Specific alpine environment land cover classification methodology: Google Earth Engine processing for Sentinel-2 data

Elena Belcore [1,2] *; Marco Piras[1]; Edyta Wozniak[3]

[1]Politecnico di Torino, DIATI, Department of Environment, Land and Infrastructure Engineering. Corso Duca degli Abruzzi, 24, 10129 Torino, Italia- Email(s): name.surname@polito.it
[2]Politecnico di Torino, DIST, Interuniversity Department of Regional and Urban Studies and Planning. Viale Pier Andrea Mattioli, 39,10125 Torino, Italia - Email(s): name.surname@polito.it
[3]Centrum Badań Kosmicznych PAN, Bartycka 18A, 00-716 Warsaw, Poland - Email: ewozniak@cbk.waw.pl

TCIII-7: Landuse and Landcover Change Detection

KEY WORDS: Classification, Land Cover, Mountainous areas, Optical imagery, Machine Learning, Sentinel-2, Google Earth Engine, Pixel-based, random forest.

ABSTRACT:

Land Cover (LC) plays a key role in many disciplines and its classification from optical imagery is one of the prevalent applications of remote sensing. Besides years of researches and innovation on LC, the classification of some areas of the World is still challenging due to environmental and climatic constraints, such as the one of the mountainous chains. In this contribution, we propose a specific methodology for the classification of the Land Cover in mountainous areas using Sentinel 2, 1C-level imagery. The classification considers some specific high-altitude mountainous classes: clustered bare soils that are particularly prone to erosion, glaciers, and solid-rocky areas. It consists of a pixel-based multi-epochs classification using random forest algorithm performed in Google Earth Engine (GEE). The study area is located in the western Alps between Italy and France and the analyzed dataset refers to 2017-2019 imagery captured in the summertime only. The dataset was pre-processed, enriched of derivative features (radiometric, histogram-based and textural). A workflow for the reduction of the computational effort for the classification, which includes correlation and importance analysis of input features, was developed. Each image of the dataset was separately classified using random forest classification algorithm and then aggregated each other by the most frequent pixel value. The results show the high impact of textural features in the separation of the mountainous-specific classes the overall accuracy of the final classification achieves 0.945.

1. INTRODUCTION

Land Cover (LC) plays a key role in many disciplines, such as sustainable land management, land resource monitoring, urban vegetation mapping landscape ecology and climate-related researches (Feng et al., 2015; Rizeei et al., 2016; Sekertekin et al., 2017; Shelestov et al., 2017; Turner and Gardner, 2015). This makes the land cover, and land use, maps highly demanded (Carrasco et al., 2019; Delalay et al., 2019; Sekertekin et al., 2017; Thanh Noi and Kappas, 2018). LC maps are commonly derived from machine learning classifications of optical imagery, which represents one of the prevalent applications of remote sensing and it has been described by many authors (Carrasco et al., 2019; Delalay et al., 2019; Rizeei et al., 2016; Sekertekin et al., 2017; Shelestov et al., 2017; Sidhu et al., 2018; Thanh Noi and Kappas, 2018). During the last years, LC classification has done great steps forward: a wide range of free satellite medium-high optical imagery, specific classification algorithms, many processing platforms and machines with more and more high computational power are now available (Carrasco et al., 2019; Rizeei et al., 2016; Sidhu et al., 2018). Besides today's achievements, we still face some major constraints in LC classification, which can be distinguished in environmental constraints and technical constraints. Among the technical constraints, the poor temporal resolution of satellites is one of the most popular, but it has been partially overcome with the introduction of medium-high resolution satellites that increase the free data available and make possible integrating the datasets from different acquisitions. Although, data integration and high-spatial resolution requires to manage a large amount of data, but also a large amount of storage, as well as significant computing power and time (Carrasco et al., 2019). Managing satellite dataset requires considerable data storage capability and the high spatial resolution further increase this requirement. Indeed, the antinomy between spatial resolution and computational power is another very common technical constraint. During the last few years, some geographic cloud computing platforms that allow the analysis and storage of geographic data were born. These services (such as Google Earth Engine) decrease the computational and storage limits related to the satellite data processing (Kumar and Mutanga, 2018). The atmospheric disturbance and the high seasonal variability are some of the main environmental constraints in LC classification along with the topography variation. The topography strongly influences the spectral values in satellite imagery, especially in the case of steep areas. Indeed, the strong variability in the reflectance and the spread shadows are the main effect of the topography on satellite imagery, which may complicate the classification (Dorren et al., 2003). This is particularly true in narrow valleys of mountainous areas, where some mountainside are permanently shadowed in winter months. The LC classification of mountainous areas is affected by other major environmental constraints due to climatic conditions. For example, the snow cover prevents the classification in the winter season and the orographic rains and clouds, which mechanisms are influenced by the terrain (Houze, 2012), lower the possibility of accurate classifications. The combination of these factors makes the generation of the land cover maps of mountainous areas particularly challenging and generally recognized as fairly-low accurate (Dorren et al., 2003; Itten and Meyer, 1993).

In this contribution, we propose a specific methodology for the classification of the Land Cover in mountainous areas using Sentinel 2, 1C-level imagery. The classification considers some specific high-altitude mountainous classes: clustered bare soil areas that are particularly prone to erosion; glaciers; and solid-rocky areas. The methodology tries to overcome the above-mentioned environmental limitations and it consists of a pixel-based multi-epochs classification using random forest algorithm. We chose to perform the analysis using Google Earth Engine (GEE) because of its high computational speed and the large dataset of satellite imagery it makes easily available.

2. MATERIAL

2.1 Test study

The study interests the Wester Alpine arch between France and Italy and it extends for approximately 340000 hectares (Figure 1). The town of Cesana Torinese (44° 57′ 0″ N, 6° 48′ 0″ E) in Italy is the reference site of this study. Two Sentinel-2 tiles cover the study area: 31TGK (K) and 32TLQ (Q).
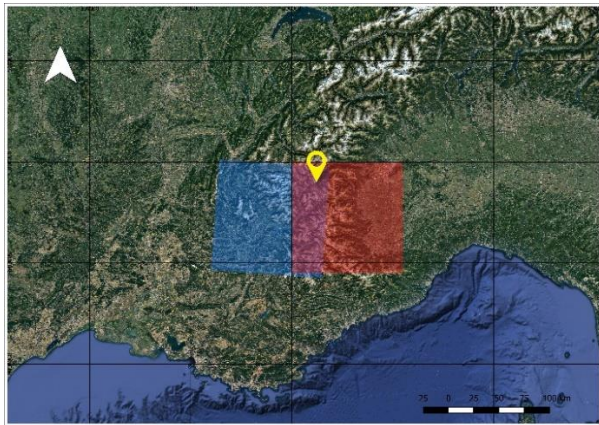

Figure 1. Study area. The Blue square covers the K tile, while the red one covers the Q tile. The yellow pointer indicates Cesana Torinese (TO).

2.2 Google Earth Engine

The main steps of a semi-authomatic classification workflow (topographic correction, extraction of training points, training of the classifier, classification) were realized on Google Earth Engine (GEE). GEE is a web-based platform for geospatial analysis lunched in 2010, and it is free for research and education purposes (Gorelick et al., 2017; Kumar and Mutanga, 2018). The GEE data catalog is composed of continuously updated geospatial datasets provided by different national and international programs, such as NASA and ESA (Hu et al., 2018; Kumar and Mutanga, 2018; Shelestov et al., 2017; Sidhu et al., 2018). The datasets included some already elaborated satellite data. The true innovation of GEE is the possibility of interacting with very large datasets and computing basic geospatial analysis directly on GEE servers through JavaScript/Python-based API (GEE API) (Goldblatt et al., 2017; Google Earth Engine, 2020). GEE makes available the imagery captured by Sentinel 2. Sentinel 2 is a mission of the European Space Agency (ESA) and it is composed of twins satellites (Sentinel-2A and Sentinel-2B) that carry multispectral optical sensors. Sentinel 2 products are images of 13 spectral bands that are published with two possible levels of processing (ESA, 2020). The dataset for the classification is composed of 31 scenes (epochs) filtered with GEE from Sentinel 2 level-1C representing the Top-Of-Atmosphere (TOA) reflectance scaled by 10000 (it includes radiometric and geometric corrections including ortho-rectification and spatial registration on a global reference system with sub-pixel accuracy (ESA, 2020)).

2.3 Training and Validation Datasets

The 1260 points that compose the training dataset were randomly selected. This process was sped up by randomly sampling (stratify samples) a training layer generated from already existing land cover classifications. This phase, besides accelerating the training points positioning, also ensured the minimization of the subjectivity in the identification of the training points (Gromny et al., 2019). The information of ESA High-resolution Layers (VHR) 10m (Forest cover, Imperviousness, Grassland and Water

and Wetness, http://land.copernicus.eu/) (Lefebvre et al., 2016), CORINE Land Cover (CLC) (Copernicus, 2020), and RUSLE 2015 at 100 m from ESDAC (Panagos et al., 2015) constitute the training layer. The classed of CORINE land cover, as well as other input data, were reassigned based on the classes of interest, as Table 1 shows. Finally, the training points were visually checked. This phase was carried out outside GEE environment. The accuracy assessment is based on 1260 points identified using the same methods for the definition of the training points.

Reference classes	VHR mosaics 2015	ESDAC 2015	CLC
Coniferous forest	Forest Type	-	312
Broadleaves forest			311
Grasslands	Grassland	-	321-231
Water	Water and Wetness	-	511,512
Clustered bare soil areas	-	RUSLE (> 20 t/he/yr)	323-331
Solid-rocky areas	-		332
Urban areas	Imperviousness Density	-	111-133
Glaciers	-		335
Agricultural lands	-		211-244

Table 1. Input datasets for the training layer and the translation to the reference classes.

2.4 Satellite images dataset

The classification was realized on Q tile and then its replicability checked on tile K. The images regarding the entire Sentinel-2 activity of sensing were filtered according to two parameters: the cloud cover percentage, data must have less than 10% the scene covered by clouds; and by the sensing period. Only images sensed during summertime (from June to August) were selected to minimize the effect of the shadows. The filter reduced considerably the data available from Sentinel 2 level 2A (Sentinel 2 highest level of processing product that geometrically and atmospherically corrected): only 16 images that satisfied the filter criteria were available. Therefore, to ensure a larger classification dataset, 1C level data were used. (Table 2).

Tile	year	No.	Sentinel Image identification code
K	2017	0	20170704T103019_20170704T103637_T31TGK
		1	20170714T103019_20170714T103022_T31TGK
		2	20170818T103021_20170818T103421_T31TGK
		3	20170823T103019_20170823T103018_T31TGK
	2018	4	20180619T103019_20180619T103559_T31TGK
		5	20180719T103019_20180719T103020_T31TGK
		6	20180719T103019_20180719T103820_T31TGK
		7	20180729T103019_20180729T103815_T31TGK
		8	20180828T103019_20180828T103013_T31TGK
	2019	9	20190604T103029_20190604T103616_T31TGK
		10	20190629T103031_20190629T103537_T31TGK
		11	20190704T103029_20190704T103317_T31TGK
		12	20190714T103029_20190714T103635_T31TGK
		13	20190719T103031_20190719T103715_T31TGK
		14	20190729T103031_20190729T103230_T31TGK
		15	20190803T103029_20190803T103728_T31TGK
		16	20190808T103031_20190808T103026_T31TGK
		17	20190818T103031_20190818T103539_T31TGK
Q	2017	0	20170704T103019_20170704T103637_T32TLQ
		1	20170714T103019_20170714T103022_T32TLQ
		2	20170818T103021_20170818T103421_T32TLQ
		3	20170823T103019_20170823T103018_T32TLQ
	2018	4	20180619T103019_20180619T103559_T32TLQ
		5	20180704T103021_20180704T103023_T32TLQ
		6	20180719T103019_20180719T103020_T32TLQ
		7	20180828T103019_20180828T103013_T32TLQ
	2019	8	20190604T103029_20190604T103616_T32TLQ
		9	20190619T103031_20190619T103536_T32TLQ
		10	20190704T103029_20190704T103317_T32TLQ
		11	20190729T103031_20190729T103230_T32TLQ
		12	20190803T103029_20190803T103728_T32TLQ

Table 2. Sentinel 2 level 1C images of the classification dataset.

31 tiles constitute the final dataset: 13 images for Q and 18 for K. Sentinel 1C level products are georeferenced and geometrically corrected, but the atmospheric disturbance and the topographic-derived distortions are not rectified. Thus, the imagery was pre-processed before the classification.

3. METHODS

In this paragraph is described the methodology applied for this classification. It consists of a machine learning classification based on random forest algorithm, realized using 1260 training points on 9 classes: Coniferous forest, Broadleaves forest, Grasslands, Water, Clustered bare soil areas, Solid-rocky areas, Urban areas, Glaciers, and Agricultural lands (Table 3).

No.	Class	Description
1	Coniferous forest	Coniferous trees areas
2	Broadleaves forest	Broadleaves trees forests
3	Grasslands	Includes pastures and moors
4	Water	Internal waters
5	Clustered bare soil areas	Clustered rocks and bare soil, generally in high slope mountainsides. Highly prone to erosion.
6	Solid-rocky areas	Continuous non-clustered rocky cover. Typical of Alpine mountain peaks.
7	Urban areas	Buildings, roads, artificial infrastructures.
8	Glaciers	Perennial snow/ice cover. In Alpine zone, only in areas above 3000 m a.s.l.
9	Agricultural lands	Areas interested by agricultural activities that require tillage, or present rows of fruit trees/bushes.

Table 3. Classes of Land Cover considered in this work.

3.1 Pre-processing

Sentinel 2 1C-level products are georeferenced, but the atmospheric disturbance and the topographic-derived distortions are not corrected. Thus, the imagery was pre-processed phase preceded the classification and consists of the cloud masking, the atmospheric correction, and the topographical correction.

3.1.1 Cloud masking
The dense clouds and the cirrus were masked in each epoch using the QA60 band. The QA60 band is provided as an additional 1C-level product and represents the cloud mask computed by ESA through the analysis of the blue spectral region (B2), the SWIR reflectance in B11 and B12 (ESA, 2020). Only cloud-free pixels were maintained, while dense clouds and cirrus were classified as *no data* values.

3.1.2 Atmospheric correction
The atmospheric correction of satellite imagery is considered fundamental in remote sensing applications, especially in the case of multi-temporal analysis (Hadjimitsis et al., 2004; Lantzanakis et al., 2017; Martins et al., 2017; Sola et al., 2018). The atmospheric correction removes the scattering effect of the Earth's atmosphere and it can be based on Radiative Transfer Models (Specific mathematical models that consider latitude,

season and atmospheric conditions) or on Image-Based Correction Techniques (that estimate atmosphere scattering using information and data within the image) (Hadjimitsis et al., 2004; Lantzanakis et al., 2017; Martins et al., 2017). Usually, the Radiative Transfer Models are more accurate and therefore more applied. For example, ESA for the production of atmospherically corrected imagery of Sentinel 2 (level-2A), applies the Sen2Cor Radiative Transfer model. For the time being, there is no atmospheric correction model to be applied to Sentinel 2-level 1C implemented in GEE. This means that Sentinel 2 at 1C processing level cannot be easily atmospherically corrected using Radiative Transfer Models in GEE environments. This is a severe limit to the performing of multi-temporal analysis. To overcome this limitation, in this work, we applied a linear model for the atmospheric disturbance reduction: Dark Object Subtraction (DOS) (Chavez, 1988), which performs similarly to radiative transfer models on homogeneous surfaces such as grass, water, and bare soil (Lantzanakis et al., 2017). The application of the DOS did not negatively affect the final results because the method we propose classify separately each epoch of the dataset, as discussed in paragraph 3.3.

3.1.3 Topographic correction
The topographic correction allows the variation in the reflectance derived by the inclination of the terrain and the sun elevation (Poortinga et al., 2019; Shepherd and Dymond, 2010). This pre-processing phase is crucial in mountainous areas, because of the steep mountainsides and the consequent alteration of reflectance values. The entire dataset was corrected by applying a semi-empirical correction. The code was originally implemented in GEE by Patrick Burns and Matt Macander, and then adapted to the dataset. The correction is based on a semi-empirical method that takes into consideration not only the topography of the area (as for the empirical methods), but also the solar angle (both zenith and azimuth) (Shepherd and Dymond, 2010). The topographic correction is based on sun-canopy-sensor with a semi-empirical moderator (c) (SCSc) method (Poortinga et al., 2019; Soenen et al., 2005). A Digital Elevation Models (DEM) to detect the slopes of the area and the solar position information (i.e. the sun inclination and sun irradiance) are the input data. These data are available from the metadata of the satellite images, while elevation information was extracted from the Shuttle Radar Topography Mission (SRTM) digital elevation data with 30m spatial resolution that is available in the GEE catalog (Farr et al., 2007).

3.2 Derivative features

Some additional bands were calculated to improve the final accuracy of the classification. The diversification of the input information is crucial for a good classification. For example, textural elements can facilitate Land cover classes discrimination (Lewiński et al., 2015), as well as the histogram-based ones (Drzewiecki et al., 2013). First, 10 radiometric features were added (Table 4). Secondly, 5 histogram-based features, 19 textural and 1 edge-detector features were computed. Particularly, the texture metrics from the Gray Level Co-occurrence Matrix in the 7x7 neighborhood of each pixel of band 8 (Near InfraRed, NIR) were computed (Table 4) (Conners et al., 1984; Haralick et al., 1973; GEE, 2020). The derivate features improved the goodness of the classification, but required high computational power. Indeed GEE exceeded the memory limit by running the entire code: the filtering of the Sentinel-2 data catalog, the topographic and atmospheric corrections of the filtered features, the derivative bands computation for each epoch and the classification itself. Thus, to slim out the classification process and to reduce the computational time of the classification a correlation analysis and the importance of the layer in the classification analysis were realized.

	Feature	Formula/note
Radiometric	Alteration	B11/B12
	Chlorophyll IndexRedEdge CRE	(B9/B5)−1
	Enhanced Vegetation Index EVI	2.5*((B9−B5)/((B9+6*B5−7.5*B1)+1))
	HUE	Arctan((2*V5−B3−B1)/30.5)*(B3−B1))
	Normalized Difference Vegetation Index NDVI	(B8−B4)/(B8+B4)
	Soil-Adjusted Vegetation Index SAVI	((B8−B4)/(B8+B4+L))*(1+0.5)
	Soil Composition Index SCI	(B11−B8)/(B11+B8)
	Specific Leaf Area Vegetation Index SLAI	B9/(B5+B12)
	Wetnss Index WET	(0.1509*B2)+(0.1973*B3)+(0.3279*B4)+(0.03406*B8)-(0.7112*B11)-(0.4572*B12)
	Triangular Vegetation Index TVI	0.5*(120*(B8-B3))-(200*(B4-B3))
Edge	Sob	Sobel edge extractor
Histogram-based	Var	Variance
	Mean	Mean
	Skew	Skewness
	Kurt	Kurtosis
	Contr	Contrast; measures the local contrast of an image
Textural GLCM	Entr	Entropy
	Asm	Angular Second Moment; measures the number of repeated pairs
	Corr	Correlation; measures the correlation between pairs of pixels
	Var	Variance; measures how spread out the distribution of gray-levels is
	Idm	Inverse Difference Moment; measures the homogeneity
	Savg	Sum Average
	Svar	Sum Variance
	Sent	Sum Entropy
	Ent	Entropy. Measures the randomness of a gray-level distribution
	Dvar	Difference variance
	Dent	Difference entropy
	Imcorr1	Information Measure of Corr. 1
	Imcorr2	Information Measure of Corr. 2
	Maxcorr	Max Corr. Coefficient.
	Diss	Dissimilarity
	Inertia	Inertia
	Shade	Cluster Shade
	Prom	Cluster prominence

Table 4. Derivative features calculated for each epoch. They are divided into 4 groups: histogram-based, radiometric, edge extractor and textural.

3.2.1 Correlation analysis

The slimming out workflow was designed to reduce the computational time without losing accuracy. First, a correlation analysis was performed between the radiometric derivative features and the original bands to avoid information redundancy. The correlation between bands was analyzed using July 2017 as reference. The analysis was based on the DN values randomly sampled within the classes. Band 4 was used as reference band. Correlation coefficients equal to 1 indicate total correlation. This analysis considered highly-correlated variables greater or equal to 0.85 (dark green in Table 5). Then, the correlation analysis between textural bands was carried out. Since the textural bands resulted to be much less correlated to each other compared to the radiometric bands, the importance of textural predictors in the classification was computed by running the classification each time excluding one different predictor. The Overall accuracies of these classifications were considered as the importance value of the removed predictor. This indirect strategy to estimate the importance of each feature within the classification was necessary because the evaluation of predictors' importance is not implemented in GEE. The importance analysis was performed on July 2017 tile. The OA of the classification computed with all textural features is 0.825. The features which exclusion caused the increasing on the OA of 0.01 points were considered as less important (negatively affecting the classification). To further reduce the computational effort, the bands were normalized and then transformed in integer values (int16). Since running the normalization on GEE required too much memory was realized a "pseudo-normalization". For each band was identified the multiplicative factor that allows to obtain integer values no bigger than 32767 (max values for signed integer data format). The pseudo-normalization does not affect the classification.

3.3 Classification

Each image was separately classified using the machine learning algorithm random forest with 50 rifle decision trees per class and 4 as the minimum size for terminal nodes. The same training dataset was used for each image. This means that for one image were sampled 37 DN values in correspondence of every training point. This datum was used to train the classifier and finally to apply it to the starting image. The results are 13 classifications that were aggregated according to the most frequent pixel value between 1 and 9 (*no data* values excluded) to obtain the final classification of the area. The aggregation allowed us to minimize the mistakes of the single classifications and to take out from the final result the *no data* values of the cloud masking (Nowakowski et al., 2017).

3.4 Accuracy assessment

The accuracy assessment is based on 1260 defined by the same method adopted for the identification of the training dataset. It consisted of the computation of the error matrix and the derived accuracy measures for each single classification and the final aggregated one. The error matrix-derived measures are the overall accuracy, the producer's accuracy, the user's accuracy, and the F1 score.

4. RESULTS

4.1 Correlation and importance Analysis

Table 5 shows the results from the correlation analysis between radiometric-based derivative features. The correlation coefficient ranges between 0 (no correlation, light green) and 1 (total correlation, dark green), (Table 5, Table 6). The radiometric derivative features with a correlation coefficient larger than 0.85

were excluded from the classification dataset: SLAI, NDVI, Alteration, and SAVI. The correlation between the original bands was not taken into consideration. Nevertheless, Sentinel bands 1, 2, 9, 10 were excluded from the analysis as too influenced by the atmospheric component. The final bands considered were: B4, B2, B3, B5, B6, B7, B8, B8A, B11, B12, CRE, EVI, SCI, HUE, WET, TVI. The correlation between the textural features is not particularly remarkable (Table 6): only GLCM Contrast feature appears highly correlated to Dissimilarity, Variance, and Shade.

Table 5. Correlation analysis of radiometric-based features. The dark green cells indicate a correlation coefficient >0.85.

Table 6. Correlation analysis of textural and edge-detector features. The dark green cells indicate a coefficient >0.85.

The importance of textural features in the classification was computed and Table 7 reports the results: only the Shade and Shade and Prominence negatively affect the classification, as they removal increased the overall accuracy of 0.01. They were excluded from the classification dataset. The correlation and the importance of the predictors analysis on the bands resulted in 7 highly-correlated and low-importance bands that were removed from the dataset (2 textural-based). The final results are 37 bands in each one of the 13 images. The following bands used: 'B3', 'B4', 'B5', 'B6','B7', 'B8', 'B8A', 'B11', 'B12', 'CRE', 'EVI', 'SCI', 'HUE', 'WET', 'TVI', 'entr', 'sob', 'var', 'mean', 'skew', 'kurt', 'asm',

'contrast', 'corr', 'idm', 'maxcorr', 'var', 'svar', 'dent', 'imcorr2', 'savg', 'imcorr1', 'diss', 'sent', 'ent', 'dvar', 'inertia'.

Excluded features	OA value
All features	0.825
entr	0.829
sob	0.826
asm	0.824
contrast	0.819
corr	0.822
var	0.822
idm	0.823
savg	0.822
svar	0.824
sent	0.827
ent	0.825
dvar	0.826
dent	0.822
imcorr1	0.824
imcorr2	0.829
maxcorr	0.825
diss	0.828
inertia	0.825
shade	**0.834**
prom	**0.835**

Table 7. Importance analysis of the textural features. The features which exclusion cause increasing on the OA of 0.01 points were considered as less important (negatively affecting the classification).

4.2 Classification

It was not possible to perform the entire classification in one GEE script. Thus, it was carried out using some expedients. First were performed the filtering and the correction. The pre-processed dataset was then exported in the GEE personal Asset. A new script was written for the computation of the classifications in which was imported the pre-processed dataset. The slimming out tests were realized in a separate script. Finally, the single classifications were converted into int8 data format, stacked in a single image and exported (Figure 2). Figure 3 provides some examples of the classification results in high-altitude areas (left) and lowlands (right).

4.3 Accuracy

The overall accuracy for single date classification varied from 0.661 to 0.747 in the case of only radiometric derivative features and from 0.791 to 0.900 for the radiometric and textural derivative features. Table 8 reports the results of the accuracy assessment deriving from the classification of i) the Sentinel radiometric bands and the radiometric-based features, and ii) the Sentinel radiometric bands, the radiometric-based features, and the textural features. The F1 scores of clustered bare soil areas along with the OA accuracy values rapidly increase. The classification with the derivative features shows the overall accuracy of 86%. The OA of 94% deriving from the dataset with radiometric and textural features is the result of the aggregation of 13 images.

	Conifers	Broadleaves	Grassland	Water	Clustered bare soils	Solid rocky	Urban areas	Agricultural lands	Overall Accuracy
Radiometric bands and radiometric-based features									
PA	0.986	0.979	0.965	1.000	0.636	0.520	0.949	0.938	
UA	0.986	0.979	0.993	0.971	0.250	0.829	0.936	0.964	0.863
F1	0.986	0.979	0.979	0.986	0.359	0.639	0.942	0.951	
Radiometric bands, radiometric-based features and textural features									
PA	1.000	0.986	0.898	1.000	0.812	0.915	0.993	0.995	
UA	0.993	1.000	0.993	0.979	0.957	0.771	0.950	0.946	0.945
F1	0.996	0.993	0.943	0.989	0.879	0.837	0.971	0.970	

Table 8. Producer's and User's accuracies and F1 score for the classification with and without textural features.
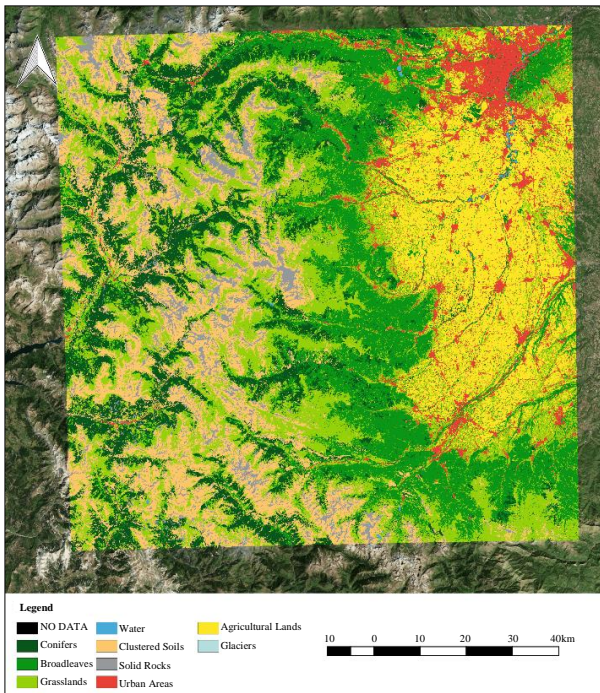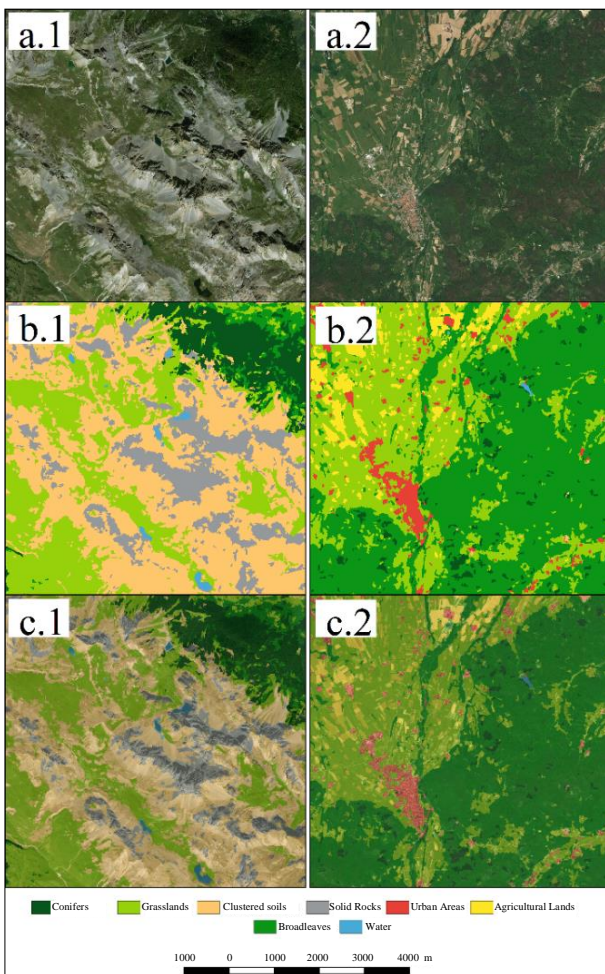
Figure 2. Classification of tile Q.

Figure 3. Some detail of the classification on tile in high-altitude mountainous areas (left, scene 1) and lowlands (right, scene 2). Images a.1 and a.2 show the RGB orthophoto (Source: Bing Satellite); b.1 and b.2 shows the classification; c.1 and c.2 shows the overlapping of the orthophoto and the classification (low opacity).

Figure 4 reports the accuracy and the F1 score for each single classification. The accuracy values have very unstable results within the classification scene. This is partially due to the incorrect classification of the areas covered by clouds (*no data*) and the pixels in the immediate surroundings of the masked area that may suffer from radiometric alteration for cloud proximity, but not detected by Sentinel cloud masking.
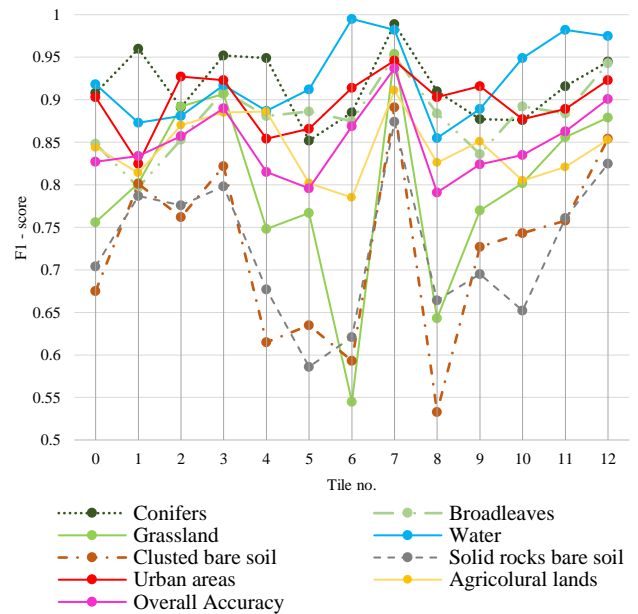
Figure 4. The changes in F1-score for each class through the epochs (tile no., on the abscissae).

4.4 Replicability

The replicability of the classification methodology was tested on the dataset of tile K and the overlapping area of tiles Q and K. The accuracy assessment from tile K Table 9 shows very similar results to the one of tile Q. It appears that on both tiles the classes that describe the mountainous areas have lower F1 value. The accuracy assessment based on 700 test points (placed *ad hoc*) was realized for the overlapping area, to check the stability of the classifications of tiles K and Q in most challenging areas (mountains peaks). Only 7 classes are present in the scene, but the rocky areas classes dominate the scene (Table 10).

	Conifers	Broadleaves	Grassland	Water	Clustered bare soils	Solid rocky areas	Urban areas	Agricultural lands	Glaciers	Overall Acc.
PA	0.993	0.943	0.927	0.914	0.824	0.727	0.899	0.905	1.000	
UA	0.979	0.950	0.907	0.993	0.771	0.836	0.829	0.886	0.974	0.907
F1	0.986	0.947	0.917	0.952	0.797	0.777	0.862	0.895	0.987	

Table 9. Producer's and User's accuracies and F1 score for the classification on tile K.

	Conifers	Broadleaves	Grassland	Water	Clustered bare soils	Solid rocky areas	Urban areas	Overall Acc.
PA	0.990	1.000	0.990	1.000	0.855	0.750	0.961	
UA	1.000	0.980	1.000	0.970	0.949	0.870	0.730	0.928
F1	0.995	0.990	0.995	0.985	0.900	0.806	0.830	

Table 10. Producer's and User's accuracies and F1 score for the classification on the overlapping area between tiles Q and K.

5. DISCUSSION

The filter for selecting only the summertime images with a very low percentage of cloud cover, on one hand, ensured uniform illumination and atmospheric conditions; but on the other hand, impeded the use of Sentinel 2 highest processing level (2A). Besides, the classification of the 1C-level dataset achieved interesting results. The independent classification of each image and the frequency-based aggregation minimized the distortions and the errors derived from the topography and the atmosphere. Moreover, the summertime filter reduced the variability in the meteorological and atmospheric conditions allowing the application of a linear model of correction. Even if the pre-processing phase is consistent, it had a low impact of the methodology thanks to GEE cloud computing and the ease with which it is replicable in GEE. Although, we faced some major constraints in GEE related to the memory available for single users, the poor information regarding the available functions and the lack of some useful functions (like layers importance). Nevertheless, it is a relatively young service that is constantly enriched with new functions and features and our analysis was quite ambitious from the computational point of view.

Considering the results in detail, the derivative features considerably increased the goodness of the classification. It is interesting to note the role of textural bands in the discrimination of confusion between clustered soil class and solid-rocks class. Indeed, F1 scores of clustered bare soils jumped from 0.359 to 0.879 by adding the textural features to the classification (Table 8). Figure 4 clearly shows that the clustered bare soil and the solid-rocky areas are the classes with a lower F1 score in every single classification as evidence of the difficulty of separation of the two classes. Generally, the accuracy of a single classification shows a similar trend throughout the classes, such as image 6 (low F1 values) and classification no. 7 (high F1 values for all its classes) from Figure 4. Regarding the aggregated classification, the overall accuracy achieves 0.945 with a clear improvement also in clustered-rocky and solid-rocky areas (respectively F1-scores 0.827 and 0.890) proving the validity of the aggregation method to reduce the main errors from the single classifications (Table 8). The replicability analysis on tile K reported in Table 9, show trends in clustered and solid rock classes close to the ones of tile Q. The overall accuracy of K is 0.907. The difference between K and Q overall accuracy can be caused by the training dataset and it is negligible. It is worth mentioning that in tile K there is one class more: glaciers. Indeed, the land cover by glaciers is relatively small and present only in K tile (in the Écrin National Park, FR). Even if the glaciers classification is very good some small glaciers were not detected. The overlapping area between K and Q provides promising results if we consider the fact that the most confuse classes clustered and rocky areas are the dominant LC (Table 10). On the other hand, the few classes present in the area provide better values of Overall accuracy. Generally, the solid-rocks are frequently misclassified as urban areas. This is attributed to the high spectral similarity of the classes and their similar textural characterization. A frequent error in Alpine areas is the classification of buildings with rocky roofs as solid-rocky areas, the same for the small rivers in which the water flow during the summertime is reduced. Also, some pixels classified as water are detected in mountainous areas, probably due to the presence of shaded rocky/snowed areas.

6. CONCLUSION

This study proposes a simple method for the classification of Land cover in mountainous areas in GEE platform, which includes some steps to reduce the computational effort and the time for the classification. The methodology minimizes the error introduced by the atmospheric component and the terrain

inclination using only images captured during a short time range in limited cloud cover conditions and by applying atmospheric and topographic corrections. The textural derivative features played a key role in distinguishing the most challenging classes (clustered bare soil and solid-rocky areas). An additional positive value of the methodology is the aggregation method, indeed, by considering the modal value of the single classifications, the final accuracy significantly raised. All the aspects allowed us to reach good accuracies in mountain areas. Since the entire classification is performed in GEE, it can be easily modified and updated.

REFERENCES

Carrasco, L., O'Neil, A.W., Morton, R.D., Rowland, C.S., 2019. Evaluating Combinations of Temporally Aggregated Sentinel-1, Sentinel-2 and Landsat 8 for Land Cover Mapping with Google Earth Engine. Remote Sens. 11, 288. https://doi.org/10.3390/rs11030288

Chavez, P.S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sens. Environ. 24, 459–479. https://doi.org/10.1016/0034-4257(88)90019-3

Conners, R.W., Trivedi, M.M., Harlow, C.A., 1984. Segmentation of a high-resolution urban scene using texture operators. Comput. Vis. Graph. Image Process. 25, 273–310. https://doi.org/10.1016/0734-189X(84)90197-X

Copernicus Land Monitoring Service [WWW Document]. URL https://land.copernicus.eu/user-corner/publications/clc-flyer (accessed 3.16.20).

Delalay, M., Tiwari, V., Ziegler, A.D., Gopal, V., Passy, P., 2019. Land-use and land-cover classification using Sentinel-2 data and machine-learning algorithms: operational method and its implementation for a mountainous area of Nepal. J. Appl. RemoteSens.13,014530. https://doi.org/10.1117/1.JRS.13.014530

Dorren, L.K.A., Maier, B., Seijmonsbergen, A.C., 2003. Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. For. Ecol. Manag. 183, 31–46. https://doi.org/10.1016/S0378-1127(03)00113-0

Drzewiecki, W., Wawrzaszek, A., Aleksandrowicz, S., Krupiński, M., Bernat, K., 2013. Comparison of selected textural features as global content-based descriptors of VHR satellite image, IGARSS. pp. 4364–4366. https://doi.org/10.1109/IGARSS.2013.6723801

ESA, User Guides - Sentinel-2 MSI - Sentinel Online [WWW Document], URL https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/processing-levels (accessed 3.16.20).

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., et Al. 2007. The Shuttle Radar Topography Mission. Rev. Geophys. 45. https://doi.org/10.1029/2005RG000183

Feng, Q., Liu, J., Gong, J., 2015. UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis. Remote Sens. 7, 1074–1094. https://doi.org/10.3390/rs70101074

GEE– Google Earth Engine Platform [WWW Document]. URL https://earthengine.google.com/platform/ (accessed 3.9.20).

Goldblatt, R., Rivera Ballesteros, A., Burney, J., 2017. High Spatial Resolution Visual Band Imagery Outperforms Medium Resolution Spectral Imagery for Ecosystem Assessment in the Semi-Arid Brazilian Sertão. Remote Sens. 9, 1336. https://doi.org/10.3390/rs9121336

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ., Big Remotely Sensed Data: tools, applications and experiences 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Gromny, E., Lewiński, S., Rybicki, M., Malinowski, R., Krupiński, M., Nowakowsk, A., Jenerowicz, M., 2019. Creation of training dataset for Sentinel-2 land cover classification, SPIE, Wilga, Poland, p. 106. https://doi.org/10.1117/12.2536773

Hadjimitsis, D.G., Clayton, C.R.I., Hope, V.S., 2004. An assessment of the effectiveness of atmospheric correction algorithms through the remote sensing of some reservoirs. Int. J. Remote Sens. 25, 3651–3674. https://doi.org/10.1080/01431160310001647993

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. IEEE Trans. Syst. Man Cybern. SMC-3, 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Houze, R.A., 2012. Orographic effects on precipitating clouds. Rev. Geophys. 50. https://doi.org/10.1029/2011RG000365

Hu, Y., Dong, Y., Batunacun, 2018. An automatic approach for land-change detection and land updates based on integrated NDVI timing analysis and the CVAPS method with GEE support. ISPRS J. Photogramm. Remote Sens. 146, 347–359. https://doi.org/10.1016/j.isprsjprs.2018.10.008

Itten, K.I., Meyer, P., 1993. Geometric and radiometric correction of TM data of mountainous forested areas. IEEE Trans. Geosci. Remote Sens. 31, 764–770. https://doi.org/10.1109/36.239898

Kumar, L., Mutanga, O., 2018. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. Remote Sens. 10, 1509. https://doi.org/10.3390/rs10101509

Lantzanakis, G., Mitraka, Z., Chrysoulakis, N., 2017. Comparison of Physically and Image Based Atmospheric Correction Methods for Sentinel-2 Satellite Imagery, Perspectives on Atmospheric Sciences, Springer Atmospheric Sciences. Springer International Publishing, pp. 255–261.

Lefebvre, A., Sannier, C., Corpetti, T., 2016. Monitoring Urban Areas with Sentinel-2A Data: Application to the Update of the Copernicus High Resolution Layer Imperviousness Degree. Remote Sens. 8, 606. https://doi.org/10.3390/rs8070606

Lewiński, S., Aleksandrowicz, S., Banaszkiewicz, M., 2015. Testing Texture of VHR Panchromatic Data as a Feature of Land Cover Classification. Acta Geophys. 63, 547–567. https://doi.org/10.2478/s11600-014-0250-5

Martins, V.S., Barbosa, C.C.F., De Carvalho, L.A.S., Jorge, D.S.F., Lobo, F.D.L., Novo, E.M.L. de M., 2017. Assessment of Atmospheric Correction Methods for Sentinel-2 MSI Images Applied to Amazon Floodplain Lakes. Remote Sens. 9, 322. https://doi.org/10.3390/rs9040322

Nowakowski, A., Rybicki, M., Kukawska, E., Malinowski, R., Krupiński, M., Lewinski, S., 2017. Aggregation of Sentinel-2 time series classifications as a solution for multitemporal analysis. SPIE, Warsaw, Poland, p. 11. https://doi.org/10.1117/12.2277976

Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusburger, K., Montanarella, L., Alewell, C., 2015. The new assessment of soil loss by water erosion in Europe. Environ. Sci. Policy 54, 438–447.https://doi.org/10.1016/j.envsci.2015.08.012

Poortinga, A., Tenneson, K., Shapiro, A., Nquyen, Q., San Aung, K., Chishtie, F., Saah, D., 2019. Mapping Plantations in Myanmar by Fusing Landsat-8, Sentinel-2 and Sentinel-1 Data along with Systematic Error Quantification. Remote Sens. 11, 831. https://doi.org/10.3390/rs11070831

Rizeei, H.M., Saharkhiz, M.A., Pradhan, B., Ahmad, N., 2016. Soil erosion prediction based on land cover dynamics at the Semenyih watershed in Malaysia using LTM and USLE models. Geocarto Int. 31, 1158–1177. https://doi.org/10.1080/10106049.2015.1120354

Sekertekin, A., Marangoz, A.M., Akcin, H., 2017. Pixel-based classification analysis of land use land cover using sentinel-2 and landsat-8 data. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII-4/W6, 91–93. https://doi.org/10.5194/isprs-archives-XLII-4-W6-91-2017

Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., Skakun, S., 2017. Exploring GEE Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping. Front. Earth Sci. 5. https://doi.org/10.3389/feart.2017.00017

Shepherd, J.D., Dymond, J.R., 2010. Correcting satellite imagery for the variance of reflectance and illumination with topography. Int. J. Remote Sens.https://doi.org/10.1080/014311602210154029

Sidhu, N., Pebesma, E., Câmara, G., 2018. Using Google Earth Engine to detect land cover change: Singapore as a use case. Eur. J. Remote Sens. 51, 486–500. https://doi.org/10.1080/22797254.2018.1451782

Soenen, S.A., Peddle, D.R., Coburn, C.A., 2005. SCS+C: a modified Sun-canopy-sensor topographic correction in forested terrain. IEEE Trans. Geosci. Remote Sens. 43, 2148–2159. https://doi.org/10.1109/TGRS.2005.852480

Sola, I., García-Martín, A., Sandonís-Pozo, L., Álvarez-Mozos, J., Pérez-Cabello, F., González-Audícana, M., Montorio Llovería, R., 2018. Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes. Int. J. Appl. Earth Obs. Geoinformation 73, 63–76. https://doi.org/10.1016/j.jag.2018.05.020

Thanh Noi, P., Kappas, M., 2018. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. Sensors 18, 18. https://doi.org/10.3390/s18010018

Turner, M.G., Gardner, R.H., 2015. Landscape Ecology in Theory and Practice: Pattern and Process, 2nd ed. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4939-2794-4