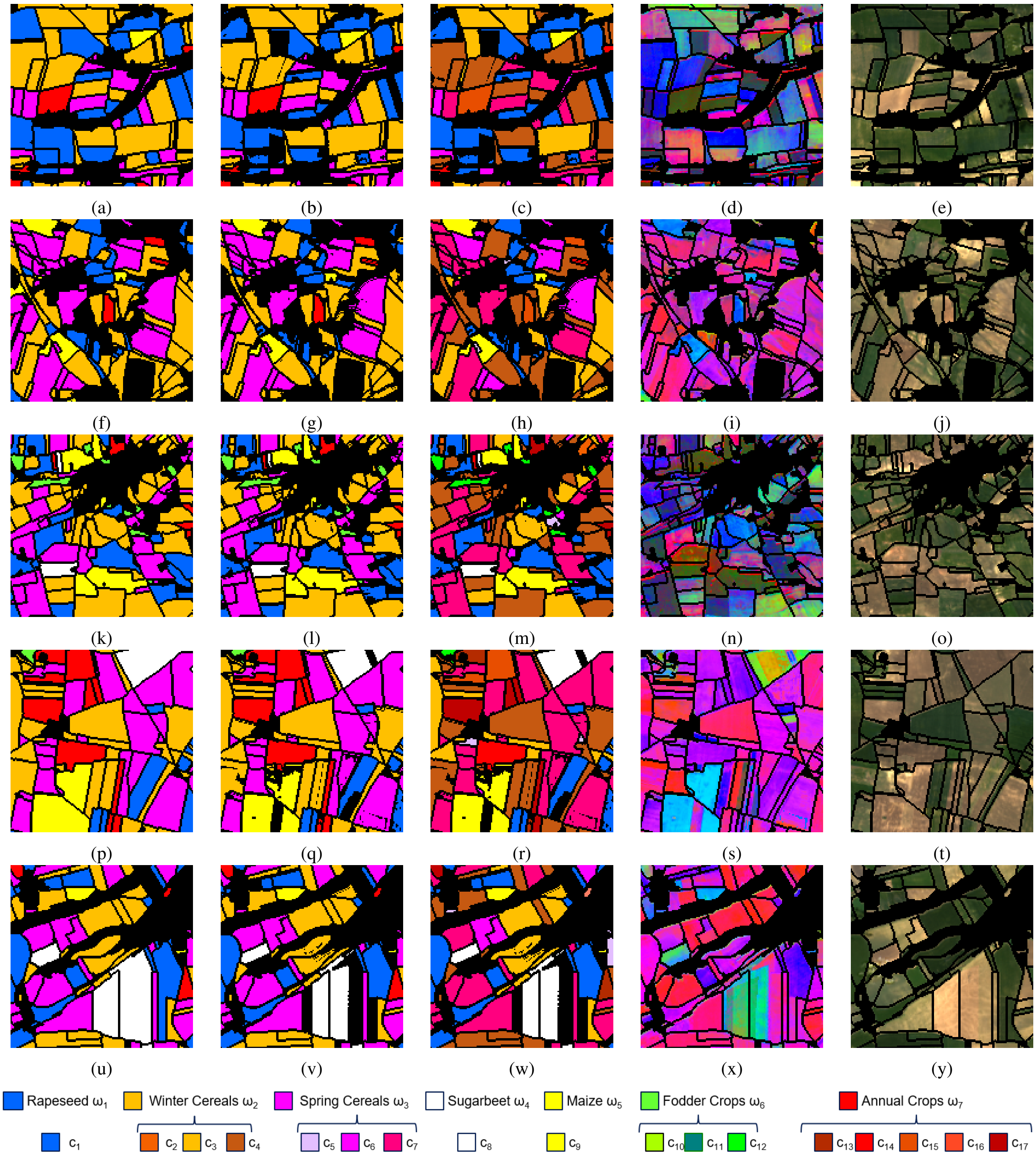


Fig. 6: Examples of map decomposition results of the 2015 crop type map: (a),(f),(k),(p),(u) original thematic products; (b),(g),(l),(q),(v) maps spatially decomposed; (c),(h),(m),(r),(w) maps semantically decomposed; (d),(i),(n),(s),(x) false color representations of three NDVI derived from the TS of the L8 images; and (e),(j),(o),(t),(y) true color compositions of the L8 image acquired on April 2015. The false color composition of the NDVI was stretched for visual enhancement to emphasize the different cultivations present in the scene. (Czech Republic dataset)

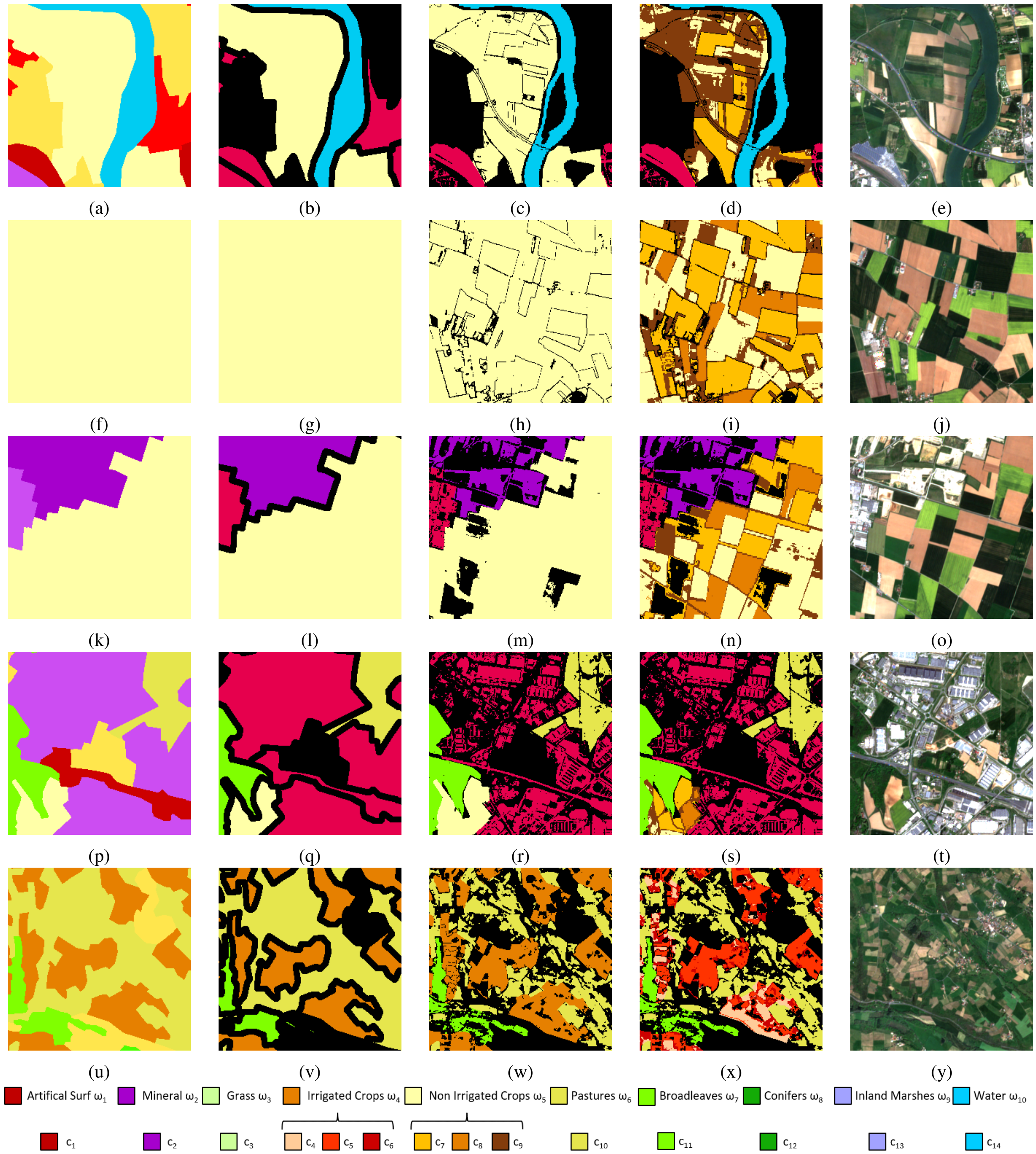


Fig. 7: Examples of map decomposition results of the 2018 CLC map: (a),(f),(k),(p),(u) original thematic products; (b),(g),(l),(q),(v) converted map; (c),(h),(m),(r),(w) maps spatially decomposed; (d),(i),(n),(s),(x) maps semantically decomposed; and (e),(j),(o),(t),(y) true color compositions of the Sentinel 2 image acquired on June 2018. (France dataset)

to all the land-cover classes belonging to the same semantic class to accurately model the class distribution. Also in this case, the qualitative evaluation confirms the effectiveness of the proposed approach. For instance, in Fig. 6b, the "winter cereals" class (i.e., ω_2) includes cultivation having different spectral behaviors (see Fig. 6d). Its semantic decomposition, reported in Fig. 6c, associates different parcels to different land-cover classes (i.e., c_3 and c_4). Fig. 6f depicts a similar example related to the "spring cereals" semantic label (i.e., ω_3), decomposed in Fig. 6h in c_6 and c_7 that clearly have different spectral behaviors with respect to most of the pixels present in the polygon (see Fig. 6i). Fig. 6s shows different crops associated with the "annual crops" label (i.e., ω_7) clearly visible in Fig. 6s and accurately discriminated in Fig. 6r (i.e., c_{13}, c_{14}, c_{15} and c_{17}). It is worth mentioning that the spatial decomposition of the previous step correctly removes minor crops associated to the wrong labels. However, since we need to transfer the labels to a multitemporal dataset, it is fundamental to accurately characterize all the land-cover classes included into the semantically aggregated ones, in order to face possible shift of the class distribution. Also in this case, no post-processing was performed on the decomposed maps in order to show the results at pixel level. Note that this step is fundamental to extract an informative database of weak labeled units from the source map. Thus, the missed selection of labeled units belonging to dominant land-cover classes present in the scene would result in a poorly representative training set that does not allow accurate land-cover map updates.

Fig. 7 reports several examples of the decomposition result obtained from the 2018 CLC map on the France dataset. Fig. 7a, 7f, 7k, 7p, 7u show the original thematic maps, Fig. 7b, 7g, 7l, 7q, 7v the converted thematic products, Fig. 7c, 7h, 7m, 7r, 7w the spatially decomposed maps, Fig. 7d, 7i, 7n, 7s, 7x the semantically decomposed maps, and Fig. 7e, 7j, 7o, 7t, 7y the true color compositions of the Sentinel 2 image acquired on June 2018. Differently from the crop type map, the 2018 CLC map presents a complex classification scheme characterized by land-cover, land-use classes, spatially and semantically aggregated classes. In the semantically converted thematic product, the spatially aggregated classes are removed. For instance, in Fig. 7a the polygons associated with the "Complex Cultivation Pattern" are discarded (Fig. 7b) since this class includes land-cover classes already present in the map legend (i.e., "Crops", "Pastures" and "Vegetation"). The land-use are converted into land-cover when possible according to the LCCS. In Fig. 7p the "Industrial Units" and "Roads" labels are converted into "Artificial Surfaces" since all these classes present similar spectral behavior (see Fig. 7q). Finally, the semantic classes are decomposed according to their number of land-cover classes. In the considered study, the semantic classes are "Irrigated Crops" and "Non Irrigated Crops". Both the classes present three land-cover classes according to the definition of the CLC map legend.

Due to the minimum mapping unit of 25 Ha, most of the polygons include many pixels wrongly associated to their labels. In such thematic product, the spatial decomposition step is fundamental to sharply increase the probability of selecting pixel correctly associated to their labels. Due to the high spatial resolution provided by the Sentinel 2 images (i.e., 10 m), we are in the condition of accurately removing wrong labeled units. For instance, Fig. 7p shows a urban area associated to the "Artificial Surfaces" label which includes also many "Grass" pixels. The spatial decomposition accurately removes those labeled units (see Fig. 7r) by correctly delineating the geometrical details of the buildings. In Fig. 7c the spatial decomposition step accurately removes the small island present in the river (see Fig. 7e), by keeping only the water pixels. Similarly, in Fig. 7k the pixels which do not belong to the mineral site are discarded from the polygon (see Fig. 7m and Fig. 7o). Accurate results are obtained also for the complex case of the semantically aggregated classes. In Fig. 7f a polygon associated to the "Non Irrigated Crops" label is reported. By removing the pixels belonging to the smallest parcels, the spatial decomposition automatically enhances the crop boundaries while keeping all the land-cover classes belonging to the semantic class (see Fig. 7h).

The importance of the semantic decomposition step can be assessed from the qualitative view point. Fig. 7n and Fig. 7i show the capability of the method of accurately detecting different cultivations belonging to the "Non Irrigated Crops" semantic class. The true color compositions of the Sentinel 2 image acquired in June (Fig. 7o and Fig. 7j) demonstrate the presence of different cultivations that should be accurately represented to obtain reliable classification results. Thus, the lack of one of those land-cover classes in the training set hampers the possibility of producing an accurate thematic product. Similar results are visible in Fig. 7b and Fig. 7v. Also in these cases, parcels characterized by different spectral responses are associated with the same semantic labels (Fig. 7e and Fig. 7y). However, the semantic decomposition allows us to accurately distinguish the land-cover classes present in the scene (see Fig. 7d and Fig. 7x).

C. Results: Updated Land-Cover Map Production

The qualitative evaluation of the decomposed maps is confirmed by the quantitative classification results of the obtained pixel land-cover maps. Tab VII and Tab. VIII report the classification accuracy of the obtained land-cover products derived by extracting the database of weak labeled units from the crop type map and the 2018 CLC map, respectively. The Producer Accuracy (PA %), the User Accuracy (UA %), the Fscore (F1 %) and the Overall Accuracy (OA %) metrics calculated on the validation set are reported for the baseline methods (on 5 trials) and the proposed system.

Let us focus the attention on the Czech Republic dataset. The Outlier Filtering method achieves an F1% ranging from 4.57 % (for the "Annual Crops" class) to 87.20% (for the "Rapeseed" class), whereas the Random Forest F1% ranges from 29.13% (for the "Sugar Beet" class) to 88.43% (for the "Winter Cereals class). By taking advantage from the multitemporal information, the LapSVM obtains better classification results with respect to the other baselines, with an F1% that ranges from a minimum of 57.82% (for the "Fodder Crops" class) to a maximum of 94.29% (for the

TABLE VII: Land-cover map update results of the Czech Republic dataset. The Overall Accuracy (OA%), User Accuracy (UA%), Producer Accuracy (PA%) and Fscore (F1%) are reported for: 1) the reference method based on a outlier filtering procedure [35]; 2) the Random Forest noise-tolerant classifier [53]; 3) the standard LapSVM [10]; and 4) the proposed unsupervised approach.

Map Legend	**Baselines**									**Proposed Method**		
	Outlier filtering [35]			**Random Forest [53]**			**LapSVM [10]**					
	PA%	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**
Rapeseed	77.46	99.80	87.22	79.57	99.89	88.58	95.28	93.20	94.23	89.50	96.78	93.00
Winter Cereals	85.56	82.06	83.77	95.48	79.57	86.80	97.23	89.01	92.94	96.14	93.52	94.81
Spring Cereals	90.30	63.36	74.47	66.58	53.66	59.43	61.52	74.84	67.53	71.92	97.96	82.95
Sugar beet	69.64	93.40	79.79	17.20	99.63	29.34	69.56	96.65	80.90	96.59	89.18	92.73
Maize	95.24	53.09	68.18	43.20	94.40	59.27	52.59	92.29	67.00	72.94	62.93	67.57
Fodder Crops	64.87	30.64	41.62	25.24	41.54	31.40	45.59	79.69	58.00	76.96	55.53	64.51
Annual Crops	2.38	100	4.65	91.85	38.25	54.01	83.65	47.08	60.25	82.51	57.42	67.71
OA%	**76.73**			**73.85**			**85.18**			**89.55**		

TABLE VIII: Classification results of the France dataset. The Overall Accuracy (OA%), User Accuracy (UA%), Producer Accuracy (PA%) and Fscore (F1%) are reported for: 1) the reference method based on a outlier filtering procedure [35]; 2) the Random Forest noise-tolerant classifier [53]; 3) the standard RBF SVM [54]; and 4) the proposed unsupervised approach.

Map Legend	**Baselines**									**Proposed Method**		
	Outlier filtering [35]			**Random Forest [53]**			**SVM [54]**					
	PA%	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**	**PA%**	**UA%**	**F1%**
Artificial Surfaces	85.82	64.39	73.58	64.18	88.03	74.24	81.82	82.42	82.12	89.09	94.23	91.59
Mineral extraction sites	14.29	6.33	8.77	36.43	92.73	52.31	50.00	60.87	54.90	92.86	92.86	92.86
Grass	74.63	83.61	78.87	78.54	90.45	84.08	85.37	90.67	87.94	90.24	97.37	93.67
Non Irrigated Crops	25.75	81.59	39.15	74.55	75.18	74.86	78.38	89.47	83.56	91.32	92.99	92.15
Irrigated Crops	48.44	19.75	28.06	49.33	28.46	36.10	64.00	48.65	55.28	91.11	61.19	73.21
Pastures	84.02	51.73	64.03	73.41	54.93	62.84	86.71	61.99	72.30	92.07	87.79	89.88
Broadleaves	53.02	78.84	63.40	81.48	77.32	79.35	77.18	84.06	80.47	88.59	92.31	90.41
Conifers	51.93	73.27	60.78	71.58	89.08	79.38	80.70	81.56	81.13	91.23	88.14	89.66
Inland Marshes	47.69	13.36	20.87	0	0	0	20.00	35.14	25.49	57.69	71.43	63.83
Water	88.41	100	93.85	92.75	99.07	95.81	83.48	91.43	87.27	91.30	100	95.45
OA%	**54.40**			**71.43**			**77.77**			**89.93**		

"Rapeseed" class). The proposed system outperforms all the baseline techniques, with a minimum F1% of 64.51% (for the "Fodder Crops" class) and a maximum F1% of 94.81% (for the "Winter Cereals" class).

Both the Outlier Filtering and the Random Forest methods obtain very poor classification accuracy on semantic aggregated classes Ω_{Sem}. In particular, the worst results are obtained on the "Annual Crops" class (i.e., F1% of 4.57 and 54.05 for the Outlier Filtering and Random Forest, respectively), which includes five land-cover classes. Due to the large amount of changes present in the scene, poor classification accuracy are achieved also on some land-cover classes (i.e., F1% of 28.16 on the "Sugar Beet" class with the Random Forest). This problem is alleviated by the use of the LapSVM. However, the most balance classification results are achieved by the proposed system. Thus, even though the considered classification problem is complex due to the crop rotation practice (which leads to many changes on the ground) and the complex structure of the semantically aggregated classes, the proposed system is able to achieve good F1% for all the land-cover classes. This is confirmed by the OA %, which is 89.55 % for the proposed approach, which is much higher than those obtained by the baseline methods (i.e., 76.73, 73.85 and 85.18 for the Outlier Filtering, the Random Forest and the

LapSVM classifier, respectively).

Similar results are obtained on the pixel land-cover method generated by extracting the labeled units from the 2018 CLC map on the France dataset. The proposed system sharply improves the classification accuracy with respect to the baseline methods by achieving an OA % of 89.93% compared to 54.40%, 71.43% and 77.77% of the Outlier Filtering, the Random Forest and the SVM classifier, respectively. In particular, the F1% achieved by the proposed system ranges from a minimum of 63.83% (for the "Inland Marshes" class) to a maximum of 95.45% (for the "Water" class). The Outlier Filtering method ranges from 8.77% (for the "Mineral Site" class) to a maximum of 93.85% (for the "Water" class), whereas the Random Forest ranges from 0% (for the "Inland Marshes" class) to 95.81% (for the "Water" class). The best results among the baseline are achieved by the standard SVM that reaches an F1% ranging from a minimum of 25.49% (for the "Inland Marshes" class) to a maximum of 87.94% (for the "Grass" class). The Outlier Filtering fails in modeling the land-cover classes penalized by the spatial aggregation rule (i.e., "Mineral Site" and "Artificial Surfaces") and the semantic aggregated classes (i.e., "Irrigated Crops" and "Non Irrigated Crops"). Thus, discarding the outliers using a spectral filtering technique for such complex land-cover class distributions leads to the removal of informative labeled units which are fundamental for accurately training the classifier. Similar problems are encountered also with the Random Forest classifier, which is not able to deal with the semantically aggregated classes as well as to manage classes having a low number of training samples (i.e., "Inland Marshes"). In contrast, the standard RBF SVM can handle the noisy training set extracted from the map, even though some classes achieves low F1% (e.g., Non "Irrigated crops" and "Inland Marshes").

Due to the capability of the system of extracting reliable an informative training samples, high classification accuracies are achieved on all the land-cover classes. In particular, the spatial decomposition results strongly increase the probability of selecting correctly labeled units. For instance, on the "Artificial Surfaces" class the proposed system achieves an F1% of 91.59% compared to the 73.58%, 74.24% and 82.12% obtained by the Outlier Filtering, the Random Forest and the SVM, respectively. Note that due to the minimum mapping unit of 25 Ha, the "Artificial Surfaces" polygons include many "Grass" pixels which are discarded by the proposed system. Similar results are obtained on the "Mineral Site" class, where the proposed system achieves an F1% of 92.86% compared to the 8.77%, 52.31% and 54.90% of the Outlier Filtering, the Random Forest and the SVM, respectively. Also in this thematic product, the baseline methods achieve low classification accuracy on the Ω_{Sem}. For instance, the F1% obtained for the "Non Irrigated Crops" are 39.15%, 74.86% and 83.56% for the Outlier Filtering, the Random Forest and the SVM, respectively, compared to the 92.15% of the proposed system.

D. Results: Weak Training Set Analysis

In this section, we evaluate the quality of the extracted weak training set. First, the sensitivity of the OA% of the proposed approach versus the considered number of training samples was analyzed. Fig. 8 reports the OA% obtained by increasing the number of samples from 1641 to 8271 for the outlier filtering procedure [35], the Random Forest classifier [53], the standard SVM [54] and the proposed method. Note that for each trial, the number of samples selected per class has been calculated according to the stratified random sampling strategy considering the original thematic product. From the results obtained, one can notice that the proposed approach outperforms the baseline methods for all the trials. Moreover, it is slightly affected by the number of training samples by obtaining an OA% that ranges from almost 85% to 90%. This proves the effectiveness of the method used for the selection of the training samples, as increasing the number of samples we increase the amount of information given to the classifier.

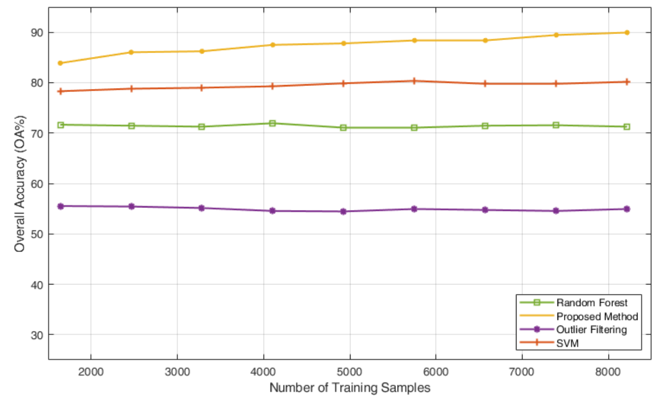


Fig. 8: Overall Accuracy (OA%) classification performance versus the number of training samples for the: 1) outlier filtering procedure [35]; 2) the Random Forest classifier [53]; 3) the standard SVM [54]; and 4) the proposed method.

Then, we evaluate the reliability of the labeled units extracted from the map. The main goal of the proposed approach is to extract training units that: (i) have the highest probability to be correctly associated to their labels, and (ii) are representative of the land-cover class distribution. Although it is reasonable to assume that classifiers trained with high quality samples achieve high classification accuracy, this is an indirect measure that does not guarantee that the training set is made up of reliable training sample. To verify the quality of the extracted labeled units, a quantitative evaluation of the training samples was performed by checking their labels via photo-interpretation of both Sentinel 2 data and ESRI ArcGIS Online World high-resolution aerial optical images. To this end, we focused the attention on one of the five training set automatically extracted by the method and we randomly selected the 10% of samples per class (for a total number of 822 samples checked). Differently from the previous experimental results, this analysis has been carried out only for the 2^{nd} dataset (France), since a reliable identification of the different crop types in Czech Republic is not possible by photo-interpretation.

The proposed method was compared with the Bayesian uncertainty evaluation strategy, which is used in sample se-

TABLE IX: Comparison between the training labels automatically extracted from the thematic product and the ones assigned by photo-interpretation and classification results obtained on the validation set. The Overall Accuracy (OA%) and Fscore (F1%) are reported for: 1) the proposed method; and 2) a Bayesian uncertainty method. The number of training units extracted per class is reported.

# training units extracted	Map Legend	Training Set		Validation Set	
		Proposed Method	Bayesian Method	Proposed Method	Bayesian Method
		F1%	F1%	F1%	F1%
587	Artificial Surfaces	82.76	83.76	91.59	76.70
282	Mineral extraction sites	94.34	65.12	92.86	47.37
281	Grass	83.58	96.30	93.67	75.79
2768	Non Irrigated Crops	95.09	95.34	92.15	82.20
950	Irrigated Crops	93.26	97.33	73.21	29.03
1473	Pastures	88.42	87.97	89.88	77.42
974	Broadleaves	76.54	85.85	90.41	75.19
367	Conifers	89.74	86.15	89.66	68.09
257	Inland Marshes	63.16	63.16	63.83	32.08
278	Water	100	88.52	95.45	96.24
8217	**OA%**	**88.81**	**90.02**	**89.93**	**74.78**

lection [55]. To this end, first the prior probabilities and the conditional density functions of the land-cover classes were estimated by using the 2018 CLC thematic product and the TS of Sentinel 2 images. Then, for each sample, we computed the Bayes decision rule that maximizes the posterior probability (i.e., that minimizes the error probability in the sense of Bayesian theory) [55]. Only the most reliable samples per class were selected to generate the training set.

Tab. IX reports the comparison between the labels of the training units automatically extracted from the map and the ones assigned by photo-interpretation for both the proposed method and the Bayesian strategy. For each class, the number of samples extracted is presented. Moreover, the classification results obtained on the validation set with the considered training sets are reported. In particular, the OA% and F1% scores are presented for the proposed method and the Bayesian uncertainty strategy. Note that the results obtained with the proposed method on the validation set are the same of Tab. VIII and are replicated here to help the reader in the comparison with the Bayesian method. As expected the Bayesian approach is able to select more reliable samples, by selecting the samples closer to the cores of the land-cover Gaussian distributions. However, the results on the validation set demonstrate the importance of selecting also training units that describes more complex classes and better represent their distributions. Although the training set extracted with the proposed method is slightly less accurate compared to the Bayesian ones, the proposed approach allows for a database of labeled units which is more representative of the considered study area. This is particularly evident for semantically aggregated classes such as “Irrigated Crops”, where the selection of most reliable training units leads to a poor representation of all the land-cover classes aggregated under the same semantic label (i.e., the F1% of the Bayesian method is 29.03 on the validation set compared to the 73.21 of the proposed method). In contrast, due to the semantic and spatial decomposition steps, the proposed method achieves high F1% scores for all the land-covers.

IX. Conclusion

In this paper we have presented a novel approach to the automatic extraction of labeled units from existing cartographic products. The goal is extract training samples having the highest probability of being correctly associated to their labels according to the information provided by the satellite RS data. The main assumptions of the approach are that: (i) RS data contemporary to the map used for extracting the labels of the units are available, (ii) the vector map has been converted into raster and accurately co-registered to the RS data, and (iii) the map legend has been converted into an exhaustive set of classes discriminable with the considered RS image. In the considered implementation we focused the attention on satellite MS optical data. To prove the effectiveness of the proposed approach we considered two thematic products characterized by different spatial properties and classifications scheme: a 2015 crop type map of the Czech Republic and the 2018 CLC map representing a study area located in France.

The crop type map has a better spatial resolution compared to the 2018 CLC map (i.e., smaller mapping units). However, it represents a complex dataset since it is characterized by a classification scheme made up very similar cultivations, where many semantic classes are present. Moreover, due to the crop rotation practice, the update of this thematic product is not trivial since many changes happened on the ground. In contrast, the 2018 CLC map is characterized by a minimum mapping unit of 25 Ha which leads to large polygons that include many pixels associated to wrong labels. Moreover, its classification scheme is characterized by spatially aggregated classes, semantically aggregated classes, land-use and land-cover classes. Thus, this dataset demonstrates the importance of performing the spatial and semantic decomposition to extract a reliable and informative database of labeled units.

From the results obtained one can observe that the proposed system outperforms the baseline methods in both the experiments. By accurately understanding the properties of the considered map, the proposed approach is able to convert the thematic product into a set of land-cover classes that can be discriminated by the spectral properties of the MS data. For each polygon, the approach accurately extracts (in an unsupervised way) the pixels which have high probability to be correctly associated to their labels. This spatial decomposition step strongly increases the probability of extracting reliable labeled units from the maps. Although the spatial decomposition is fundamental to increase the probability of selecting correctly labeled units, to generate an informative training set it is fundamental to accurately decompose the thematic product from the semantic view point. The importance of this step is highlighted by the capability of the proposed approach to achieve accurate classification results on the semantically aggregated classes.

As future developments, we aim to exploit the proposed system to extract huge databases of labeled units from existing thematic products to train deep network tailored to the specific properties of RS data. Indeed, even though deep architectures typically outperform standard machine learning classification systems, their main bottleneck is the need of hundred of labeled units to train the network to avoid over-fitting problem. While in the computer vision community, huge databases of training samples have been created, when moving to the RS community we clash with the major problem of limited training data. In this context, the proposed approach is promising to generate in an unsupervised way large databases of weak training samples to train the network. Moreover, we plan to investigate the possibility of integrating the proposed method with a further step which aims to detect new land-cover classes that may appear in the most recent RS data.

X. Acknowledgment

This work was supported by the European Space Agency (ESA) under project "S2-4Sci Land and Water - Multitemporal Analysis". Authors would like to thank Dr. Šavelková Lucie Ing. Ph.D., the Czech Paying Agency (SZIF) and the Czech-Agri national demonstration of the Sen2Agri project funded by the ESA Data User Element for providing the validation datasets.

References

[1] S. Patra, K. Bhardwaj, and L. Bruzzone, "A spectral-spatial multicriteria active learning technique for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 12, pp. 5213–5227, 2017.

[2] A. Singla, S. Patra, and L. Bruzzone, "A novel classification technique based on progressive transductive svm learning," *Pattern Recognition Letters*, vol. 42, pp. 101–106, 2014.

[3] 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, nov. 2006.

[4] D. Tuia, M. Volpi, M. Trolliet, and G. Camps-Valls, "Semisupervised manifold alignment of multimodal remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, 2014.

[5] M. Chi and L. Bruzzone, "An ensemble-driven k-nn approach to ill-posed classification problems," *Pattern Recognition Letters*, vol. 27, no. 4, pp. 301–307, 2006.

[6] J. Huang, A. Gretton, K. M. Borgwardt, B. Schölkopf, and A. J. Smola, "Correcting sample selection bias by unlabeled data," in *Advances in neural information processing systems*, 2006, pp. 601–608.

[7] G. Jun and J. Ghosh, "Spatially adaptive classification of land cover with remote sensing data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 7, pp. 2662–2673, 2011.

[8] 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.

[9] 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.

[10] 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.

[11] T. S. Caetano, T. Caelli, D. Schuurmans, and D. A. C. Barone, "Graphical models and point pattern matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1646–1663, 2006.

[12] 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.

[13] B. Luo and E. R. Hancock, "Structural graph matching using the em algorithm and singular value decomposition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1120–1136, 2001.

[14] 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.

[15] 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.

[16] S. Rajan, J. Ghosh, and M. M. Crawford, "Exploiting class hierarchies for knowledge transfer in hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3408–3417, 2006.

[17] H. L. Yang and M. M. Crawford, "Spectral and spatial proximity-based manifold alignment for multitemporal hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 1, pp. 51–64, 2016.

[18] L. Bruzzone and C. Persello, "A novel approach to the selection of spatially invariant features for the classification of hyperspectral images with improved generalization capability," *IEEE transactions on geoscience and remote sensing*, vol. 47, no. 9, pp. 3180–3191, 2009.

[19] 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.

[20] 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.

[21] 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.

[22] M. A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang, "Modis collection 5 global land cover: Algorithm refinements and characterization of new datasets," *Remote Sensing of Environment*, vol. 114, no. 1, pp. 168–182, 2010.

[23] 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," *In-*

ternational Journal of Applied Earth Observation and Geoinformation, vol. 13, no. 2, pp. 207–219, 2011.

[24] 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.

[25] B. Kosztra, G. Büttner, G. Hazeu, and S. Arnold, "Updated clc illustrated nomenclature guidelines," *Final Report by European Environmental Agency.*, 2017.

[26] A. Pérez-Hoyos, F. J. García-Haro, and J. San-Miguel-Ayanz, "A methodology to generate a synergetic land-cover map by fusion of different land-cover products," *International Journal of Applied Earth Observation and Geoinformation*, vol. 19, pp. 72–87, 2012.

[27] H. Balzter, B. Cole, C. Thiel, and C. Schmullius, "Mapping corine land cover from sentinel-1a sar and srtm digital elevation model data using random forests," *Remote Sensing*, vol. 7, no. 11, pp. 14 876–14 898, 2015.

[28] M. Lesiv, E. Moltchanova, D. Schepaschenko, L. See, A. Shvidenko, A. Comber, and S. Fritz, "Comparison of data fusion methods using crowdsourced data in creating a hybrid forest cover map," *Remote Sensing*, vol. 8, no. 3, p. 261, 2016.

[29] D. Schepaschenko, L. See, M. Lesiv, I. McCallum, S. Fritz, C. Salk, E. Moltchanova, C. Perger, M. Shchepashchenko, A. Shvidenko *et al.*, "Development of a global hybrid forest mask through the synergy of remote sensing, crowdsourcing and fao statistics," *Remote Sensing of Environment*, vol. 162, pp. 208–220, 2015.

[30] S. Fritz, I. McCallum, C. Schill, C. Perger, L. See, D. Schepaschenko, M. Van der Velde, F. Kraxner, and M. Obersteiner, "Geo-wiki: An online platform for improving global land cover," *Environmental Modelling & Software*, vol. 31, pp. 110–123, 2012.

[31] B. A. Johnson and K. Iizuka, "Integrating openstreetmap crowdsourced data and landsat time-series imagery for rapid land use/land cover (lulc) mapping: Case study of the laguna de bay area of the philippines," *Applied Geography*, vol. 67, pp. 140–149, 2016.

[32] C. Pelletier, S. Valero, J. Inglada, N. Champion, and G. Dedieu, "Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas," *Remote Sensing of Environment*, vol. 187, pp. 156 – 168, 2016.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] C. Lin, P. Du, A. Samat, E. Li, X. Wang, and J. Xia, "Automatic updating of land cover maps in rapidly urbanizing regions by relational knowledge transferring from globeland30," *Remote Sensing*, vol. 11, no. 12, p. 1397, 2019.

[38] 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.

[39] 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.

[40] K.-T. Chang, *Introduction to geographic information systems*. McGraw-Hill Higher Education Boston, 2006.

[41] 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.

[42] 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.

[43] 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.

[44] C.-h. Chen, *Handbook of pattern recognition and computer vision*. World Scientific, 2015.

[45] O. Chapelle, B. Scholkopf, and A. Zien, "Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]," *IEEE Transactions on Neural Networks*, vol. 20, no. 3, pp. 542–542, 2009.

[46] 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.

[47] P. Defourny, S. Bontemps, N. Bellemans, C. Cara, G. Dedieu, E. Guzzonato, O. Hagolle, J. Inglada, L. Nicola, T. Rabaute *et al.*, "Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the sen2-agri automated system in various cropping systems around the world," *Remote sensing of Environment*, vol. 221, pp. 551–568, 2019.

[48] "Final report: Czech agriculture national demonstrator," Tech. Rep. [Online]. Available: http://www.esa-sen2agri.org/wp-content/uploads/resources/technical-documents/CzechAgri-Final-Report-1.2.pdf

[49] V. Sagris, P. Wojda, P. Milenov, and W. Devos, "The harmonised data model for assessing land parcel identification systems compliance with requirements of direct aid and agri-environmental schemes of the cap," *Journal of Environmental Management*, vol. 118, no. Supplement C, pp. 40 – 48, 2013.

[50] S. Foga, P. L. Scaramuzza, S. Guo, Z. Zhu, R. D. Dilley, T. Beckmann, G. L. Schmidt, J. L. Dwyer, M. J. Hughes, and B. Laue, "Cloud detection algorithm comparison and validation for operational landsat data products," *Remote Sensing of Environment*, vol. 194, pp. 379–390, 2017.

[51] X. Zhu, F. Gao, D. Liu, and J. Chen, "A modified neighborhood similar pixel interpolator approach for removing thick clouds in landsat images," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 521–525, 2012.

[52] Scihub.copernicus.eu. Open Access Hub. URL: https://scihub.copernicus.eu/ (Last accessed on 30 May 2020).

[53] 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.

[54] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1778 – 1790, Aug. 2004.

[55] 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, 2011.