GAVIN PUBLISHERS

OPEN ACCESS

Current Trends in Forest Research

Pettinato S, et al. Curr Trends Forest Res: CTFR-117.

Short Commentary

DOI: 10.29011/ 2638-0013. 100017

Forest Classification in Mediterranean Areas Using a Quadratic Bayesian Classifier Method Applied to SAR Images

Simone Pettinato[1*], Emanuele Santi[1], Simonetta Paloscia[1], Giacomo Fontanelli[2], Andrea Garzelli[3]

[1]Institute of Applied Physics-National Research Council of Italy (IFAC-CNR), Via Madonna del Piano, Florence, Italy

[2]Rothamsted Institute, London, UK

[3]Department of Information Engineering and Mathematical Sciences, University of Siena, Siena, Italy

[*]**Corresponding author:** Simone Pettinato, Institute of Applied Physics-National Research Council of Italy (IFAC-CNR), Via Madonna del Piano 10, 50019 Florence, Italy. Tel: +390555226463; Email: s.pettinato@ifac.cnr.it

Citation: Pettinato S, Santi E, Paloscia S, Fontanelli G, Garzelli A (2018) Forest Classification in Mediterranean Areas Using a Quadratic Bayesian Classifier Method Applied to SAR Images. Curr Trends Forest Res: CTFR-117. DOI: 10.29011/ 2638-0013.100017

Received Date: 24 July, 2018; **Accepted Date:** 09 August, 2018; **Published Date:** 17 August, 2018

Abstract

In this paper, multi frequency SAR images from ALOS/PALSAR, ENVISAT/ASAR, and Cosmo-SkyMed sensors have been studied for highlighting forest features in a test area in Central Italy (San Rossore), where detailed in-situ measurements were available. A preliminary discrimination of the main land cover classes and forest types was carried out by exploiting the synergy among various frequencies (L, C and X bands). After a pre-processing of the available SAR images by applying a multilook approach, SAR data were used for discriminating forest from non-forest land covers and separating broadleaved from coniferous forest types by using RGB compositions of multi-temporal and multi-frequency images. The mean backscattering coefficient (σ°) was computed for each sensor and available polarization from the pixels associated to coniferous and broadleaf obtained from the reference classification map. The classification has then been performed by applying to the SAR images, in different configurations of polarizations and frequencies, a new method based on a quadratic Bayesian classifier, which is able to overcome the limits of ground-truth classes that contain not homogenous targets (i.e. non-forest class). The obtained results indicated that the different surface types were best identified by the joint use of X and L bands (the correct classification show 80.13%, 83.03% and 75.07% for coniferous, broadleaf forests and non-forest respectively). The best overall accuracy is also obtained by considering the joint use of L and X bands (80.06%).

Keywords: Bayesian Classifier; Forest Features; Forest/Non-Forest Areas; Land Classification; SAR

Introduction

Forest monitoring is commonly recognized as a vital task due to the role played by forests in carbon cycle evolution, which act as the main terrestrial carbon sinks [1]. Since maps of forest status and their temporal evolution are increasingly required, the combined use of in-situ and remote sensing data is desirable, especially in regions with scarce accessibility and limited ground data availability [2]. Optical sensors have been for a long time widely used, although these sensors have the limitation of operating in clear-sky conditions and are sensitive to the upper layer of the canopy only (e.g., [3,4]). This makes the investigation of equatorial, boreal, and mountain areas rather difficult, due to the frequent and consistent cloud cover over these regions.

Microwave frequencies have the advantage to be independent of cloud cover and solar radiation and can significantly penetrate into vegetation cover, and both emission and backscatter are considerably influenced by moisture content and by the geometrical features of plant constituents, according to frequency and polarization. A very suitable sensor for forest investigations is the Synthetic Aperture Radar (SAR), which is carried onboard historical satellites, such as ALOS1, RADARSAT1, and ENVISAT (the latter active until April 2012), and recently launched satellites (e.g. Cosmo-SkyMed, TerraSARX, ALOS2, RS2, Sentinel-1), the latter being able to obtain images at high spatial resolution and frequent revisit time. The total backscatter from a forest is a

combination of contributions from ground, vegetation components and their interactions. Each of these components affects the total scattering depending on the microwave wavelength: at wavelengths shorter than 10 cm, i.e. C (6 cm) and X (3 cm) bands, the backscatter is mainly due to leaves or needles and twigs of the upper crowns, thus making it possible to obtain information on the first layer of vegetation cover, and therefore allowing the identification of forest/non-forest areas, forest thermal state (frozen/thawed), and the estimate of some forest parameters as well, e.g. tree density, tree height and above ground biomass [5-12]. At longer wavelengths, i.e. L (21 cm) and P bands (60 cm), due to the higher penetration power, information on thicker layers of vegetation and of soil under vegetation cover can be retrieved, due to the interactions with major branches, trunks and ground [13-16]. Experimental and theoretical investigations carried out for many years mainly focused on the boreal forests of North America and Eurasia, in view of their influence on climate changes [17,18], and on tropical regions due to the frequent cloud cover that hampers the use of optical sensor [14,19,20]. Interesting recent works have been carried out by Deutcher et al. [21] and Perko et al. [22] on Tropical rainforests but also on European forest sites by using high-resolution (Spotlight) X band data for estimating the tree height by using interferometric methods. A few research works have also been focused on the investigation of Mediterranean forests, although with the specific aim of studying forest fire events and the successive re-growth of biomass (e.g. [23,24]). The aim of this research is to investigate the sensitivity of multi-frequency, multi-polarization SAR data to forest characteristics and to identify the role of each frequency in the assessment of forest classification. This study, which can be considered as a preparatory step in evaluating the image information content in view of the estimate of growing stock volume, deals with the exploitation of SAR backscatter capabilities for classifying forest types in Mediterranean areas, which is a non-trivial task, due to the high spatial fragmentation and heterogeneity typical of these forests [25]. A forest area, which is a coastal plain mostly covered by evergreen conifers located in Central Italy (San Rossore), has been examined. In this area, a forest classification map was available together with ground measurements of forest parameters. SAR images were acquired from ALOS/PALSAR (L band), ENVISAT/ASAR (C band), and Cosmo-Sky Med (X band). SAR data were used for discriminating forest from non-forest land covers and separating broadleaved from coniferous forest types by using compositions of multi-temporal and multi-frequency images. The mean backscattering coefficient (σ°) was computed from the pixels associated to coniferous and broadleaf obtained from the in-situ classification map. Analysis of SAR image compositions by using various frequencies and polarizations, pointed out a notable capability of higher frequencies (C and X band) in identifying broad-leaf and coniferous forests, whereas L band is mostly able to separate forest from non-forest areas. The classification was then performed by supervised Bayesian prediction to the multi-polarization or multi-temporal features. The originality of the research relies on the application of multi-frequency SAR images to heterogeneous and mixed forests of Central Italy and on the use of a robust method based on a quadratic Bayesian classifier applied to the SAR images in different combinations of frequencies and polarizations. This specific classifier is able to overcome the limits of ground-truth classes that contain not homogenous targets (i.e. non-forest class). This peculiarity can be considered innovative with respect to conventional classification of multi-polarization/ multi-frequency SAR images. Moreover, the role of each frequency was better identified by integrating the different contributions.

The paper is organized as follows. First, the test area and the experimental data are introduced, and then the processing methods applied to the SAR images are defined. The capability of SAR images at different bands (L, C, and X) in mapping the forest area and the main forest types is subsequently described in section 3, by identifying the role of each frequency in this task. A clear separation between non-forest and forest classes, the latter subdivided in broadleaf and coniferous, was subsequently achieved by applying a quadratic Bayesian classifier. A discussion on the achieved results and some conclusions are drawn in the final section.

Materials

The investigation was carried out in a forest area in Central Italy, where ground measurements, meteorological information and other ancillary data were available (see Figure 1). The natural park of San Rossore (43.72° N, 10.30° E) is a protected flat area of about 4800 ha located along the coast of Tuscany. The land is covered by forests, pastures and fields; forests are dominated by Mediterranean pines (Pinus pinaster Ait. and Pinus pinea L.), holm oak (Quercus ilex L.) and several broadleaved species. The ground truth is represented by the forest type map produced by 'Dimensione Ricerca Ecologia Ambiente', DREAM (2003) [26]. The original classification map was provided at the 1:15000 scale and was derived from field observations collected in the whole Park. According to the definition used by the Tuscany Regional authority, forests correspond to areas having a minimum extension of 2000 m^2 and length greater than 20 m; tree cover has to be greater than 20%. Unfortunately, evergreen broadleaf forests (Holm oak) cover only a marginal area (0.2%) of the whole Park, thus, it was not possible to separate them. Therefore, the identified classes correspond to coniferous and deciduous broadleaf forests. Logging activities have interested part of the forest area since 2009; therefore, a preliminary check was done to exclude these areas from the training and the test phases. Felled areas were identified using a Landsat TM of 2009 and Google Earth images

of 2010.Additional conventional measurements were carried out on 72 forest stands covered by three forest types: Mediterranean pines, holm oak and deciduous trees, whose area ranged from 1 to 170 ha [27]. A series of SAR images, listed in Table 1, was collected at L (ALOS/PALSAR), C (ENVISAT/ASAR), and X (COSMO-SkyMed) bands in 2009 and 2010 in different seasons and different orbital configurations in terms of incidence angle, polarization, and mode.

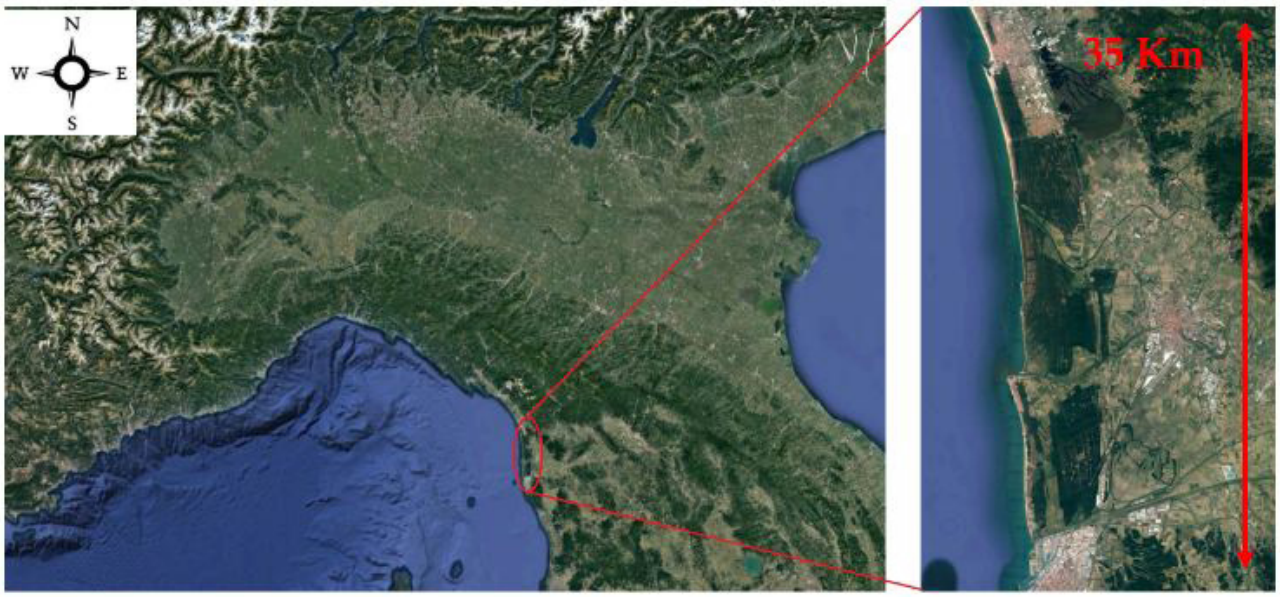


Figure 1: Italian forest test area of San Rossore in Tuscany (from Google Earth).

Group	Sensor	Date	Time (UTC)	Pass	Inc. Ang.	Pol.	Mode
1	PALSAR	28/02/2009	21:43:00	Asc	38	HH	FBS
	ASAR	26/02/2009	9:38:29	Desc	23	VV	IMS
	CSK2	6/3/2009	5:13:15	Asc	33	HH	Himage
2	PALSAR	7/6/2009	21:03:08	Asc	22	HH/HV/VH/VV	POL
	ASAR	26/05/2009	20:59:38	Asc	23	VV	IMS
	CSK2	25/05/2009	5:12:24	Asc	33	HH	Himage
3	PALSAR	29/06/2009	21:41:48	Asc	38	HH/HV	FBD
	ASAR	27/06/2009	9:35:57	Desc	23	VV	IMS
	CSK2	25/05/2009	5:12:24	Asc	33	HH	Himage
4	PALSAR	16/07/2009	21:44:03	Asc	38	HH/HV	FBD
	ASAR	16/07/2009	9:38:31	Desc	23	VV	IMS
	CSK2	13/08/2009	5:11:29	Asc	33	HH	Himage
5	PALSAR	29/09/2009	21:42:14	Asc	38	HH/HV	FBD
	ASAR	24/09/2009	9:38:26	Desc	23	VV	IMS
	CSK2	29/08/2009	5:11:29	Asc	33	HH	Himage

6	PALSAR	16/10/2009	21:44:25	Asc	38	HH/HV	FBD
	ASAR	13/10/2009	20:59:34	Asc	23	VV	IMS
	CSK	*N/A	*		*		*
7	PALSAR	30/12/2009	21:42:15	Asc	38	HH	FBS
	ASAR	22/12/2009	20:59:34	Asc	23	VV	IMS
	CSK3	20/12/2009	5:10:04	Asc	34	HH	Himage
8	PALSAR	16/01/2010	21:44:22	Asc	38	HH	FBS
	ASAR	10/1/2010	21:02:24	Asc	23	VV	IMS
	CSK2	20/01/2010	5:09:38	Asc	33	HH	Himage
9	PALSAR	14/02/2010	21:42:06	Asc	38	HH	FBS
	ASAR	11/2/2010	21:02:24	Desc	23	VV	IMS
	CSK2	21/02/2010	5:09:15	Asc	33	HH	Himage
10	PALSAR	1/4/2010	21:41:07	Asc	38	HH	FBS
	ASAR	3/4/2010	9:35:32	Desc	23	VV	IMS
	CSK3	25/03/2010	5:08:25	Asc	33	HH	Himage
11	PALSAR	18/04/2010	21:23:47	Asc	38	HH	FBS
	ASAR	22/04/2010	9:38:20	Desc	23	VV	IMS
	CSK	*N/A	*	*	*		*
12	PALSAR	19/07/2010	21:42:52	Asc	38	HH/HV	FBD
	ASAR	20/07/2010	20:59:36	Asc	23	VV	IMS
	CSK3	1/8/2010	5:07:28	Asc	34	HH	Himage

Table 1: SAR images available in the test area of San Rossore.

Methods

Preliminary Image Analysis

This investigation aimed at evaluating the use of the available SAR data for discriminating forest from non-forest land covers and separating broadleaved from coniferous forest types. Thus, a preliminary analysis oriented toward a better understanding of the role of each frequency in the assessment of forest features, was carried out in order to choose the optimal set of bands to be used in the classification process. First of all, all the available SAR images were pre-processed by using SARSCAPE©. The entire image dataset is in single-look complex slant range format, which does not include radiometric corrections; consequently, a significant radiometric bias is still present. The first step was therefore the radiometric calibration that was obtained by modifying the image pixel values through multiplication and division of calibration factors, as the antenna gain, the calibration constant and the Local Incidence Angle (LIA). These parameters are contained in the ancillary data of each SAR image, while LIA was computed through the Digital Elevation Model (DEM) of the observed area combined with the orbital parameters (i.e. satellite orbit state vector contained in the ancillary data). In this processing phase, orbital parameters and DEM were also used for generating layover and shadow maps, in order to exclude pixels that belong to areas affected by these phenomena. The radiometric correction provided imagery in which pixel values truly represent the radar backscatter of the reflecting surface. This step is necessary for the comparison of SAR images acquired with different sensors, or acquired from the same sensor but at different times and modes, or processed by different processors. The next step concerns the despeckling, obtained by using a multilook filter. As each SAR sensor has a different spatial resolution, the window filter size was selected according which SAR image had to be despeckled. In this case, the window sizes (in range and azimuth, respectively) were selected as follows: 1×2 pixels (PALSAR), 1×5 pixels (ASAR), 2×2 pixels (CSK). With respect to the low values of the despeckling window size, this choice allowed to do not lose spatial resolution inside the test area. Successively, since the SAR images were in the 2D raster

radar geometry (i.e. slant range view), a precise geolocation process was applied to the images by using orbit state vector information, radar timing annotations, the slant to ground range conversion parameters, and the reference DEM data. The resulting geocoded images have a pixel size of 10 m × 10 m and the same projection (UTM 32). Later on, the SAR images were co-registered by shifting each image in the stack file by using ENVI, in order to relate the same pixel to the same geocoded target and allowing a 'pixel by pixel' comparison among the various images.

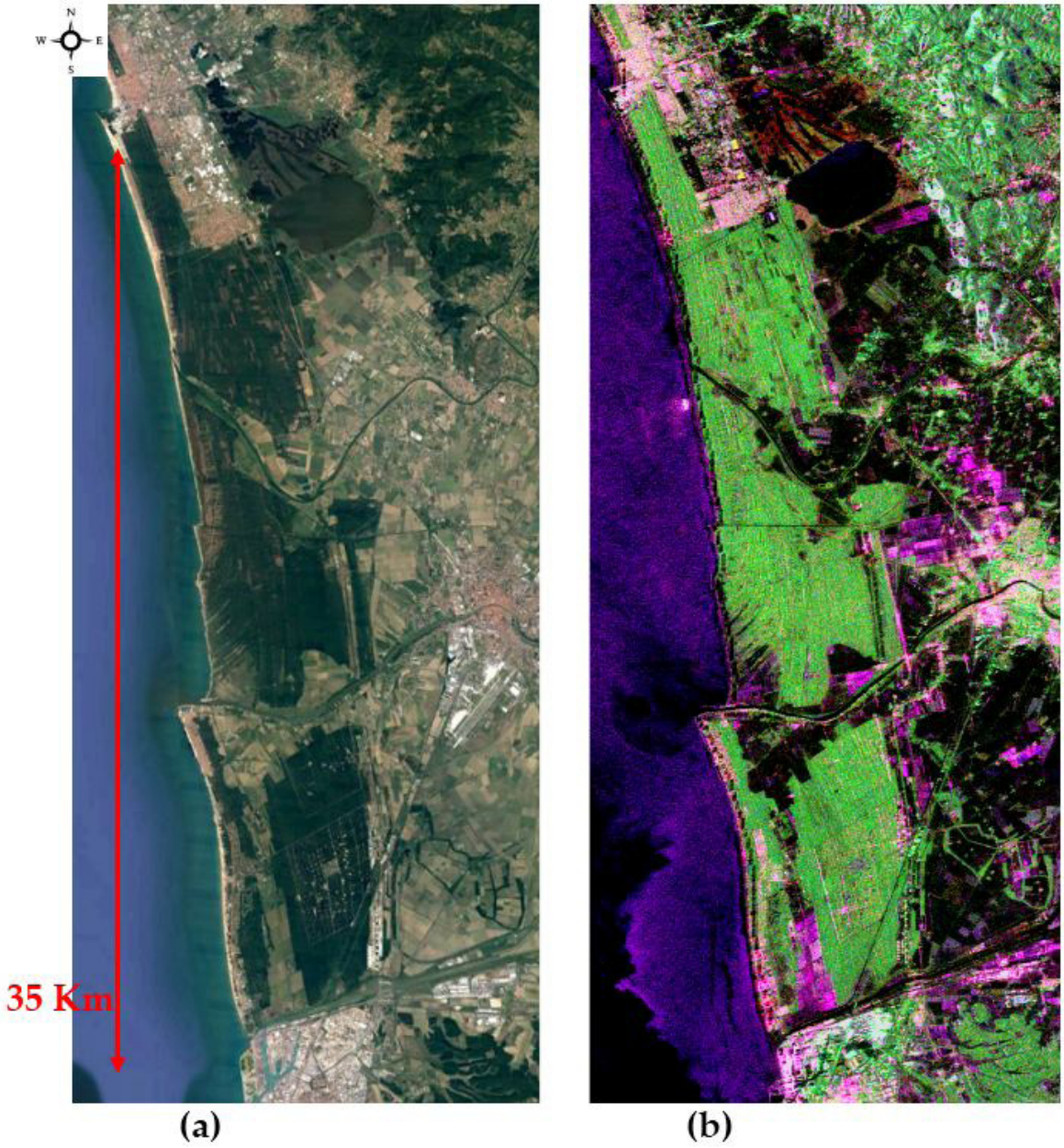


Figure 2: RGB image of San Rossore area on top right side, compared with a Google Earth image **(a)**. PALSAR (L-band) 7/6/2009. R:HH, G:HV, B:VV. Bright green areas of RGB exactly correspond to forest areas in the Google image **(b)**.

Afterwards, some RGB compositions, derived from the available SAR images of Table 1, were prepared for each area in order to check the classification capability of different frequencies. RGB images have been compared with the available ground truth. As an example, in Figure 2, a RGB visualization of ALOS/PALSAR images collected on San Rossore area in June 7th 2009 at three polarizations is shown (R: HH pol., G: HV pol., B: VV pol.), which allowed a preliminary but clear identification of forest areas with respect to other type of surfaces. In this figure, an optical image (gathered from Google Earth) is also shown as a visual comparison. Bright green areas of RGB exactly correspond to forest areas in the Google image, pointing out the role of HV polarization in identification of forest areas. Cross polarization at L band is, in fact, mainly sensitive to inclined cylinders of dimensions similar to the observation wavelength (about 20 cm) and therefore represented by thick branches and trunks. When inclined cylinders are observed, σ° shows a maximum whose position tends toward lower diameter values as the frequency increases [10,28]. For each frequency the range of diameters producing the maximum values of σ° is a fraction of the wavelength (1/10-1/20).

A more in-depth analysis was carried out on the entire SAR images dataset collected on the area, taking into account only the portion of the image covered by forests and considering the in-situ classification of the area produced by DREAM (2003) [26] and considered as a reference for validating the image classification methods (Figure 3).

The backscattering coefficient value (σ°) of each pixel contained in every image, for each SAR frequency and polarization, has been added to the corresponding co-registered pixel of the other available images of Table 1. σ° has then been normalized by the number of images, thus obtaining a single image, in order to maximize the signal intensity with respect to speckle and consequently to attain a better classification. The σ° averaging was also carried out for minimizing the seasonal effects (such as soil moisture variations, presence/absence of leaves, higher or lower tree water content) on the classification procedures. As a second step, the mean σ° was computed from the pixels associated to coniferous and broadleaf forests obtained from the classification map (red and green areas, respectively). These averaged values are represented in Table 2 for each SAR sensor (PALSAR, ASAR, and CSK) at different polarizations, along with their standard deviation values. We can generally note that σ° values are similar for coniferous and broadleaf at L-band (differences lower than 0.5 dB), while σ° at C and X band is higher on coniferous forests than on broadleaf, with differences of about 2 dB at both frequencies. Looking at this table, the different sensitivities at various frequencies to forest characteristics are well pointed out. At X and C bands, the observation wavelength (between 3 and 6 centimeters, respectively) is comparable with the dimensions of needles and leaves, being therefore able to better identify these surface characteristics. On the other hand, at L band, the longer wavelength (about 20 cm) has a higher penetration power inside vegetation cover and is less influenced by crown characteristics.

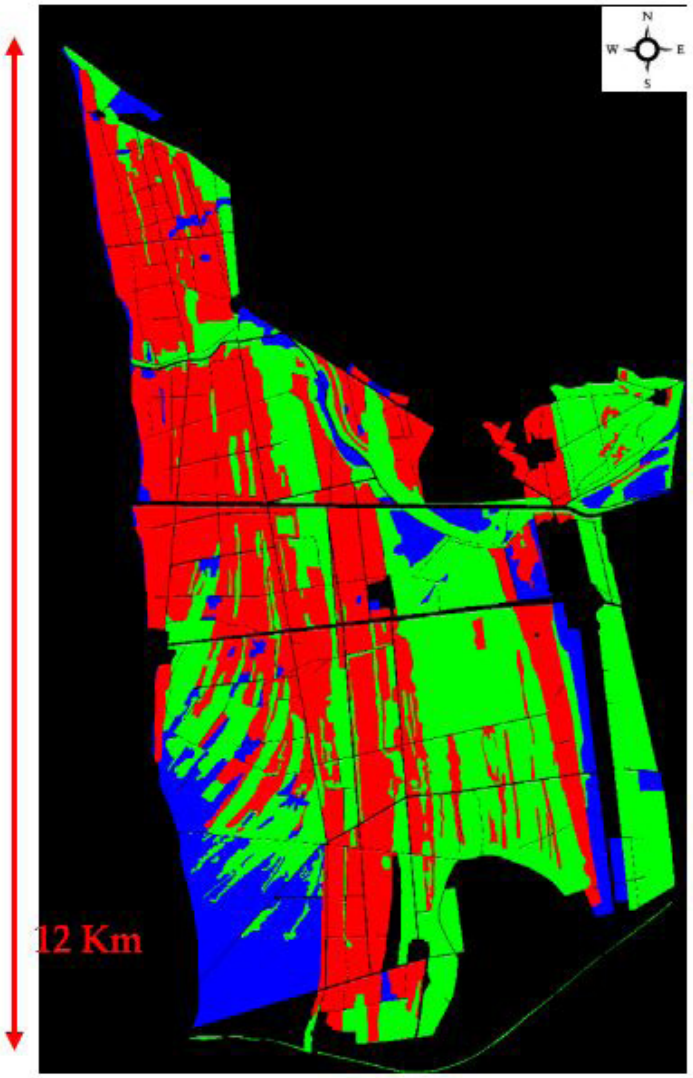
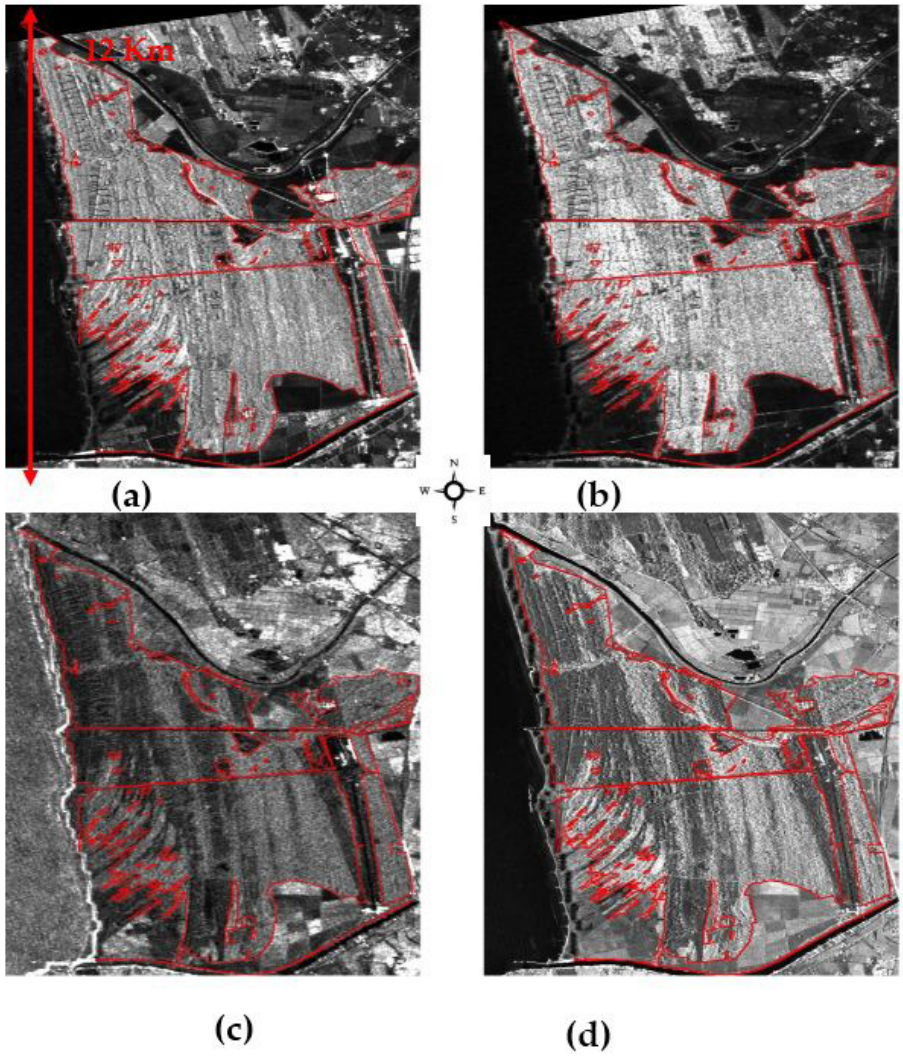


Figure 3: Forest type map produced by DREAM (2003) [26]. Red: coniferous, Green: broadleaf, Blue: Non-forest, Black: unclassified.

Sensor/Polarization	Frequency/Band	Coniferous	Broadleaf	Coniferous	Broadleaf
		Mean s^o (dB)		St. Dev. S^o (dB)	
PALSAR HH	1.27 GHz/L	-8.41	-8.14	1.31	1.37
PALSAR HV	1.27 GHz/L	-13.49	-13.95	1.46	1.36
ASAR VV	5.3 GHz/C	-9.8	-8.13	1.09	0.93
CSK HH	9.65 GHz/X	-12.67	-10.53	1.95	1.81

Table 2: Mean values of σ° obtained from PALSAR, ASAR and CSK images, for the two forest types (coniferous and broadleaf) classified in Figure 3.



Figures 4(a-d): σ° mean values computed over the San Rossore test site for each SAR sensor and polarization available from the dataset of Table 1. The red line identifies the contour of forest area, derived from the map produced by DREAM (2003) of figure 3. **(a)** PALSAR σ° HH pol.; **(b)** PALSAR σ° HV pol.; **(c)** ASAR σ° VV pol.; **(d)** CSK σ° HH pol.

This behavior is also visible looking at the SAR images of Figure 4a and Figure 4b, which represent the averaged images at L band in HH (a) and HV (b) polarizations. The different forest features can be weakly identified in HH polarization, whereas they are more visible in HV polarization, the latter being more influenced by inclined cylinders represented by branches and trunks. However, also

in this case, the discrimination between the two types of forest is not trivial. In Figures 4c and 4d the average σ° at C and X-band is represented, and we can note that the features inside the forest area are more evident, and σ° at these bands seems to be able to better identify coniferous and broad-leaf forests than L-band. In order to better check the sensitivity of C-band to forest features, a multi-temporal RGB image obtained at C band from ENVISAT/ ASAR data in 2009 is shown in Figure 5 (R: 26 Feb., G: 16 Jul., B: 24 Sept.). At this frequency, as expected, the temporal variability is highest on agricultural surfaces and very low on urban areas (bright, white areas). On forests, the moderate variability is due to the presence/absence of leaves from winter to summer season in broadleaf forests.

The information of interferometric coherence [29], which was generated by using suitable couples of images for each SAR sensor (i.e. repeat-pass interferometry: images couples acquired from slightly different orbits at different times) was also added for improving the classification of the area. Figure 6a shows a composite RGB image obtained using PALSAR averaged data of σ° of 12 HH, (R), σ° of 6 HV (G), and interferometric coherence (B). The latter parameter was obtained from the couple of PALSAR images of 29 June - 29 September, 2009 (with orthogonal baseline, Bort, ≈ 740 m) in HH polarization. Several combinations of polarizations and coherence have been preliminarily tested with ranges of Bort between 323 m and 2954 m. However, all the considered image couples produced almost the same results: the coherence intensity prevails in agricultural and anthropic (i.e. non-forest) areas, while σ° in HH and mainly in HV polarization dominates in forested areas, due to volume scattering contribution. Figure 6b shows a multi-temporal RGB composite image from CSK data. In this case, RGB was applied to multi-temporal images (August 13 and 29, 2009), since at X band only VV or HH polarizations were available. As in the previous case, the coherence dominates in non-forest areas, while the intensity in co-polarized bands prevails in forest area. Differences inside the forest areas are visible and roughly correspond to the two types of forest. The same investigation, carried out by using interferometric coherence at C band, did not produce any significant result and it was therefore not reported in the paper.

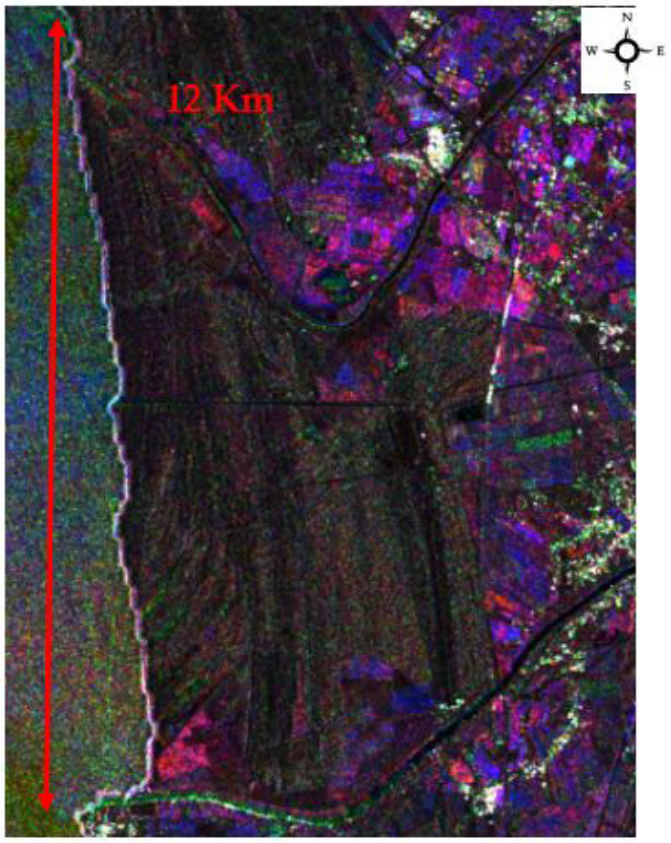
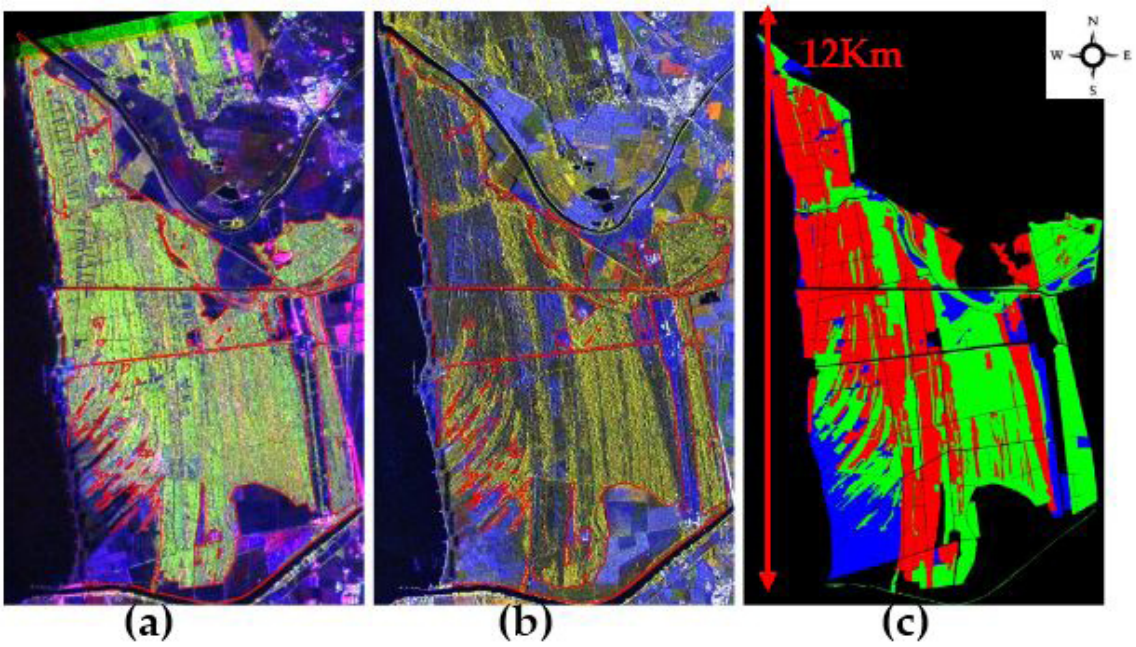


Figure 5: RGB map of ENVISAT/ASAR data collected in 2009 on San Rossore area. (R: 26 February, G: 16 July, B: 24 September).



Figures 6(a-c): RGB composite images for (a) PALSAR - R: Mean σ° HH pol., G: Mean σ° HV pol., B: Coherence obtained by using the PALSAR images of 29 June - 29 September, 2009, in HH pol.; (b) CSK - R: 13 Aug. 2009, G: 29 Aug 2009, B: Coherence. The red line identifies the forested area. (c) Forest classification map produced by DREAM (2003). Legend: Red: coniferous, Green: broadleaf, Blue: Non-forest, Black: unclassified.

Image Classification

After evaluating the sensitivity of various frequencies to the different surface features and forest characteristics, a preliminary statistical analysis was carried out in order to obtain a correct classification map according to the available ground truth data.

The dataset considered for this task consists of six (P=6) coregistered SAR images as follows:

1. one PALSAR image that is the mean of 12 HH PALSAR images;
2. one PALSAR image that is the mean value of 6 HV available image;
3. PALSAR coherence map obtained by the couple of 29 June - 29 September, 2009, HH pol.;
4. CSK image acquired on 13 August 2009;
5. CSK image acquired on 29 August 2009;
6. CSK coherence map generated by the CSK images 13, 29 August 2009.

The classification of SAR images is generally not trivial when the number of looks is low, as shown in the previous paragraph. Because of this reason, the speckle was further reduced by applying the Kuan filter on a 7x7 sliding window. The use of kuan filter is used to improve classification results because it minimizes the variability due to speckle effect and consequently supports the classification task in identifying the various classes.

The histogram of the resulting class features was generated in order to find out, for each class, the kind of statistical distribution.

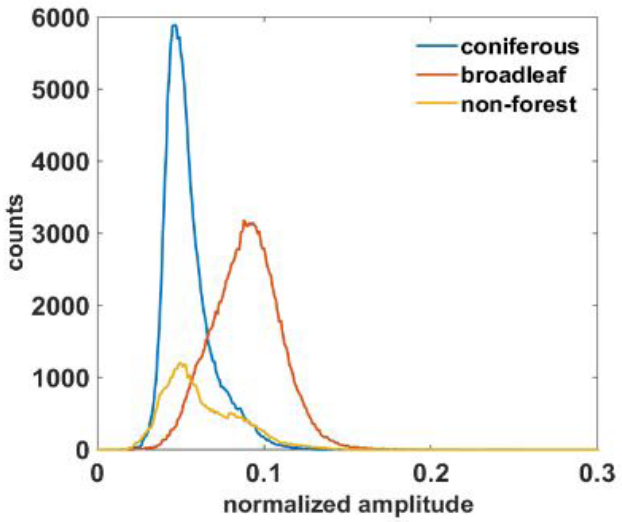


Figure 7: Histogram of the despeckled HH polarized CSK amplitude values.

As an example, in Figure 7 the histogram of despeckled HH polarized CSK amplitude values is shown for all classes. It can be observed that coniferous and broad-leaf features are well represented by Gaussian distributions, as expected for despeckled amplitude values. The class 'non-forest' does not perfectly fit the Gaussian distribution, although this fact can be attributed to the heterogeneity of this class, where different surface features as marine environment, agricultural fields and small anthropic settlements are contained. The classification of the test site is then accomplished by applying a Bayesian quadratic discriminant analysis [30] according to which both means and covariances of each class can vary according to an image probabilistic model, by assuming a Gaussian mixture class distribution. Other types of discriminant functions could be applied in order to maximize the discriminatory power and separate the available classes effectively [31,32] as, e.g., nearest neighbor, decision trees, linear functions, nonlinear functions, and so on. The main problem of these classic supervised classifiers is their sensitivity to the Hughes effect [30] which is critical for this specific classification problem. In fact, the dimension of the selected training set (1% of the ground truth samples are randomly selected, i.e. 34472 samples) is very high with respect to the number of features, which is 6, when both CSK and PALSAR data are processed. As a consequence, non-linear strategies such as Quadratic Discriminant Analysis (QDA) are more suitable in this case and the proposed classifier based on the Bayesian quadratic discriminant analysis provides the best performance when data can be modeled with a Gaussian mixture distribution. Looking at the diagram of Figure 7, the classification accuracy of coniferous and broad-leaf classes is therefore expected to be very high, while the non-forest class would need further processing step. Since the non-forest class is not homogeneous, an unsupervised kmeans [30] classification algorithm was applied to separate two sub-classes inside this class (i.e. NFa and NFb). The number of subclasses was suggested by the shape of the histogram of the non-forest class, which looks like a bimodal distribution. The classifier used in this paper is better described in the following. Let $X = \{x_1, \dots, x_N\}$ denote the N-pixel multiband image to be classified. In case of the classification of CSK data, X contains P features, where P is the number of bands of the dataset stack. Three configurations of the considered dataset were selected: i) only PALSAR data ($P = 3$), ii) only CSK data ($P = 3$) and iii) combined PALSAR with CSK datasets ($P = 6$). In the CSK case, we selected the coherence image acquired on August 13, 2009, and the two HH polarized CSK amplitude images, acquired on August 13 and 29, 2009, respectively. Therefore pixel x_n can be represented by a column vector in a PP-dimensional space. The supervised classification of X into $K = 4$ classes $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, namely coniferous, broadleaf, and non-forest (the latter split in two subclasses), is performed according to the following procedure. In the training phase, a small percentage of the ground truth pixels, typically 5-10% of each class, is randomly selected as training samples. The sample mean of each class is then computed as an estimate of the class mean:

$$\hat{\mu}_k = \frac{1}{N_t}\sum_{n=1}^{N_t^{(k)}} x_n \qquad k = 1, \dots, K \qquad (1)$$

where $N_t^{(k)}$ is the number of training pixels belonging to class k.

Then the sample covariance is computed for each class by first subtracting the sample mean of each class from the observations of that class, and taking the empirical $P \times P$ covariance unbiased estimate:

$$\hat{\Sigma}_k = \frac{1}{N_t^{(k)} - 1}\sum_{n=1}^{N_t^{(k)}} (x_n - \hat{\mu}_k)(x_n - \hat{\mu}_k)^T \qquad k = 1, \dots, K \qquad (2)$$

A maximum a-posteriori criterion is finally applied to classify each pixel $x_n \in X$, which reduces to the maximum likelihood decision by assuming uniform prior class probabilities (Hastie et al., 2001):

$$\omega(x_n) = arg \min_{\omega = 1, \dots, K} \sum_{k=1}^{K} P(x_n \mid k) C(y \mid k) \qquad (3)$$

where, by hypothesis,

$$P(x \mid k) = \frac{1}{(2\pi|\hat{\Sigma}_k|)^{1/2}} \exp\left(-\frac{1}{2}(x - \mu_k)\hat{\Sigma}_k^{-1}(x - \mu_k)^T\right) \qquad (4)$$

As it can be noted from equation (3), the Bayesian classifier was applied to the resulting four classes by applying the following cost matrix C, in which each element adjusts the relationships among the classes:

$$C = \begin{array}{c} \\ CO \\ BL \\ NFa \\ NFb \end{array} \begin{array}{c} \begin{array}{cccc} GTCO & GTBL & GTNFa & GTNFb \end{array} \\ \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{bmatrix} \end{array}$$

where:

GTCO: Ground-Truth Coniferous class.

GTBL: Ground-Truth Broad-Leaf class.

GTNFa: Ground-Truth Non-forest subclass a.

GTNFb: Ground-Truth Non-forest subclass b.

CO: Coniferous class.

BL: Broad-Leaf class.

NFa: Non-forest subclass a.

NFb: Non-forest subclass b.

Each element of the *C* matrix affects the Bayesian classifier by forcing the classifier to promote or limit the migration of pixels from a class to another. More in depth, the selected cost matrix was the following:

$$C = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0.2 & 0.2 & 0 & 0 \\ 0.2 & 0.2 & 0 & 0 \end{bmatrix}$$

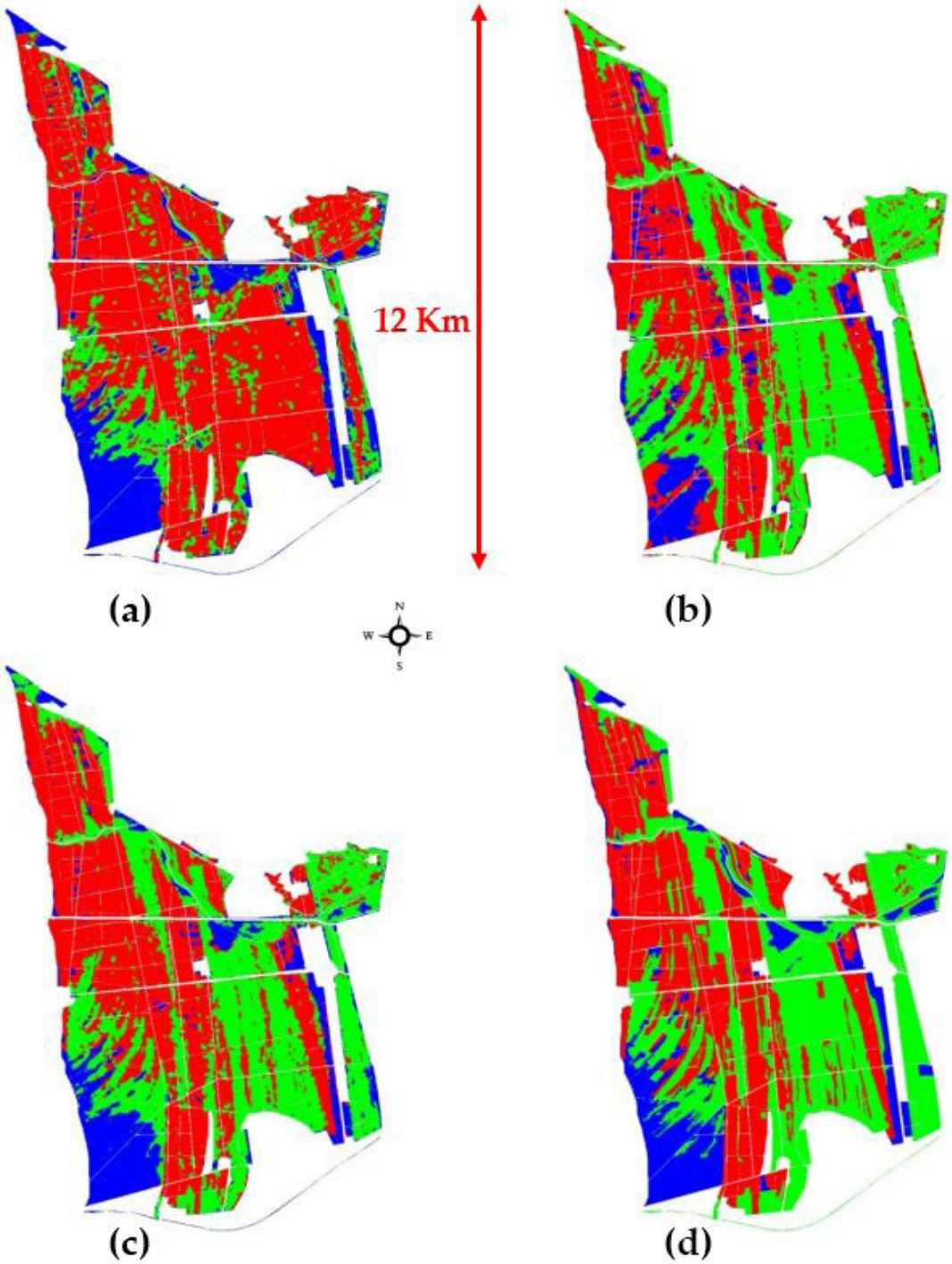
This matrix forces the classifier to operate as follows:

- It does not penalize the transitions the pixels correctly classified because each element of the diagonal is set to zero.
- It does not penalize transition of the pixels between the subclass ***NFa*** and ***NFb*** because C_{43} and C_{34} are set to 0.
- It strongly hampers the transition of pixels from ***CO*** and ***BL*** classes, since C_{12} and C_{21} are set to one.
- It penalizes the transition of pixels belonging to ***CO*** and ***BL*** towards ***NFa*** and ***NFb*** because C_{13}, C_{23}, C_{14} and C_{24} are set to 1.
- It lightly penalizes the transitions of pixels from ***NFa*** and ***NFb*** towards ***CO*** and ***BL*** C_{31}, C_{32}, C_{41} and C_{42} are set to 0.2.

This specific setting of the cost matrix *C* enables the classifier to overcome the limits of ground-truth classes that contain not homogenous targets (i.e. non-forest class in this case). This peculiarity can be considered innovative with respect to conventional classification of multi-polarization/multi-frequency SAR images.

Results

In order to quantitatively assess the performance of the classification method, the two non-forest subclasses (namely ***NFa*** and ***NFb***) were merged into a single non-forest class. This step was performed simply relabeling ***NFa*** and ***NFb*** as a unique label ***NF,*** which means to join the classified pixels in ***NFa*** together with ***NFb***. The classification maps, obtained from both X and L band, are presented in Figure 8, while a quantitative result, expressed as confusion matrices, is shown in Table 3. As stated above, non-forest areas correspond to agricultural surfaces, anthropic/ urban areas and marine environment. Figure 8a presents the classification results obtained by using PALSAR data shown in Figure 6a, i.e. PALSAR averaged backscatter image in HH and HV polarization in addition to the interferometric coherence from PALSAR image couple of 29 June - 29 September, 2009 ($B_{ort} \approx$ 740 m), in HH polarization. In this case, the identification between coniferous and broad leaf is scarce, whereas the non-forest areas are almost correctly identified as blue areas (see ground truth map of Figure 8d). Conversely, the classification map (see Figure 8b) obtained from CSK data only is really close to the in-situ map (i.e. Figure 8d), with a better discrimination between coniferous (red areas) and broadleaf (green) forests. Figure 8c shows a further classification result combining L and X band image (i.e. the entire stack considered at the beginning of this paragraph, P=6). In this case, C band data have been omitted since from preliminary classification tests they did not add any significant improvement of the classification results. The quantitative results shown in Table 3, confirmed that X band (CSK) is the more suitable frequency for identifying different types of forest: in fact, the overall classification accuracy at X band ranges from 75.21% (coniferous) to 80.49% (broadleaves); at L band (PALSAR) the average percentage of correctly identified pixels denotes nevertheless a lower accuracy for coniferous and broadleaf class. However, the L band σ° is able to identify the non-forest areas with 67.31% of accuracy, better than X band (55.63%). The confusion matrix, represented in the last columns of Table 3, shows a better classification accuracy for all the three classes (80.13%, 83.03% and 75.07% for coniferous, broadleaf, and non-forest, respectively). The joint use of L and X band allowed obtaining better classification results in comparison with L band or X band dataset only. Finally, considering the overall accuracy, the best result was obtained, once again, with the joint use of PALSAR and CSK data (see Table 3), with an overall accuracy of 80.06%.



Figures 8(a-d): Classification maps of San Rossore test site obtained by applying the Bayesian quadratic discriminant analysis by using **(a)** PALSAR, **(b)** CSK, **(c)** PALSAR and CSK; in **(d)** Reference forest classification map produced by DREAM (2003) [22]. Legend: Red: coniferous, Green: broadleaf, Blue: Non-forest, White: unclassified.

Confusion Matrix	PALSAR			CSK			PALSAR+CSK		
	Ground Truth (%)			Ground Truth (%)			Ground Truth (%)		
Coniferous	54.14	26.13	11.46	75.21	11.28	32.93	80.13	11.02	8.76
Broadleaf	43.92	66.09	21.23	15.18	80.49	11.44	16.71	83.03	16.16
Non-forest areas	1.94	7.78	67.31	9.61	8.23	55.63	3.16	5.95	75.07
Total	100	100	100	100	100	100	100	100	100
Overall accuracy	58.55			75.34			80.06		

Table 3: Confusion matrices at X and L bands for three classes: coniferous, broadleaf and non-forest areas.

Discussion

The use of SAR images for the estimation of forest features presents particular complexities in Mediterranean areas, where the vegetation cover is spatially fragmented and heterogeneous due to the long-term effect of human activities. The variable aridity conditions typical of the Mediterranean climate represent an additional challenge, since they irregularly limit foliage density, creating incomplete and discontinuous plant canopies that are difficult to characterize from remote sensing observations. Similar investigations were performed in [33] by applying Support Vector Machine (SVM) to the test site of the entire Tasmania. Here, the kind of forest and the age of the plantations was quite diversified in comparison with San Rossore test site, while the classifier algorithm was implemented to retrieve two discriminated features, i.e. forest and non-forest classes. The best results obtained with L-band, obtaining and overall accuracy of 92.1%. However, it has to be remarked that the nature of the forest in Tasmania is substantially different than the one is present in the Mediterranean areas and the proposed approach considered only two output classes, while here, the output classes consider coniferous and broadleaf classification also.

The current study addressed this issue by performing a multi-step data analysis in a forest area of Central Italy (San Rossore) representative of Mediterranean conditions. The investigation was performed using a consistent dataset of SAR images at L, C and X bands collected by ALOS/PALSAR, ENVISAT/ASAR, and Cosmo-SkyMed satellites. These images characterized by different frequencies and polarizations, were used to verify the capability of SAR in mapping forest areas and for discriminating main forest types (coniferous vs. broadleaved). The information retrieved from interferometric coherence images has also been added.

The results obtained by applying a Bayesian quadratic discriminant analysis confirmed that also in Mediterranean areas the backscattering coefficient (σ°) at L band is generally able to better identify forest areas than other frequencies, and to separate them for other land covers (with an overall accuracy of 67.31%). On the other hand, σ° at X band, in HH polarization, proved to be sensitive to forest type and capable in separating broadleaved from coniferous (accuracy of about 75-80%, resulting from the mean value between coniferous and broadleaf true-positive). The use of coherence maps added information to the classifier algorithm, allowing obtaining the results showed in table 3. Moreover, the joint use of L and X band shows that the classification task is more performing not only in terms of forest and non-forest areas discrimination (75%), but also in terms of coniferous and broadleaf classification inside forested areas, reaching an overall accuracy equal to 80%.

This behavior can be explained by the different penetration powers of these frequencies: at X-band the short wavelength (3 cm) interacts with only the first layers of the observed surfaces and with the scattering elements of similar dimensions, such as leaves (for broad-leaf) and needles (for coniferous) from the forest canopy. Due to the different response of these scattering elements, the backscattered radar signal varies according to the geometries of leaves, enabling the discrimination between coniferous and broadleaved dominated forest types. On the other hand, due to

the scarce penetration, the difference between forest and crop leaves is almost negligible and therefore at this frequency it is tough to separate forest from agricultural areas. L band signal (21 cm) instead has a great penetration power inside the surfaces and therefore is less sensitive to the differences in leaf geometry and more related to the amount of biomass over soil making the separation between forest and non-forest easier. C band backscatter, due to the intermediate penetration inside forest canopy, is more related to the conditions of crowns and to the absence/presence of leaves, and did not therefore provided significant contribution to the classification purposes.

Conclusions

The analysis performed in this study allowed drawing the following conclusions. First of all, the originality of the research consists of the application of multi-frequency SAR images to heterogeneous Mediterranean forests, which have not been so far extensively investigated by using microwave remote sensing methods. The role of L and X bands in land classification was analyzed, by applying the proposed classification method to each SAR frequency and also considering the integration of the different frequencies. On the basis of the available SAR data set, the joint use of multi-frequency SAR images increases the ability in correctly classifying heterogeneous forests, allowing the separation of forest areas from non-forest class, as well as the identification of broad-leaf and coniferous classes inside the forest class. The overall accuracy reaches 80% by integrating both L and X band contributions, whereas by considering separately L and X band the reached accuracies are lower (58.55% and 75.34%, respectively).

Another asset of this study is the particular setting of the cost matrix inside the quadratic Bayesian classifier method. The application of this method, according the procedure presented here, allowed overcoming the limits of ground-truth classes that contain not homogenous targets (i.e. non-forest class). This peculiarity can be considered innovative with respect to conventional classification methods used for multi-polarization/multi-frequency SAR images. This research can be also interesting in view of OptiSAR Constellation mission, devoted to the Earth surface observation by means of spaceborne optical, L and X band SAR sensors, with the aim of developing consistent applications in environmental, hazard and safety monitoring.

Further test and validation of the presented methodology should be advisable by extending the investigation to other datasets and test areas.

Acknowledgments

Italian Space Agency (ASI) partially supported this research by providing the Cosmo-SkyMed images of the test areas through the 'SAR for Biomap' project. The ALOS/PALSAR images were provided under the AO project by JAXA (PI PIGM111T), whereas the ENVISAT/ASAR images were provided by ESA under the Category-1 proposal "SAR for Biomap" (Project id 28510). The authors wish to warmly thank Drs. Fabio Maselli and Marta Chiesi and Prof. Gherardo Chirici for providing ground truth data of the San Rossore area and their useful comments and suggestions.

Author Contributions

All the authors equally contributed to the work reported.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Waring RH, Running S (2007) Carbon Cycle, in: Forest Ecosystems (Third Edition). Academic Press Pg No: 59-98.
2. McRoberts RE (2008) Using satellite imagery and the k-nearest neighbors' technique as a bridge between strategic and management forest inventories. Remote Sens Environ 112: 2212-2221.
3. Guyot G, Guyon D, Riom J (1989) Factors affecting the spectral response of forest canopies: a review. Geocarto Int 3: 3-18.
4. Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ 8: 127-150.
5. Ulaby FT, Dobson MC (1989) Handbook of Radar Scattering Statistics for Terrain, Artech House. Norwood, MA.
6. Ulaby FT, Whitt MW, Dobson MC (1990) Measuring the Propagation Properties of a Forest Canopy Using a Polarimetric Scatterometer. IEEE Trans. Antennas Propag 38: 251-258.
7. Hoekman DH (1993) Radar signature and forest vegetation., in: In: Land Observation by Remote Sensing, Theory and Applications, H.J. Buiten, P.G.W. Clevers (Eds.). Gordon and Breach Sci. Publ., Switzerland. Ch.11 (1993) 219-235. New York, USA. Pg No: 219-235.
8. Paloscia S, Macelloni G, Pampaloni P, Sigismondi S (1999) The potential of C- and L-band SAR in estimating vegetation biomass: the ERS-1 and JERS-1 experiments. IEEE Trans. Geosci. Remote Sens 37: 2107-2110.
9. Le Toan T, Laur H, Mougin E, Lopes A (1989) Multitemporal and dual-polarization observations of agricultural vegetation covers by X-band SAR images. IEEE Trans. Geosci. Remote Sens 27: 709-718.
10. Ferrazzoli P, Paloscia S, Pampaloni P, Schiavon G, Sigismondi S, et al. (1997) The potential of multifrequency polarimetric sar in assessing agricultural and arboreous biomass. IEEE Trans Geosci Remote Sens 35: 5-17.
11. Kasischke ES, Melack JM, Dobson MC (1997) The use of imaging radars for ecological applications - A review. Remote Sens. Environ 59: 141-156.
12. Ackermann N, Thiel C, Borgeaud M, Schmullius C (2012) Cosmo-SkyMed backscatter intensity and interferometric coherence signatures over Germany's low mountain range forested areas, in: International Geoscience and Remote Sensing Symposium (IGARSS). Pg No: 5514-5517.
13. Le Toan T, Beaudoin A, Riom J, Guyon D (1992) Relating Forest Biomass to SAR Data. IEEE Trans Geosci Remote Sens 30: 403-411.
14. Saatchi S, Marlier M, Chazdon RL, Clark DB, Russell AE (2011) Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. Remote Sens. Environ 115: 2836-2849.
15. Woodhouse IH, Mitchard ETa, Brolly M, Maniatis D, Ryan CM (2012) Radar backscatter is not a "direct measure" of forest biomass, in: Clevers and HJ Buiten (Ed.). Nature Climate Change. Gordon and Breach Science Pub, New York, USA. Pg No: 556-557.

16. Baghdadi N, Le Maire G, Bailly JS, Osé K, Nouvellon, et al. (2015) Evaluation of ALOS/PALSAR L-Band Data for the Estimation of Eucalyptus Plantations Aboveground Biomass in Brazil. IEEE J Sel Top Appl Earth Obs Remote Sens 8: 3802-3811.

17. Santoro M, Eriksson L, Askne J, Schmullius C (2006) Assessment of stand□wise stem volume retrieval in boreal forest from JERS□1 L-band SAR backscatter. Int J Remote Sens 27: 3425-3454.

18. Santoro M, Askne J, Beer C, Cartus O, Schmullius C, et al. (2008) Automatic model inversion of multi-temporal C-band coherence and backscatter measurements for forest stem volume retrieval, in: International Geoscience and Remote Sensing Symposium (IGARSS).

19. Grover K, Quegan S, da Costa Freitas C (1999) Quantitative estimation of tropical forest cover by SAR. IEEE Trans Geosci Remote Sens 37: 479-490.

20. Santos JR, Freitas CC, Araujo LS, Dutra LV, Mura JC, et al. (2003) Airborne P-band SAR applied to the aboveground biomass studies in the Brazilian tropical rainforest. Remote Sens. Environ 87: 482-493.

21. Deutscher J, Perko R, Gutjahr K, Hirschmugl M, Schardt M (2013) Mapping tropical rainforest canopy disturbances in 3D by COSMO-SkyMed spotlight in SAR-stereo data to detect areas of forest degradation. Remote Sensing 5: 648-663.

22. Perko R, Raggam H, Deutscher J, Gutjahr K, Schardt M (2011) Forest Assessment Using High Resolution SAR Data in X-Band. Remote Sensing 3: 792-815.

23. Minchella A, Del Frate F, Capogna F, Anselmi S, Manes F (2009) Use of multitemporal SAR data for monitoring vegetation recovery of Mediterranean burned areas. Remote Sensing of Environment 113: 588-597.

24. Tanase M, de la Riva J, Santoro M, Pérez-Cabello F, Kasischke E (2011) Sensitivity of SAR data to post-fire forest regrowth in Mediterranean and boreal forests. Remote Sensing of Environment 115: 2075-2085.

25. Scarascia-Mugnozza G, Oswald H (2000) Forests of the Mediterranean region: gaps in knowledge and research needs. For Ecol Manage 132: 97-109.

26. DREAM (2003) Tenuta di San Rossore. Note illustrative della carta forestale e della fruizione turistica. Published by S.E.L.C.A. Firenze, Italy.

27. Corona P, Chirici G, Marchetti M (2002) Forest ecosystem inventory and monitoring as a framework for terrestrial natural renewable resource survey programmes. Plant Biosyst 136: 69-82.

28. Ferrazzoli P, Guerriero L (1995) Radar sensitivity to tree geometry and woody volume: a model analysis. IEEE Trans. Geosci. Remote Sens 33: 360-371.

29. Bamler R, Hartl P (1998) Synthetic aperture radar interferometry Synthetic aperture radar interferometry. Inverse Probl 14: 55.

30. Richards JA (2013) Remote Sensing Digital Image Analysis, Remote Sensing Digital Image Analysis: An Introduction. Berlin Heidelberg.

31. Bandos TV, Bruzzone L, Camps-Valls G (2009) Classification of hyperspectral images with regularized linear discriminant analysis. IEEE Trans. Geosci Remote Sens 47: 862-873.

32. Bioucas-Dias JM, Plaza A, Camps-Valls G, Scheunders P, Nasrabadi, N, et al. (2013) Hyperspectral Remote Sensing Data Analysis and Future Challenges. IEEE Geosci Remote Sens Mag 1: 6-36.

33. Mitchell AL, Tapley I, Milne AK, Williams ML, Zhou ZS, et al. (2014) C- and L-band SAR interoperability: Filling the gaps in continuous forest cover mapping in Tasmania. Remote Sensing of Environment 155: 58-68.