Multitemporal Level-1β products: Definitions, Interpretation, and Applications

Donato Amitrano, *Student Member, IEEE,* Francesca Cecinati, Gerardo Di Martino, *Member, IEEE,*, Antonio Iodice, *Senior Member, IEEE,* Pierre-Philippe Mathieu, Daniele Riccio, *Fellow, IEEE,* Giuseppe Ruello, *Member, IEEE*

***Abstract*—In this paper, we present a new framework for the fusion, representation, and analysis of multitemporal SAR data. It leads to the definition of a new class of products representing an intermediate level between the classic Level-1 and Level-2 products. The proposed Level-1β products are particularly oriented toward non-expert users. In fact, their principal characteristics are the interpretability and the suitability to be processed with standard algorithms. The main innovation of the paper is the design of a suitable RGB representation of data aiming to enhance the information content of time-series. The physical rationale of the products is presented through examples in which we show their robustness with respect to sensor, acquisition mode and geographic area. A discussion about the suitability of the proposed products with Sentinel-1 imagery is also provided, showing the full compatibility with data acquired by the new ESA sensor. Finally, we propose two applications based on the usage of Kohonen's self-organizing maps dealing with classification problems.**

***Index Terms*—Human-machine interaction, multitemporal SAR, high-level processing, image enhancement, Level-1β products, Sentinel-1, self-organizing maps**

I. Motivations

The use of synthetic aperture radar (SAR) data has been so far limited in applicative contexts, because the information content held by images rarely emerges without a high technical expertise. The development of new products providing user-friendly representations of the physical information is a necessary condition for the full exploitation of SAR sensors.

High-level processing is a key task for enhancing the interpretability and for emphasizing the underlying information content of remote-sensing images. This depends on the sensor acquiring the data. As an example, for hyperspectral acquisitions, principal components analysis is probably the most popular linear projection method. However, it presents several drawbacks [1]. Therefore, refined methods have been proposed in literature to overcome its limitations [1]–[3], and to make it possible to display the information contained in N channels (where N can be in the order of few hundreds) on a standard tristimulus RGB device through a consistent dimensionality reduction.

Dealing with SAR images, data interpretability problems are principally related to: i) the geometrical distortions induced by the side-looking acquisition geometry [4], ii) the scattering mechanisms, and iii) the gray-scale displaying, at least in non-polarimetric modality. In fact, humans can easily interpret color images, which support fast searching and comprehension of data [2].

Multiple acquisitions of the same scene represent a reasonable mean for increasing the dimensionality of SAR data through the combination of information collected along the time axis. The fusion problem is brilliantly focused in [5], where the authors suggested a series of questions an image fusion technique should answer. Indeed, these questions refer to multisensor data fusion. Therefore, we slightly reworked the aforementioned framework as follows:

1. What is the objective/application of the user?
2. What are the necessary pre-processing steps involved? (See Section III).
3. Which combination of the data is the most successful? (See Section IV-A.
4. Which is the "best" technique to fuse these data for that particular application? (See Section IV-B).

The objective of this paper is the definition of a new class of RGB SAR products exploiting multiemporal acquisitions. This answers to the first question, establishing the purpose of the whole processing to data representation.

Multitemporal SAR data have been widely exploited in the recent literature for, as an example, forest monitoring [6], flooding events [7] and land cover mapping [8]. Other, more general, suggestions and answers to the above listed questions can be found in [9] or in the past distinguished literature [10]–[12]. In this paper, we want to emphasize the usefulness of an appropriate and comfortable representation, allowing for an easier visual data mining and a better design of the information extraction process. In fact, a better knowledge/understanding of the scene implies a higher awareness of the processing necessary for information extraction.

The above listed questions can be summarized in the following one (see [13]):

5. How the enormous amount of information contained in remotely-sensed images can be extracted?

We will not provide a comprehensive answer to this question. In fact, this work mainly deals with a part of the whole problem, i.e. data representation. However, as stated in [14], "a good knowledge representation design is the most important part of solving the understanding problem". In particular, the proposed solution aims at creating an intermediate, user-

Donato Amitrano, Gerardo Di Martino, Antonio Iodice, Daniele Riccio and Giuseppe Ruello are with the Department of Electrical Engineering and Information Technology, University of Napoli Federico II, Napoli, Italy.

Francesca Cecinati and Pierre-Philippe Mathieu are with the European Space Agency, ESA ESRIN, Frascati (Rome), Italy

oriented product level between the classic Level-1 and Level-2 products. In fact, it is designed to: i) favour interpretation and visual data mining; ii) favour the use of standard processing algorithms, which are the most popular in the end-users' community. These characteristics lower the expertise required to handle data, and make the retrieval of the physical information (which is usually seen as Level-2 processing) simplified, since it takes place exploiting a product level higher than 1. As an example, dealing with crops classification, the discrimination between summer and winter cultivations is not Level-2 information, but anyway more than Level-1 one, and could be helpful for moving toward higher levels following a processing chain composed by:

- A model-driven data representation, which is just the purpose of this paper (see also [15]);
- An application-oriented information extraction process;
- A decision, which should be taken among a very limited range of options [16].

Information extraction can be performed at two different levels: analysis and inversion. They can exist autonomously, but the knowledge they can achieve is different. In particular, the analysis (data-driven) typically allows for reaching high-level information. Inversion (model-based) allows for the retrieval of objects' physical parameters, but is usually possible only in canonical situations (see as an example [17]). Therefore, the most reliable solution for solving the understanding problem is probably the integration of the two levels. This results in a framework where the analysis makes data suitable for the implementation of inverse models. In other words, data analysis should reduce the complexity of the case-study in an application-oriented environment, thus supporting the extraction of the physical information/parameter through inversion. We think that the best examples of such integration can be found in the image understanding systems literature of '70s and '80s [18]–[23]. Recently, a very interesting mixed framework has been proposed for through-the-wall radar image understanding in [24].

However, a lack of integration between these two levels was found in the recent remote-sensing literature. In fact, a great effort was made to develop automatic (often model-free) techniques for data analysis, trying to make SAR data available/accessible to end-users. Indeed, these techniques are helpful and maybe necessary in the whole information extraction process, but they can not totally replace the usage of electromagnetic models, which represent the best way to fully understand data.

SAR data interpretation requires the knowledge of the scattering models underlying the image formation and of the cognitive mechanisms ruling their perception. Thus, two kinds of knowledge coexist: an objective knowledge, represented by electromagnetic models, and a subjective knowledge, related to the users. In Section II, some aspects related to cognitive processes guiding humans' mining of information from the surrounding environment are recalled. In Section III, the general workflow and the pre-processing chain for the generation of the proposed products are described. Section IV, is devoted to the definition of the proposed multitemporal synthesis. In Section V, we present the obtained products and provide their physical interpretation through examples. In Section VI, we face two applicative scenarios in which the proposed products are exploited as input of neural networks to solve high-level classification problems. Conclusions are drawn at the end of the work.

II. Background

In Reference [25], Mendel argues that two distinct forms of knowledge exist: an *objective knowledge* and a *subjective knowledge*. SAR image interpretation is mainly a matter of subjective knowledge. An expert SAR user has more possibilities to successfully interpret data thanks to his/her habit to that semantic context. For many multidisciplinary users, mining information from the typical grey-scale Level-1 SAR image is not straightforward. Therefore, they often prefer different data sources, or leave the process of extraction of information to automatic algorithms.

This brings to the concept of *emergent semantics*. According to Reference [26], the informative content of an image, or its *meaning*, is assumed to be not an intrinsic property, but an emergent characteristic through the interaction with users. Therefore:

- Image meaning becomes *contextual*, i.e. depends on the condition in which a particular query is made and on the user performing it;
- The knowledge about the image is assessed by the user experience which, in this context, is built also by exploration/inspection of data.

The necessity of restoring the user's centrality in remote-sensing data analysis (especially concerning SAR data) has been already manifested in the past literature [15], [27]–[29] and appears evident if the data analysis process is approached from a semiotic viewpoint.

In Fig. 1, we report the Peirce's semiotic triangle [30]. It depicts how a concept is formed in our mind. Pierce claims that this mechanism consists of three inter-related elements: a sign, an object, and an interpretant. The sign is everything that can be perceived, i.e. the world as filtered by the sensor. The object is what the sign concerns, and consists in the real-world physical object. Finally, the interpretant is the understanding that the observer reaches about some sign/object relation.

This schema should highlight that the machine (at the moment) can be only a support (although in many cases indispensable) to the human activity. In fact, it can not participate to the cycle depicted in Fig. 1, apart from helping a better formation of the interpretant through its capacity of executing complex tasks in a short time (see also [27]).

The purpose of this work is to highlight the potentialities of an user-driven framework (provided that he/she is placed in the best condition to produce the association sign-object), starting from the basic activity of the information extraction process, i.e. the inspection of data. In fact, in a context in which the automatic extraction of information from large databases is still rather limited, the enhancement of user experience with data is crucial, he/she being in most cases designated for assigning the scene semantics. The mean we use to improve

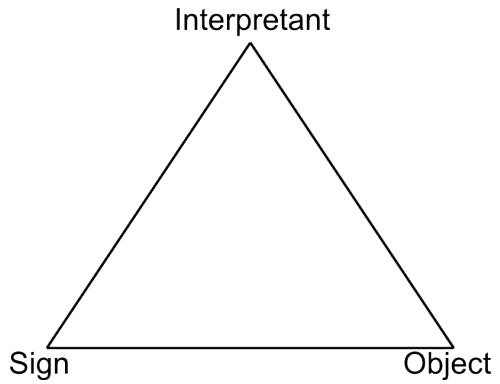


Fig. 1: Peirce's semiotic triangle. It consists of three objects: the sign (i.e. the world as filtered by the sensor), the object (i.e. the physical object), and the interpretant (i.e. the understanding reached by the observer of some sign/object relation).

the human-machine interaction is the fusion, or *synthesis*, of the information contained in a multitemporal SAR series in a unique RGB frame, whose aim is to convey to the user as much information as possible. Therefore, from this viewpoint, the word *synthesis* assumes both its acceptation of *combination of separate things into a complete whole* and of *overviewing* a series of information. This can be useful as a first step for the understanding of dynamic phenomena, segmentation and mining.

These considerations should clarify the necessity to introduce the rather general purpose, user-oriented data representation we are going to define. We used the locution "rather general-purpose" because the quantity (and the quality) of information one can mine from data is always dependent on the type of representation chosen for it. In fact, as argued by Marr, "a *representation* is a formal system for making explicit certain entities of types of information" [20]. In other words, data can be represented in different ways, and the choice made greatly affects the information extraction process, since it makes explicit some information at the expense of other that is pushed back and could be quite hard to recover [20].

Summarizing, in this Section we explained the philosophical background this work refers to. At the end of the work, the reader should be successful in figuring out an operative model that, placing the users at the center of the processing chain, allows for an effective management of several applications, replying to the fifth question introduced in Section I.

III. General workflow and pre-processing chain

The general workflow to generate the proposed products is depicted in the block diagram of Fig. 2. The pre-processing block answers to the second question posed in Section I, and involves all the operations necessary for obtaining a dataset suitable for the multitemporal fusion. This block, exploded in Fig. 3, is inspired from the MAP3 framework introduced in [15], and here briefly recalled.

The input is a stack of SLC images. As first, they are coregistered with standard algorithms. Complex data are used for the extraction of the interferometric coherence; intensity

Fig. 2: Multitemporal syntesis general workflow. Pre-processing is necessary to obtain images spatially and radiometrically comparable, thus suitable for the fusion. Analysis is devoted to the extraction of the temporal features involved into the synthesis.

data are subject to calibration. For more details on this topic the reader should refer to [31] and [32] for COSMO-SkyMed and TerraSAR-X data, respectively.

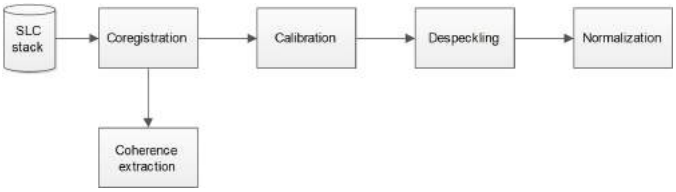


Fig. 3: Exploded view of the pre-processing block. Coregistration, calibration, despeckling, and normalization allows for obtaining a dataset suitable for the multitemporal fusion.

Despeckling is a key passage for making the semantics of the time series emergent. The availability of multiple acquisitions allows for exploiting filtering techniques which use temporal information for reducing speckle with no loss in spatial resolution [33].

Normalization is necessary to reduce images dynamics to its more relevant part (such as natural cycles, agricultural activities, etc.) through a saturation of spikes due to high reflective targets. The variable amplitude levels equalization (VALE) method introduced in [15] is used for this task. The only remark is that here we keep the floating point values of data, postponing the quantization step at the end of the fusion, as detailed in the following sections.

IV. Analysis and synthesis

Reference [2] proposed some goals that the image fusion should achieve for preserving information and enabling interpretability. Here, we want to highlight the four properties of *summarization, consistent rendering, computational ease* and *natural palette*, recalling some concepts expressed in the aforementioned work.

As for the summarization, an effective synthesis should accurately summarize the original dataset, transferring to the user an information he/she can not reach otherwise. This is possible if the rendering of data is consistent, i.e. if the user is able to easily produce the association sign-object or color-feature.

To this end, the natural color palette is of course the best intermediary, but coded techniques for SAR images displaying recalling this representation are not available. Hence, an advance in this sense, with short computational times (consistent with real-time or near real-time applications), is desirable. The products we are going to define aim to provide an

effective answer to these requests by appropriately solving the challenges concerning selection and fusion of the most suitable temporal features. This makes it possible to meet the four proprieties discussed above, and also represents the subject of the third and fourth question introduced in Section I.

A. Multitemporal analysis

In this Section, an answer to the third question raised in Section I will be provided: what is the most successful combination of data in order to obtain a RGB representation meeting the four properties of summarization, consistent rendering, computational ease and natural palette?

The idea is to compare (and combine) the mean intensity values of the time series with some variability indicators. The usage of the variance is a simple choice in order to evaluate the deviation with respect to the mean behaviour.

Another information about the scene dynamics comes from the maximum excursion of the backscattered energy. It allows for the identification of outliers. Therefore, we use as third element of our synthesis the saturation map, as defined in [34], which exploits the information derived by the absolute maxima $\sigma_{\max}$ and minima $\sigma_{\min}$, calculated pixel-wise over the time series as follows:

$$\sigma_s = \frac{\sigma_{\max} - \sigma_{\min}}{\sigma_{\max}}, \quad \sigma_s \in [0,1]\,, \tag{1}$$

The drawback of this quantity is a pdf typically strongly skewed to the right. This results, in the final composition, in a sort of watermark covering the image and corresponding to the color assigned to that band. In order to avoid this visually unpleasant effect and reach a more balanced composition, we consider a saturation index defined by:

$$\hat{\sigma}_s = \frac{\sigma_{\max} - \sigma_{\min}}{\sigma_{\max} + \sigma_{\min}}, \quad \hat{\sigma}_s \in [0,1]\,. \tag{2}$$

This formulation of the saturation index allows for reducing the skewness of the distribution with respect to Equation (1), making it possible to obtain a higher balance of the RGB channels.

B. Multitemporal synthesis

As detailed in the previous Section, the quantities that will be used for the multitemporal synthesis are the mean intensity, the variance, and the saturation index expressed in (2). The answer to the fourth question posed in Section I, i.e. how to fuse the selected data, is now in order.

One of the basic problems in computer vision is to allow the observer to segment the image into meaningful regions [35], preventing the emergence of bright saturated regions, which has been judged as distracting and confusing [2]. Hence, it is fundamental that the visualization favours the pre-attentive processing [36], i.e. the unconscious accumulation of information from the environment. As a consequence, we design our fusion mechanism in order to output highly contrasted images (which makes edge detection easier), with limited occurrences of saturated regions, and a good classes separability.

The fusion is implemented by maximizing the entropy of the channels involved in the composition. In particular, this is performed by clipping recursively the image pdf (at both left and right edges) for different percentages of the cumulative histogram until the maximum entropy value is reached. This procedure allows for obtaining more stretched histograms and a higher contrast. In fact, the flatter the histogram the higher the entropy, see [37] and [15] for more details about entropy maximization in SAR images).

As shown in Fig. 3, the synthesis has a fourth participant, i.e. the interferometric coherence. This quantity is useful for separating high-reflective natural targets from man-made surfaces. In order to insert this information in our RGB composite we proceeded as follows:

a. A master image is fixed, and assumed as reference for the entire time series;
b. The interferometric coherence between the master image and all the slaves is computed;
c. The mean coherence value γ is extracted;
d. The mean coherence map is linearly quantized between a user-defined maximum and minimum. Reasonable values for these parameters are $\in [0.3, 0.4]$ for the minimum and $\in [0.5, 0.6]$ for the maximum. This means that, assuming $\gamma_{\min} = 0.3$ and $\gamma_{\max} = 0.5$ all the pixels whose coherence value is below 0.3 are set to null coherence; pixels with a coherence value in the range $[0.3, 0.5]$ are linearly distributed in the range $[0, 255]$; pixels whose coherence value is above 0.5 are set to 255. In the following we refer to this map as time-series coherence map.
e. The obtained time series coherence map is used in combination with the saturation index map. In particular, we adopted the following rule: if the mean interferometric coherence after quantization is 0 (i.e. if the coherence is below the user-defined threshold) then display the saturation index; otherwise display the mean interferometric coherence.

In the following sections, when the saturation index map is referred, the reader should have in mind the above described modification with respect to its original definition.

As for the combination of the selected bands, we propose a solution aiming to satisfy the requirements of consistent rendering and natural palette introduced in Section IV. Therefore, the products shown in the following sections have been built using the following sequence:

- Red band: time series variance;
- Green band: time series mean;
- Blue band: saturation index map powered with the time series coherence map as explained above.

C. Data

The proposed synthesis algorithm has been tested on four multitemporal series acquired by three different sensors. In particular:

- The first stack is composed by 15 COSMO-SkyMed stripmap images with three meters spatial resolution in azimuth/ground range directions, and acquired over the

TABLE I: Summary of the available data. Resolution is expressed in azimuth/ground range.

Location	# of images	Sensor	Resolution [m]	Polarization/Modality	Period
Campania (Italy)	15	COSMO-SkyMed	3×3	HH/Stripmap	12/2009-10/2011
Calabria (Italy)	35	TerraSAR-X	3×3	HH/Stripmap	4/2008-6/2010
Saxony (Germany)	6	Sentinel-1	20×5	VV/IWS	10/2014-12/2014
Yatenga (Burkina Faso)	21	COSMO-SkyMed	1×1	HH/Spotlight	5/2014-9/2014

Campania region, Italy, between December 2009 and October 2011. Data are HH polarized and shot in ascending orbit;
- The second stack is composed by 35 TerraSAR-X stripmap images with three meters spatial resolution in azimuth/ground range directions, and acquired over the Calabria region, Italy, between April 2008 and June 2010. Data are HH polarized and shot in descending orbit;
- The third stack is composed by six Sentinel-1 interferometric wide swath (IWS) images with 20×5 meters spatial resolution in azimuth/ground range directions, and acquired over Saxony region, Germany, between October 2014 and December 2014. Data are VV polarized and shot in ascending orbit;
- The fourth stack is composed by 21 COSMO-SkyMed spotlight images with one meter spatial resolution in azimuth/ground range directions, and acquired over the Yatenga region, Burkina Faso, between May 2014 and September 2014. Data are HH polarized and shot in descending orbit.

The characteristics of the processed datasets are summarized in TABLE I.

V. Products and their physical interpretation

In this Section, we propose three applicative scenarios for introducing the multitemporal products derived from the datasets described in Section IV-C. In the following, we will adopt the nomenclature introduced in [15], referring to the proposed products as Level-1β products.

A. Scenario 1 - Monitoring seasonal crops in temperate environment with Mediterranean climate

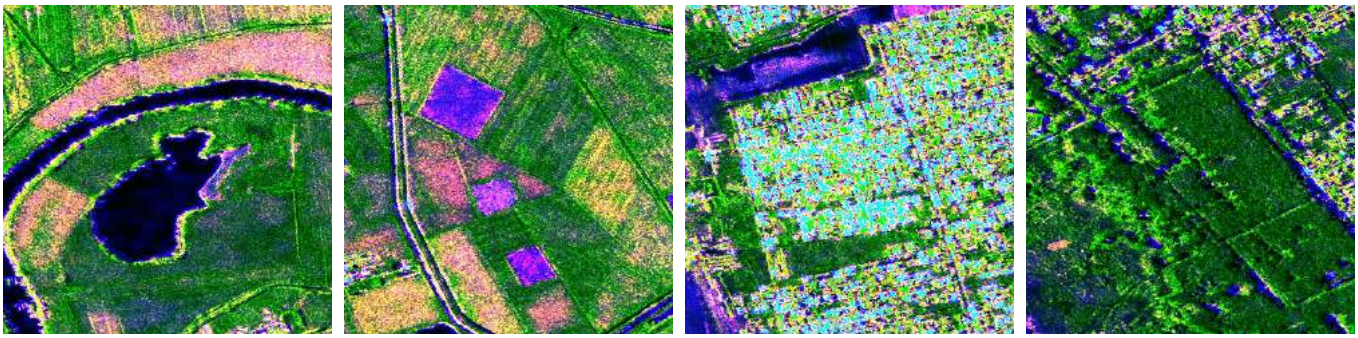
In this scenario, our objective is the detection of summer cultivations using the Campania dataset. To this end, a subset of the available time series concerning the summer season was used. Indeed, since the sowing time is slightly moved up and the harvest could be delayed by some weeks, we considered 9 images between 5 April 2010 and 28 October 2010. The obtained Level-1β product is shown in Fig. 4.

Before examining the monitoring activity, it is worthwhile to linger on this representation, which allows for highlighting some characteristics of the proposed products and the identification of some phenomena, as listed below:

- Sea appears in almost pure blue due to the low contribution of the mean and variance bands, and the high values of the saturation index. This is due to different sea states producing different backscatter in time. Indeed, this behavior (i.e. the blue color) can be observed on other weak scatterers, which could exhibit an unstable response, such as roads, shadows, or surfaces in backslope;
- A more stagnant surface water, typical of rivers and irrigation tanks turns the response of the composition toward the black (see Fig. 5a);
- Some irrigation tanks are subject to a more intensive usage during summer and dry up completely. Hence, their response acquires a strong red (i.e. variance) component which, combined with an increased contribution of the mean band, results in a violet color, as shown in Fig. 5b. In the following we will refer to these structures as "temporary tanks";
- Urban areas appears in cyan due to the combined contribution of the interferometric coherence and of the mean backscattering. The former contribution, in particular, allows for the distinction of man-made targets from the pine-wood which is characterized by a strong response of the mean, as shown in Fig. 5c and Fig. 5d.

Fig. 4: Campania dataset, summery Level-1β product obtained by processing eight images from April to October 2010. Roughly, the following color-object association can be made: blue-sea, black-stagnant water/weak scatterers, green-grasslands/unchanged land cover, cyan-built-up, yellow/pink-growing vegetation. The original image dimension is about 4400×5000 pixels.

As a general comment, a high contribution of the blue band can be found both on very stable targets (such as buildings, due to the contribution of the time series coherence), and on very dynamic objects (such as cropfields or the sea surface, due to the contribution of the saturation index). Indeed these two phenomena, in principle ambiguous if only the blue band



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

Fig. 5: Campania dataset, particulars of the Level-1β product: (a) stagnant water in rivers and in) "permanent" irrigation tanks, (b) "temporary" irrigation tanks, (c) a urban area, and (d) a pine-wood. The original patch dimension is 400×400 pixels.

is considered, are well separable. In fact, on stable (man-made) targets, a negligible contribution of variance is observed. Conversely, a high variance is expected on dynamic natural surfaces.

As argued by Marr, "the usefulness of a representation depends upon how well suited it is to the purpose for which it is used" [20]. When in Section II we referred to Level-1β imagery as a "rather general-purpose data representation", we wanted to remark this aspect. In fact, the proposed composition, besides the advantage for human operators, would have limited applicability, as an example, in sea parameter estimation or ship detection. On the other hand, it is particularly indicated for land monitoring and agricultural applications.

For mining information about cultivated fields, we used as reference the CORINE land cover map [38] (see Fig. 6a). In Fig. 6b, CORINE contours have been overlayed to the Level-1β product. These two representations should highlight the large-scale correspondence between the RGB product semantics and CORINE classes.

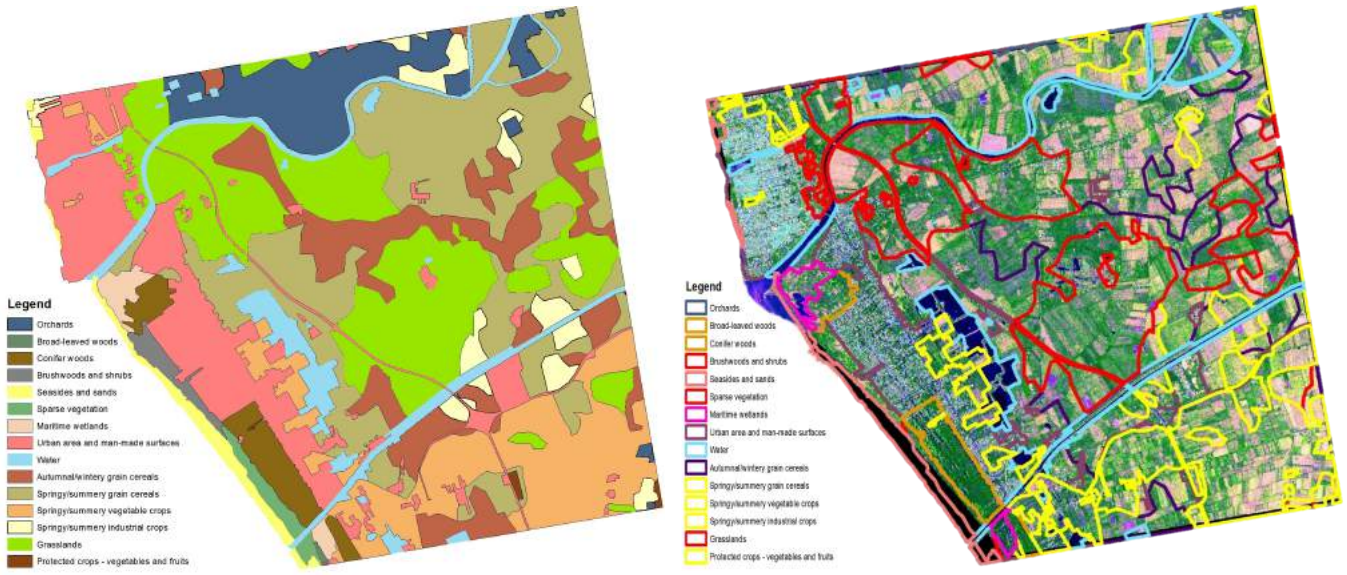
In fact, the reader should easily distinguish, for example, the classes urban area and water on the Level-1β product. Their topology is respected compared to the CORINE map. The pine-wood at the bottom left corner of the scene, as well as the stripe of sparse vegetation and brushwoods close to the seaside are in the same way identifiable. Actually, the classes brushwoods and sparse vegetation are indistinguishable on the Level-1β product, but they form a whole close to the coast that is clearly separated from the adjacent urban area and conifer woods.

It is worthwhile to remark that the product we are analyzing has been produced considering a series of images belonging to summer. Hence, seasonal cultivations are expected to have high values of the mean intensity and variance bands due to the enhancement of backscattering derived by the volumetric scattering contribution of plants growth [39]. Indeed, the saturation index is also expected to be high due to harvesting. Therefore, summery cultivations exhibit high values in all the bands of the composition, but their balance is dictated by cultivation type (determining the harvesting time), terrain roughness, foliage density, plants height and fruits dimensions. This means that there is no unique colour association that identify all agricultural fields. However, given that the variance contribution is significant, a higher contribution of the mean intensity (green band) results into a yellow color. A dominance of the saturation index (blue band) turns the response of the composition into a pink tonality. As an example, if the field appears in yellow, a high foliage density or taller plants have to be expected, as in the case of orchards (see the upper part of thematic map of Fig. 6a).

Coming back to the overlay depicted in Fig. 6b, the reader should note how the contour color for the classes "Springy/summery grain cereals", "Springy/summery vegetable crops", "Springy/summery industrial crops" and "Protected crops - vegetables and fruits" is the same (yellow), i.e. these categories have been grouped in a macro-class. This helps to highlight that the Level-1β product is congruent on average with the CORINE thematic map. In fact, fields whose response turn into pink or yellow are enclosed in the contour of that macro-class.

The classes "Autumnal/wintery grain cereals" and "Grasslands", identified in Fig. 6b by red and amethyst contours can be also fused in an unique class. In fact, grasslands are objects almost stable and therefore their response is dominated by the mean intensity band (green). During summer, winter-cultivated fields are usually destined to fallow; hence, they are most likely covered by vegetation, and their response turns into green, too. In our case, grasslands and wintery cultivations are grouped in the center of the scene (see the thematic map of Fig. 6a) and this is confirmed by the Level-1β product which, in that zone, exhibits a large dominance of the mean intensity band response.

The above listed outcomes are sensor independent. In Fig. 7a and Fig. 7b we show two Level-1β products relevant to the city of Cirò Marina, Calabria, (southern) Italy. In both cases, the SLC images were acquired by the sensor TerraSAR-X. In particular, the product depicted in Fig. 7a has been obtained by fusing a time series acquired between April-October 2008. The product depicted in Fig. 7b is the result of the fusion of a time series acquired in the same period of the year 2009. In order to make the two time series comparable, a metric was established. Therefore, we assumed as reference the series relevant to the year 2008, fixing the maximum and



(a) (b)

Fig. 6: Campania dataset: (a) CORINE land cover map and (b) CORINE land cover contour overlayed to the Level-1β product. This representation should allow the reader for appreciating the large scale correspondence between Level-1β colors and CORINE classes.

minimum values of the quantities involved in the synthesis of the Level-1β product. In other words, when the byte-scaling of the 2009 series is performed, the zeroth and 255th amplitude levels coincide for each band with those of the reference series.

The reader can appreciate how the semantics of these compositions is consistent with that illustrated above for the Campania dataset. The principal cultivation of the study area is the grapevine, which is a summer cultivation. It is easily identifiable in the coastal stripe of Fig. 7a with the fields with a yellow response. Moreover, due to the calibration, variations in the colorimetric response between the two products can be exploited for studying the behavior of the cultivations during the two farming seasons, as explained in Section VI-A.

B. Scenario 2: A first look on Sentinel-1 data in regions with continental climate

Sentinel-1 started to acquire data on April 2014. The success of the mission requires the development of repeatable and reliable processing to produce attractive products for the end-user community. The potential of the mission has been widely explored in the remote sensing literature [40], also through the use of simulated data [41]. Here, we will show that Level-1β imagery is fully compatible with the characteristics of the new sensor.

To prove this claim, we processed a multitemporal series of six images acquired between 3 October 2014 and 2 December 2014 in IWS mode over the Saxony region (Germany). Preliminary processing, such as TOPS [42] deburst and calibration, has been performed using the ESA Sentinel-1 Toolbox software.

In Fig. 8 we show a detail of the acquisition made on 3 October 2014, and relevant to the city of Dresden. The image has been processed up to 15 equivalent number of looks using multilooking with factor 1×4 in azimuth/ground range (to make the pixel square), and multitemporal despeckling. It has a spatial resolution of about 15 meters. The whole product has a coverage of about 180×250 kilometers and can be handled in stack with the other images of the time series by the algorithms described in Section IV and in Section IV-A with a processing time of about one hour on a four cores, 12 GB RAM machine.

In Fig. 9 we show the full Level-1β product of the study area. The spatial resolution is 15 meters. The yellow dot on the map identifies the city of Dresden. In the same image, the relevant CORINE land cover portion is also reported.

Macroscopically, the reader should note that the colors restituted by the composition are constantly associated to the same image feature. Cities are rendered in cyan. In the middle of the image, at South of Dresden, a wide stripe of woods is identifiable, due to the dominant green component. Looking at the CORINE land cover, in the top-left corner of Fig. 9, the dominant classes in this area are “Broad-leaved forest” and “Conifer Forest”, rendered in green and dark green, respectively.

At East and West of the city of Dresden (see the light yellow class on the CORINE land cover), croplands are distinguishable, as well as at South, above the woodland stripe. At North-West of Dresden, instead, another wide woody area can be identified with some water bodies of different size.

Summarizing, the proposed Sentinel-1 Level-1β product respects macroscopically the semantic indicated by the CORINE land cover.

The same reasoning can be repeated on a finer scale considering the close up of the Dresden city depicted in Fig. 10. The relevant CORINE land cover is displayed in the bottom-

(a) (b)

Fig. 7: Calabria dataset: (a) April-October 2008 and (b) April-October 2009 Level-1β products. The semantic exhibited by these compositions (i.e. the association color-object) is consistent with that of the Campania dataset previously analyzed. The original dimension of images is about 3200×2300 pixels.

Fig. 8: A fragment of the Sentinel-1 acquisition of the Dresden city. The image has been processed up to 15 ENL using spatial multilooking and multitemporal despeckling. It has a spatial resolution of about 15 meters. The original patch size is about 1000×1700 pixels.

left corner of the image. This representation should allow the reader to better appreciate the matching between the semantics extractable from the Level-1β product and that established by the CORINE land cover.

As an example, it is easy to link the green-dominant response in the area at North-West of the Elbe river with the conifer forest indicated in dark green in the CORINE map. In the same way, the Dresden city park, rendered in light pink on the reference land cover, is clearly distinguishable in the Level-1β image thanks to its green response.

In synthesis, based on the discussion provided in Section V-A about the Campania dataset, the reader should be able to reconstruct the semantics of the Level-1β response of the city of Dresden. This testifies the robustness of our framework in the color-object association (or sign-object in the Peirce view), and its suitability to Sentinel-1 imagery. For a quantitative analysis of this dataset, the reader is addressed to Section VI-B.

C. Scenario 3: Level-1β products in semi-arid environment

In this section, we discuss the reliability of the proposed framework when the climatic zone is changed. In this case, we deal with a semi-arid environment, where agricultural activities are strongly related to the cycle of rainfalls [43], determining the effectiveness of seasonal cultivations and the quantity of water which can be harvested for facing the dry season [44].

In Fig. 11 we show a Level-1β product depicting the Laaba basin. It has been obtained by processing eight spotlight images acquired between 3 July 2014 and 4 August 2014, i.e. in the middle of the rainy season [43]. In this scenario, it is of interest to monitor water and cultivation dynamics, due to their importance for local community wellness, as well as the presence of small human settlements, since updated maps are not always available.

Level-1β imagery provides useful information regarding all these activities. As an example, the Laaba basin is clearly distinguishable at the center of the image. As stated above, it is highly influenced by rainfall cycles. In fact, it dries up completely during the dry season due to evapotranspiration and

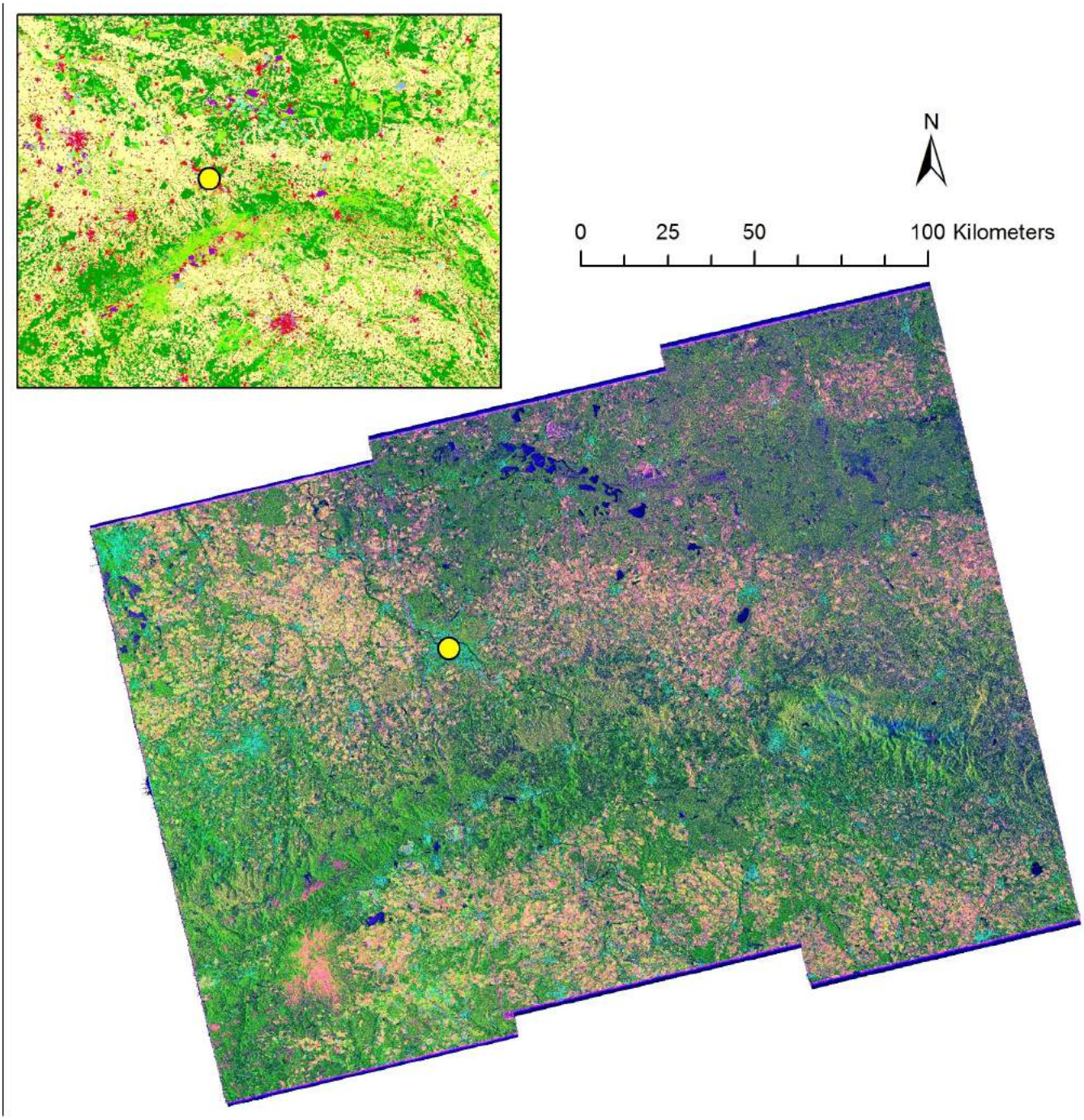


Fig. 9: Sentinel-1 full Saxony Level-1β product with 15 meters spatial resolution. The yellow dot identifies the city of Dresden. The relevant CORINE land cover is reported at the top left corner of the image in order to show, macroscopically, the correspondence between the Level-1β color meaning and the classes of the reference map.

human consumption, and it is filled by rain water during the wet season [44]. This behavior is almost constant in different years. However, basin contours are strongly variable even during the wet season, since the available water is used for irrigation and human activity. This emerges when analyzing the Level-1β product in the basin area. In fact, in its bottom part, regions characterized by a pink response are present (see also Fig. 12a), due to the combination of variance and saturation index. Physically, it is possible to associate this response to areas characterized by an unstable water coverage.

The same phenomenon characterizes bright regions within the basin area and on its contour. However, in this case, the high contribution of the mean indicates that the terrain is not covered by water in the most part of the considered acquisitions. Therefore, if a mask of the maximum basin extension is available, Level-1β imagery allows for the extraction of information about the water coverage extent in the considered time interval.

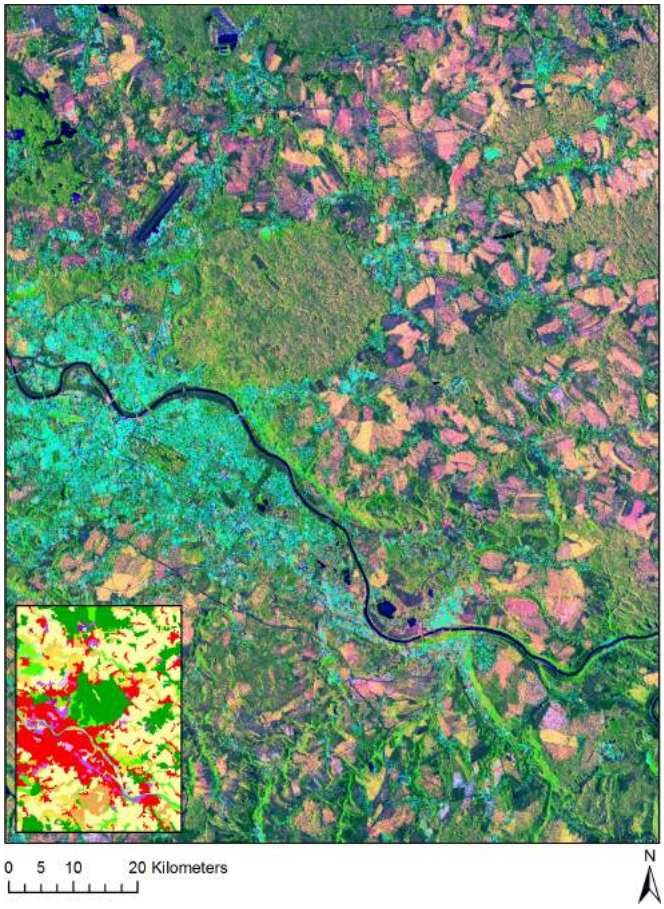


Fig. 10: Sentinel-1 Level-1β product of the Dresden city area with its relevant CORINE land cover map.

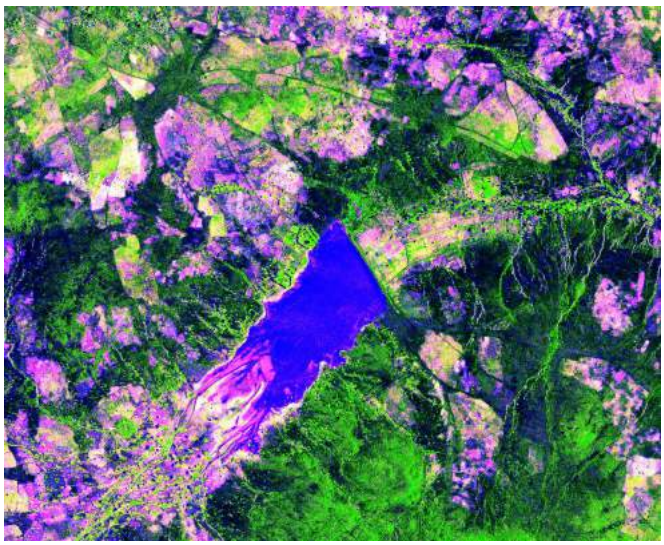


Fig. 11: Yatenga dataset (Burkina Faso): spotlight COSMO-SkyMed Level-1β product obtained by fusion of eight images acquired between July and August 2014. The blue area in the middle of the image identifies the Laaba basin. The original image dimension is about 4700×4900 pixels.

In Fig. 12b and Fig. 12c, the reader should note how both the change in acquisition modality and climatic area do not affect the response of features such as cultivated fields (Fig. 12b) and "urban areas" (Fig. 12c). We used quotation marks since in this case we are considering very small settlements, composed by few constructions behaving as stable targets.

In Fig. 12d, we show a detail relevant to a feature characteristic of a semi-arid environment, that is an eroded area [45], i.e. a region which has lost its capability to retain and absorb water. For this reason, it is completely barren and characterized by a low reflectivity. Thus, the response of Level-1β imagery for this feature turns out toward the black or a very dark blue because of the contribution of the saturation index.

It is worthwhile to remark that, dealing with semi-arid environments, in which the imaged scenes are mostly natural and, during the rainy season, characterized by a strong dynamics, the temporal sampling of the images involved in the Level-1β composition is very important, as well as the considered time span. In fact, a thin set of images or a long time interval could give rise to anomalous response due to the effects of variance and/or of the saturation index.

VI. Applications

In this Section, we explore the suitability of the proposed products to practical scenarios. In fact, visual data mining and interpretation are not enough to extract information at large scale from satellite images. As a consequence, it is highly desirable that such an easy-to-read data representation can be effectively processed with standard algorithms.

It is quite clear that a more interpretable product is very well suited for supervised classification [46]. However, the recent remote-sensing literature paid great attention to neural network techniques because they can handle effectively large multidimensional datasets. Therefore, in the following, we will linger on the usage of such techniques, with particular reference to one of the most consolidated and widespread, i.e. the Kohonen's self-organizing map (SOM) [47]. In particular, we will show that, for the considered applications, Level-1β products convey most of the information, so that useful clustering results are obtained.

A. Scenario 1: Monitoring two-year agricultural activities

In this Section, we face the problem of information mining from the Calabria scene introduced in Section V-A. We considered images acquired in two successive years (2008 and 2009) for generating two calibrated Level-1β products representative of the two summer seasons. Operatively, these representations can be exploited for making qualitative evaluations about the status of cultivations using SOMs.

1) Objective and approach: SOMs, also known as Kohonen Maps, are a machine learning technique of the artificial neural networks (ANN) family. They are used for the classification of the most diverse data types. Their application, since their introduction by Kohonen [47], has been experimented in different sectors, like gene expression analysis [48], financial diagnosis [49], synoptic climatology [50], microbial community dynamics [51], bankrupt prediction [52] and political science [53]. SOMs have been widely applied to remote sensing since the

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

Fig. 12: Fragments of the Yatenga image displayed in Fig. 11: (a) an area characterized by instable water coverage, (b) crops in the nearby of the dam, (c) a small human settlement, and (d) an eroded area. The original patch dimension is of 400×400 pixels.

'90s [54] as well, and new applications are still studied and proposed today.

This widespread use of the SOMs is due to the extreme flexibility of the tool. Compared to other classical statistical methods, ANN do not make assumptions on the statistical distribution of the data. They can be easily applied to large datasets and modified and/or integrated to be adapted to different data structures [55] and learning techniques [56]. The robustness to large amounts of data makes them a suitable instrument for unsupervised or semi-supervised classification in a big data scenario, which is, and will increasingly more be in the future, a key issue in remote sensing. SOMs proved to be an outstanding method for unsupervised classification of multispectral images, because of their dimensionality reduction capacity [56], [57].

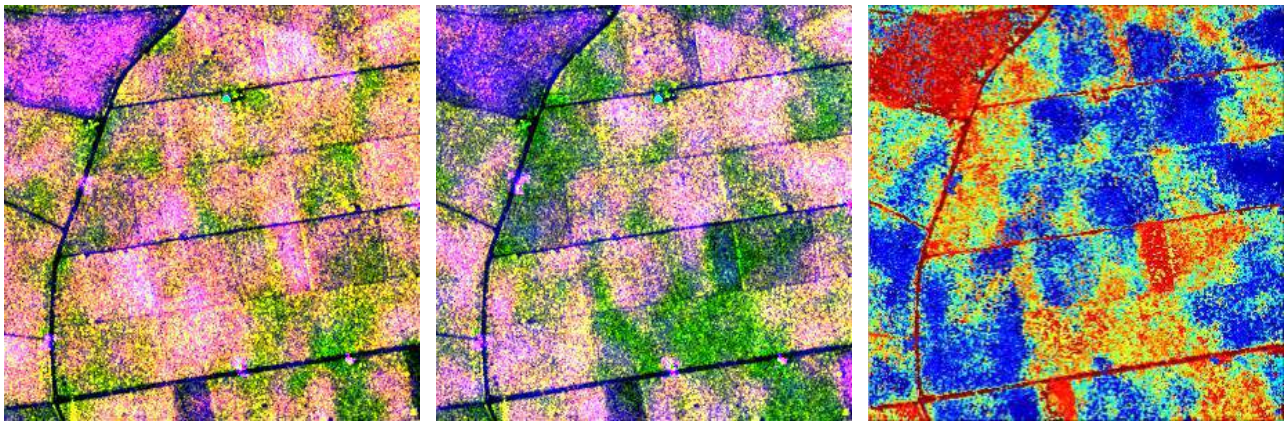
Dealing with SAR data, SOMs have been employed, for example, for polarimetric data classification [58], change detection [59] or for flood mapping [60]. Here, we want to evaluate the cropfields behavior in the two considered years running a SOM on the stack composed by the two Level-1β products. The network is set up with nodes containing together the RGB values for the two considered years and is initialized and trained with couples of RGB triplets selected randomly from the two Level-1β products to classify. Presenting the two year values together allows the network not only to classify the images based on the characteristics of the pixels, but also on their relative change during the two years. As a consequence, the changes in the two following years detected by the SOM can be interpreted as changes in the crops behavior or production.

This application is critical since it involves two different characteristic times relevant to the synthesis of the single multitemporal products and to the two-years cropfields behavior classification. This forced us to consider a strategy in which the SOM clustering was supported by hierarchical agglomeration and object-based reasoning. This allowed for reconstructing a more homogeneous cluster map and a better understanding of the study area. In particular, we used an agglomerative hierarchical clustering method to unify similar classes after classifying the input Level-1β products in 64 categories through an overdimensioned 8×8 SOM [61]. Therefore, a top-down model was followed [18]. In fact, our world model suggests that 64 categories are too much to describe the dynamics of our scene. As a consequence, the initial number of clusters will be reduced according to the characteristics of the scene and to the application we are dealing with, reducing the clusters fragmentation at the same time.

2) SOM principles and application: A SOM is a network composed by a predefined number of nodes, connected with a usually hexagonal or rectangular structure, like in our case. The nodes are elements of the same typology of the training elements. Each time a training element is presented to the network, the most similar node is detected and identified as Best Matching Unit (BMU). The BMU and its neighborhood, defined by a radius (which decreases as the training epoch increases), are updated to become more similar to the presented training element. Updating the neighbourhood of the BMU allows the SOM to maintain topological properties of the training data, i.e. to maintain topologically close on the network the identified classes that have similar characteristics [47]. At the end of the training phase, the obtained nodes are used to classify each element of the data.

Several parameters can be adjusted to change the behaviour of the SOM classification. Usually, a Gaussian neighborhood function is used, so that the strength of the adaption to the presented training vector decays exponentially around the BMU [62]. The number of training elements and the number of training epochs are adjusted in order to obtain a stable result, i.e. to obtain similar resulting patterns with multiple repetitions of the training and classification algorithm. In fact, different realizations of the Kohonen's algorithm with the selected parameters do not show exactly the same topology, because of the random initialization and training set, but show the same final patterns. Here, we used 1000 training elements and 200 epochs.

In Fig. 13a and Fig. 13b, the two subsets of the Level-1β products relevant to the year 2008 and 2009 are depicted. SOM clusters are reported in Fig. 13c. In Fig. 14, a realization of the 8×8 Kohonen network is provided. Each element is



(a) (b) (c)

Fig. 13: Calabria dataset subset: Level-1β products relevant with years (a) 2008 and (b) 2009. (c) Clustered product. The 64 clusters identified by the SOM depicted in Fig. 14 (counted from top to bottom and then from left to right) are associated to 64 different colors. The original patch size is 400×400 pixels.

represented by a plot showing the change of RGB values in the 2008 and 2009 years.

The clusters map depicted in Fig. 13c, although oversegmented, has a smooth representation because at neighbouring pixels, having a similar behavior, are associated neighbouring classes, thus similar colors. In fact, observing the Kohonen network of Fig. 14, the reader can see that adjacent nodes represent similar patterns in the pixel behavior between the two years.

3) Agglomerative hierarchical clustering and object-based reasoning: In order to reduce the map fragmentation, we performed an agglomeration of the obtained clusters to fuse those carrying similar information.

A dendrogram represents a quick solution for evaluating the relation between SOM clusters (see Fig. 15). Here, we adopted as metric the Ward's distance [63]. The red dashed line represents the distance under which dendrogram's branches are considered for fusion. Clusters interested by fusion have been displayed in yellow, green, magenta, blue and red. This representation is consistent with the SOM grid depicted in Fig. 14, in which the colored rectangles enclose fused clusters, for a total of five categories. Looking at the dendrogram, the output of the fusion process is given by the intersection between the horizontal red dashed line and the black leaves above the colored groups of clusters.

The output of the hierarchical clustering is depicted in Fig. 16a, in which an unpleasant granularity arises. It consists in small regions mainly sunk into an uniform background. In this picture, the cluster-color association is consistent with the dendrogram of Fig. 15 and with the SOM grid of Fig. 14, in which the semi-transparent colored rectangles enclose the agglomerated clusters.

Physical-based consideration and object-based reasoning can be used to improve the quality of the obtained clusters map.

We propose a simple processing based on connected components labeling [64]. This algorithm assigns an increasing numeric index to each connected region found within the image. For these objects, as suggested in [65], some shape parameters can be computed, as well as spatial relations between them. Here, we propose to compute some parameters, such as area, compactness, number of neighbours and percentage of shared borders between adjacent regions. These parameters were used to generate an appropriate physical-based and application-tailored set of rules. In this case, we want to reconstruct the homogeneity of the clusters representing the behavior of the cropfields. Therefore, the implemented rules aim at fusing small objects (also with irregular shape) to the background.

In Fig. 16b, the result of the object-based image analysis is reported. It is quite evident that the fusion operation has definitely improved the quality of the cluster map. In fact, with respect to the map presented in Fig. 16a, the number of regions is dramatically reduced, decreasing from 12529 to 207, also bringing undebatable benefits to the physical interpretation of the map.

4) Interpretation: The physical interpretation of the retrieved clusters is now in order.

From the available data extracted from the CORINE land cover [38], it appears that the study area is mainly destined to vineyards. Thus, observing the Level-1β subsets depicted in Fig. 13a and Fig. 13b, the clustered image reported in Fig. 16b, and the SOM grid of Fig. 14, it is possible to infer the following:

- Yellow cluster: The RGB values in this cluster are medium to high, constant between the two years or with a slightly increasing red value. The red (variance) and the green (mean) values are generally higher than the blue (saturation index). This means that the yellow cluster is associated with agricultural areas that have a pink/yellow color in the Level-1β product in 2008 and remained similar in 2009, or slightly turned to pink when they were yellow in the first year. These areas exhibit an almost constant behavior in the two examined years. Assuming that the study area is destined to vineyard, as

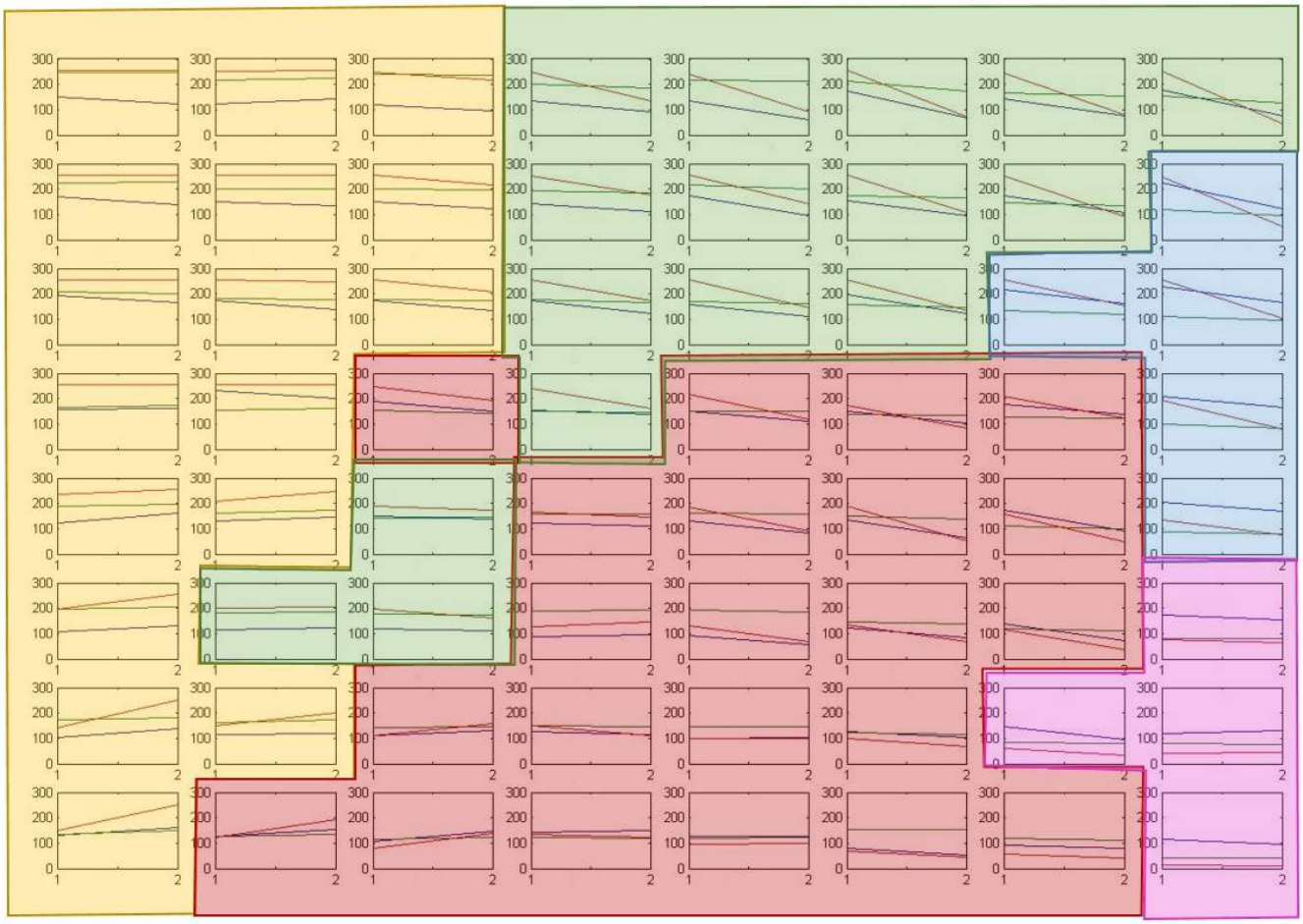


Fig. 14: 8×8 SOM grid showing the RGB values (reported on the y-axis) for year 1 and their change to year 2 (see x-axis). The colored rectangles represent the clusters after hierarchical agglomeration (see Section VI-A3 and the dendrogram in Fig. 15).

stated in the CORINE map, we can argue that this cluster represent the fields in which the optimum behavior of the cultivation is reached. In fact, vineyards experience their full development between April and October, which is actually the period considered for building the Level-1β products used for the classification (see Section V-A). Therefore, in this period the growth of the cultivation results in high values of the considered multitemporal quantities;

- Green cluster: In this cluster the green value is medium and remains constant, but red and blue values are generally decreasing from high to low. This is associated with areas that were pink in the 2008 Level-1β product and turned into green. Physically, the decrease in variance and saturation index could be related to a smaller development of the cultivation, thus to a reduction of the volumetric scattering contribution. This can be associated to a lower development of the cultivation in the year 2009. However, since the contribution of the mean remains high in both years, despite the decreasing in variance and saturation index, it is possible to argue that the cultivation is structurally characterized by a rather dense foliage, such as that of a vineyard, whose structure and dynamics are fully compatible with the phenomenology inferable from the cluster analysis;
- Red cluster: The RGB values are constant or slightly decreasing, but generally lower than the values in the Yellow cluster. This cluster is mainly associated with green areas which remains almost constant in the two years. From the above considerations, we can infer that the behavior of the vineyards grouped in this cluster is not optimal in both the considered years;
- Blue cluster: Areas in this cluster have high blue and red values in the first year, decreasing in the second year, while the green remains constant. The red value generally decreases more than the blue one. This cluster is associated with areas that were pink in the 2008 Level-1β image and turned into blue in 2009. It is mainly located in the field at the top-left corner in Fig. 13a, Fig. 13a and Fig. 16b. It represents an anomaly with respect to the behavior of the previously analyzed cluster. In fact, the abrupt fall in variance and the low values of the mean bring us to argue that this cluster is a fallow land on which some activities have been performed in the year 2008;
- Magenta cluster: these areas have very low values of RGB. They are almost constant in the two years. This is associated with dark blue areas in the Level-1β products. The very small variance indicates that the pixels are quite

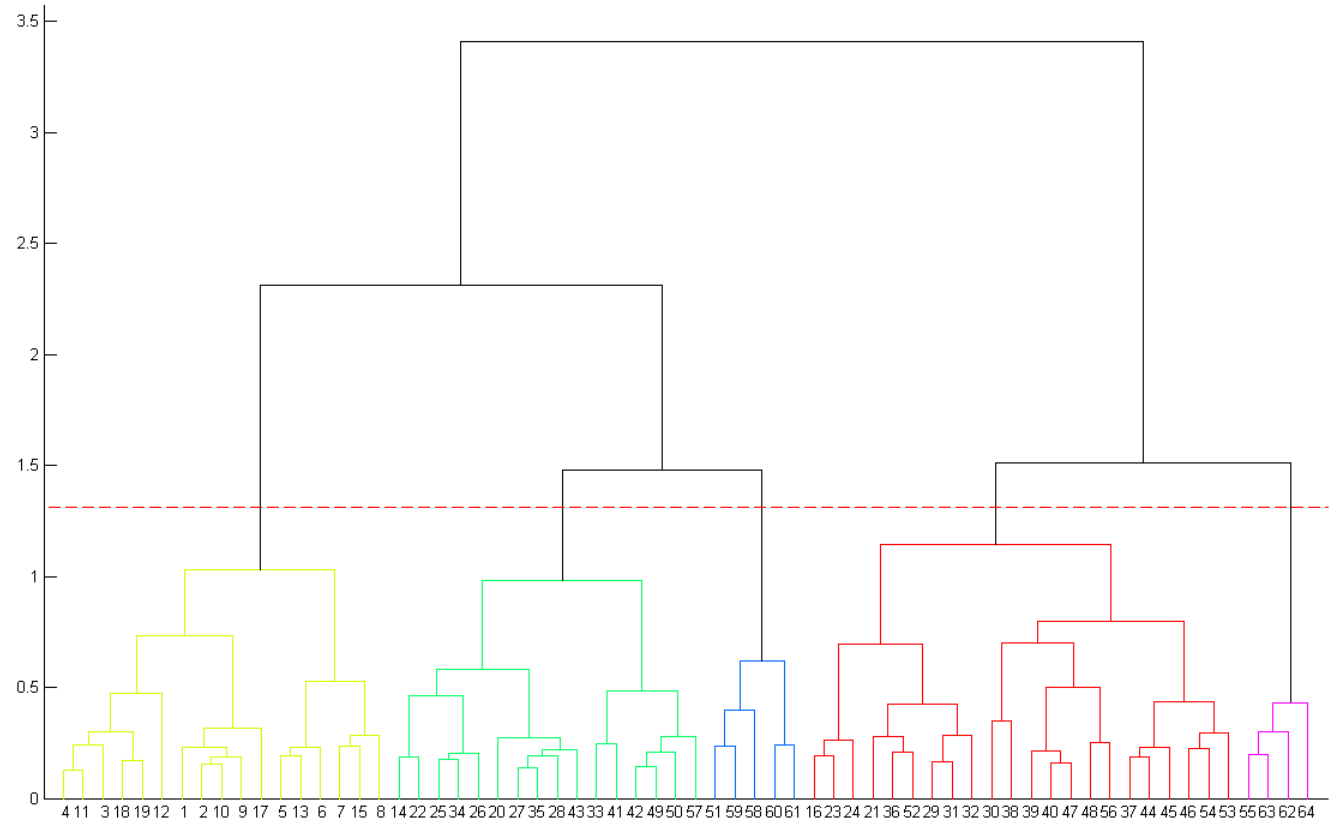
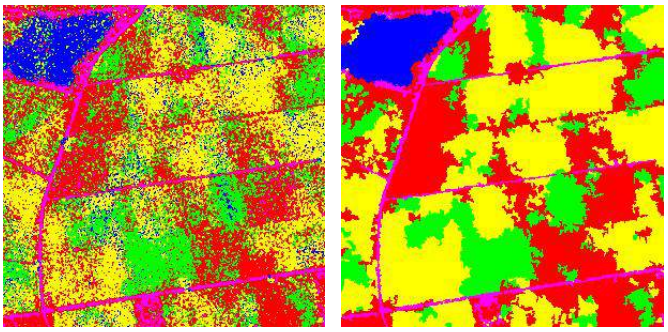


Fig. 15: SOM cluster dendrogram. The red dashed line represents the distance under which diagram's branches are considered for fusion. Clusters interested by the fusion are depicted in yellow, green, blue, red and magenta (see also the SOM grid depicted in Fig. 14).



(a) (b)

Fig. 16: Clusters map (a) after the hierarchical agglomeration and (b) after the object-based region fusion. The cluster-color association is consistent the dendrogram of Fig. 15 and with the SOM grid of Fig. 14.

stable. In fact, looking at the shapes, it is clear that they are associated to roads.

The outcomes of clusters interpretation are summarized in TABLE II. In this table, RGB attributes refer to the 2008 Level-1β product and their derivatives are qualitatively evaluated based on the SOM grid depicted in Fig. 14. However, we want to remark that this analysis is inferred by available data, but no ground truth for this experiment is available.

B. Scenario 2: Sentinel-1 land cover mapping

In the previous Section, we used Level-1β products and SOMs for solving a critical problem, i.e. the qualitative evaluation of the agricultural production in two successive years. Here, we face a more classical problem, such the land cover mapping derived from multitemporal data [34].

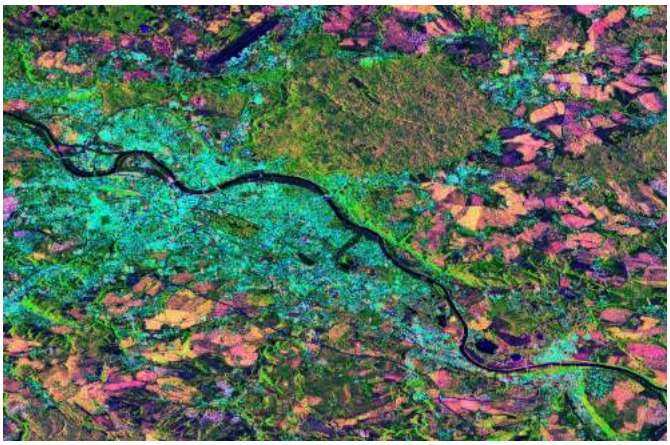
In this case, SOMs offer the possibility of an immediate semantic transferring from the Level-1β product to the classified map. In fact, during the training, the SOM (randomly initialized) nodes are updated to be representative of the training elements, which are chosen within the RGB product. Thus, the resulting node colors will have the same semantics of the input Level-1β products.

We performed a land cover classification on a fragment of the Sentinel-1 Level-1β product presented in Section V-B relevant to the Dresden city area. Unlike the technique presented in [66], here we used a 2×2 SOM grid for producing a 4-class land cover product. The SOM training parameters were the same used for the Calabria experiment discussed in Section VI-A. The original Level-1β product and the SOM land cover classification are shown in Fig. 17a and Fig. 17b, respectively.

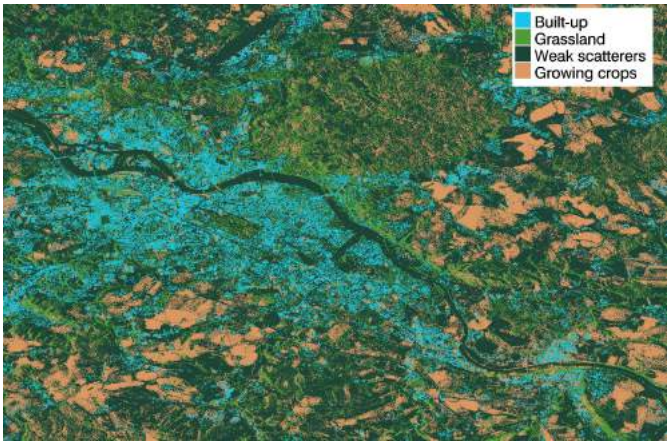
As stated above, the use of a SOM allows for transferring immediately the semantics from one product to another. In

TABLE II: Summary of cluster interpretation outcomes. RGB attributes are referred to the 2008 Level-1β product. In table headings, t refers to time.

Cluster color	R	G	B	R(t)	G(t)	B(t)	Interpretation
Yellow	High	High	High	$=$	$=$	$=$	Optimal crops behavior
Green	High	Medium/High	Medium	$\downarrow\downarrow$	$=$	$\downarrow$	Non-optimal behavior in year 2
Red	Medium	Medium	Medium/Low	$\downarrow$	$=$	$=$	Non-optimal behavior in both years
Blu	High	Low	High	$\downarrow\downarrow$	$=$	$\downarrow$	Fallow lands
Magenta	Low	Low	Low	$=$	$=$	$=$	Roads/Low scatterers



(a)



(b)

Fig. 17: Dresden area: (a) Sentinel-1 Level-1β product and (b) its 4-classe land cover map. The use of SOM allows for obtaining a classification map whose appearance is very similar to the input RGB products. This way, an immediate semantic transfer between the two products is possible. The original images dimension is about 1500×2300 pixels.

particular, the interpretation of the class meaning, following the guidelines given in Section V, is provided below:

- Dodger Blue class - Urban area: This class corresponds to cyan areas in the Level-1β product due to the high contribution of the mean and of the interferometric coherence;
- Light green class - Grasslands: It corresponds to light green areas in the Level-1β product, showing low variance and saturation index and a higher contribution of the mean due to the backscattering contribution of terrains;
- Orange class - Growing crops/vegetation: This category includes areas characterized by tones from yellow to pink in the Level-1β product due to a medium/high contribution of all the considered indicators;
- Dark green class - Low scatterers: This is the more heterogeneous class, enclosing the darker objects of the scene (water bodies, bare soils and shadows).

This experiment outputs a very heterogeneous cluster (the Dark green one) enclosing objects of different nature. Thus, with respect to the Calabria scene analyzed in Section VI-A, the problem is inverted, since we may be interested in splitting this cluster in at least two more significant categories.

To this end, we analyze the reasons that led to the association of black and dark green objects in the Level-1β product within a unique cluster.

As stated above, the SOM nodes are trained with RGB triplets randomly selected within the input product. However, looking at Fig. 17b, it arises that the "black object class" is the less represented in the data. Therefore, it is highly uncertain that a SOM node can gain such color, since few training sets relevant to this class are presented to the network.

The flexibility of a SOM, joined to a knowledge about the mapping of the real world into Level-1β imagery [18], allows us to easily solve this problem. In fact, in a Level-1β product, an insufficient presence of an object class can occur mainly for water surfaces (as in the analyzed case) and urban areas, if an extended natural scene is considered. In the presented experiment, this caused substantially the aggregation of the class water bodies into the class woods, which is the closer for chromatic characteristics.

In order to split this class into its two major features, it is sufficient to force a significant number of training sets (in the order of 15-20% of the total) to point toward the less represented class. In such way, we induced one of the SOM nodes to assume the corresponding color. In this case, we used a 2x1 SOM, acting the modification of the training sets toward the black color after masking out all the classes except the Dark green one. It is worthwhile to note that when the dimension of the SOM is reduced, the neighbor influenced by the winner neuron should be modified accordingly.

In Fig. 18, we show the final 5-class land cover map, after the application of the above described splitting of the Dark Green class into a Dark Green cluster (again) mainly representative of woods and a Black cluster relevant to water

surfaces. It is remarkable how the land cover map looks very similar to the input Level-1β product.

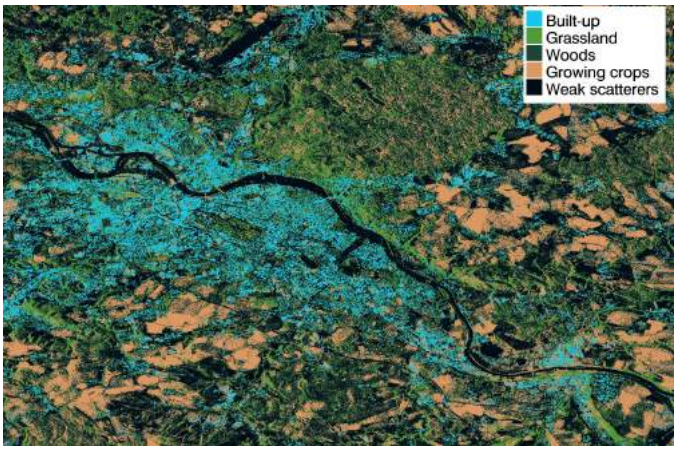


Fig. 18: Dresden area, Sentinel-1 5-classe land cover map after the splitting of the Dark Green class.

However, the Black cluster does not include only water surfaces, since other features, such as shadows and bare soils, are chromatically similar to them. This allows for enforcing the consideration made at Section I about the extraction of information from remote-sensing data. SOMs, although representing, as demonstrated, a very flexible and suitable tool, can provide only a partial answer to this problem. In fact, they act on a single image attribute, i.e. on its chromatic characteristics. An object-based reasoning, as discussed in Section VI-A3 can improve the effectiveness of a SOM-based analysis, but a complete image understanding system can not overlook the integration of image analysis techniques (such as SOMs) and scattering models. In fact, the separation of water surfaces from backslope areas or bare soils can be obtained by analyzing the electromagnetic scattering of these features [67], [68]. In the same way, as an example, the characteristics of the urban area can be retrieved only by evaluating the complex scattering mechanisms triggered by an urban environment [69], [70]. We think that all the ambiguity derived by the SAR scattering mechanism should be resolved at model level. In fact, however user-friendly may be this or that technique, including also commercial suites, it can not solve problems as those discussed above without taking into account scattering models. Thus, the role of the SAR experts should consist in the integration of electromagnetic models into user-oriented data analysis techniques for making them fully effective in the solution of disambiguation problems and, at the same time, manageable with the lowest level of expertise by the end-user community.

VII. Conclusions

In this paper, we introduced a new class of multitemporal SAR products. The proposed Level-1β products have been designed to help the human photo-interpreter, thanks to a rendering as much as possible consistent with human visual understanding. The aim of the proposed products is to lower the expertise required to correctly interpret the data. In fact, in a scenario in which the automatic extraction of information from large databases is rather limited, users are in many cases appointed to extract information from the analyzed scene, i.e. to transfer a semantics to the imaged objects. Level-1β products comply with this necessity, which is more and more felt in the end-users community, making the human-machine interaction easier and more comfortable.

We tested the reliability of our products by implementing the proposed processing chain using images acquired by three different sensors on four scenes (Campania, Calabria, Yatenga and Saxony). All the available acquisition modalities (stripmap, spotlight and scan) were tested, as well as different climatic zones (temperate with Mediterranean/continental climate and semi-arid). The obtained results confirmed the semantic stability of the proposed products, i.e. the consistency of the association between the displayed colors and the objects on the scene.

In Section VI we tested the minability of the proposed products and their suitability for data analysis techniques. The application of Self-Organizing Maps was selected for the robustness to big data, flexibility, and the capacity to maintain the chromatic semantics of the input image. In the first application, we demonstrated the suitability of the products and of the tool to convey a complex information, i.e. the qualitative evaluation of changes in the agricultural production of two consecutive years. In the second example, we tested the capability of the Self-Organizing Map to classify a Sentinel-1 composite relevant to the city of Dresden, maintaining the semantics associated to the color response of the input Level-1β product. In both applications, the combination of the Level-1β images and self-orgainizing maps performed particularly well and demonstrated the suitability of the product to data analysis techniques.

In Section I, we recalled five questions proposed in the past literature concerning the reliability of a data fusion technique. We proposed an answer for all of them with regard to the problem of data representation. However, question 5 (concerning how to extract the enormous amount of information from remotely sensed data) is only partially answered. It could not be otherwise, considering that this work mainly deals with a part of the whole problem, i.e. data representation. Here, we can linger on the following consideration: what is the role of the electromagnetic models in remote sensing data analysis?

At first, electromagnetic models constitute the basis of the interpretation of any SAR product. In fact, unlike the case of optical data, humans have no experience of radar imaging directly linked to the world they live, and from which they can take inspiration for understanding data. Therefore, Level-1β imagery gains its semantics from the knowledge of the phenomenology dictated by the interaction of the electromagnetic fields with Earth surfaces. The proposed products subtend the expertise required for understanding these complex mechanisms, re-elaborating the scene dynamic in a more user-friendly color display in which the non-expert user can more easily encounter his/her expectation about object appearance. However, models are often left aside in favour of the development of techniques, which are highly conditioned by the application and by the correct selection of free parameters.

Actually, we think that a data analysis technique can be a valid alternative to the application of electromagnetic models, which are probably out of reach for non-expert users, when high-level information are sought. Instead, understanding the scene at its lowest level requires the application of scattering models. As an example, the knowledge of the precise destination of a terrain belonging to the class "Cropfields" requires the mastery of a vegetation scattering model. In the same way, the retrieval of the height of an object classified as "Urban structure" requires the knowledge of the complex scattering mechanism triggered by an urban environment.

At the end of the day, the message we convey is that a more effective integration of techniques and models is needed, especially looking toward automatic systems of image understanding which seem to become the essential core of remotely sensed data analysis in a big data scenario. We think that the full development of such systems can provide a complete answer to the question we left partially suspended.

ACKNOWLEDGMENTS

The authors would like to thank the Italian Aerospace Research Center (CIRA), the Italian Space Agency (ASI) and Dr. Davide Notti from University of Pavia for providing the Campania, Yatenga and Calabria datasets under the aegis of the projects "Intelligent Data Extraction System", "HydroCI-DOT", and "Use of SAR-satellite data to monitor and model landslides and subsidence hazards", respectively. Moreover, the authors thank the European Space Agency for making available for free Sentinel-1 data and the basic tools for their processing, and the anonymous reviewers for their help in the improvement of some ideas expressed in the work.

REFERENCES

[1] N. P. Jacobson and M. R. Gupta, "Design Goals and Solutions for Display of Hyperspectral Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 11, pp. 2684–2692, 2005.

[2] N. P. Jacobson, M. R. Gupta, and J. B. Cole, "Linear Fusion of Image Sets for Display," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3277–3288, 2007.

[3] T. A. Wilson, S. K. Rogers, and M. Kabrisky, "Perceptual-Based Image Fusion for Hyperspectral Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 1007–1017, 1997.

[4] T. Toutin, "Geometric processing of remote sensing images: models, algorithms and methods," *Int. J. Remote Sens.*, vol. 25, no. 10, pp. 1893–1924, 2004.

[5] C. Pohl and J. L. Van Genderen, "Multisensor image fusion in remote sensing: Concepts, methods and applications," *Int. J. Remote Sens.*, vol. 19, no. 5, pp. 823–854, 1998.

[6] X. Dong, S. Quegan, U. Yumiko, , C. Hu, and T. Zeng, "Feasibility Study of C- and L-band SAR Time Series Data in Tracking Indonesian Plantation and Natural Forest Cover Changes," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 8, no. 7, pp. 3692–3699, 2015.

[7] C. Dumitru, S. Cui, D. Faur, and M. Datcu, "Data Analytics for Rapid Mapping: Case Study of a Flooding Event in Germany and the Tsunami in Japan Using Very High Resolution SAR Images," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 8, no. 1, pp. 114–129, 2015.

[8] O. Antropov, Y. Rauste, H. Astola, J. Praks, T. Hame, and M. Hallikainen, "Land Cover and Soil Type Mapping From Spaceborne PolSAR Data at L-Band With Probabilistic Neural Network," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5256–5270, 2014.

[9] F. Bovolo and L. Bruzzone, "The Time Variable in Data Fusion: A Change Detection Perspective," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 8–26, 2015.

[10] J. A. Richards, *Remote Sensing Digital Image Analysis*, 2nd ed. Berlin, Germany: Springer-Verlag, 1995.

[11] G. Franceschetti and R. Lanari, *Synthetic Aperture Radar Processing*. Boca Raton, FL: CRC Press, 1999.

[12] J. B. Campbell and R. H. Wynne, *Introduction to Remote Sensing*. New York, NY 10012: The Guilford Press, 2011.

[13] M. Schroder, H. Rehrauer, K. Seidel, and M. Datcu, "Spatial Information Retrieval from Remote-Sensing Images - Part II : Gibbs-Markov Random Fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1446–1455, 1998.

[14] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*. Pacific Grove, CA, 93950, USA: Brooks/Cole Publishing Company, 1999.

[15] D. Amitrano, G. Di Martino, A. Iodice, D. Riccio, and G. Ruello, "A New Framework for SAR Multitemporal Data RGB Representation: Rationale and Products," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 117–133, 2015.

[16] M. Thiriet and K. H. Parker, "Physiology and pathology of the cardiovascular system: A physical perspective," in *Cardiovascular Mathematics. Modeling and Simulation of the Circulatory System*, L. Formaggia, A. Quarteroni, and A. Veneziani, Eds. Milan: Springer-Verlag, 2009.

[17] R. Guida, A. Iodice, D. Riccio, and U. Stilla, "Model-Based Interpretation of High-Resolution SAR Images of Buildings," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 1, no. 2, pp. 107–119, 2008.

[18] T. Matsuyama and V. S.-H. Hwang, *SIGMA - A Knowledge-Based Aerial Image Understanding System*. New York: Plenum Press, 1990.

[19] M. Nagao and T. Matsuyama, *A Structural Analysis of Complex Aerial Photographs*. New York: Plenum Press, 1980.

[20] D. Marr, *Vision*. San Francisco: W. H. Freeman, 1982.

[21] R. A. Brooks, "Symbolic Reasoning Among 3-D Models and 2-D Images," *Artificial Intelligence*, no. 1-3, pp. 285–348, 1981.

[22] T. Poggio, "Marr's computational approach to vision," *Trends in NeuroSciences*, vol. 4, no. 10, pp. 258–262, 1981.

[23] R. A. Brooks, *Model-Based Computer Vision*. Ann Arbor, MI: UMI Research Press, 1984.

[24] G. Gennarelli, G. Vivone, P. Braca, F. Soldovieri, and M. Amin, "Multiple Extended Target Tracking for Through-Wall Radars," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6482–6494, 2015.

[25] J. M. Mendel, "Fuzzy Logic Systems for Engineering: A Tutorial," *Proc. IEEE*, vol. 83, no. 3, pp. 345–377, 1995.

[26] S. Santini, A. Gupta, and R. Jain, "Emergent semantics through interaction in image databases," *IEEE Trans. Knowl. Data Eng.*, vol. 13, no. 3, pp. 337–351, 2001.

[27] V. Madhok and D. A. Landgrebe, "A Process Model for Remote Sensing Data Analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 680–686, 2002.

[28] M. Datcu and K. Seidel, "Human-Centered Concepts for Exploration and Understanding of Earth Observation Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 52–59, 2005.

[29] R. Gaetano, D. Amitrano, G. Masi, G. Poggi, A. Verdoliva, G. Ruello, and G. Scarpa, "Exploration of Multitemporal COSMO-SkyMed Data Via Tree-Structured MRF Segmentation," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 7, no. 7, pp. 2763–2775, 2014.

[30] A. Atkin, "Peirce's theory of signs," in *The Stanford Encyclopedia of Philosophy*, Summer 2013 ed., E. N. Zalta, Ed., 2013.

[31] e geos. COSMO-SkyMed Image Calibration. [Online]. Available: http://www.e-geos.it/products/pdf/COSMO-SkyMed-Image_Calibration.pdf

[32] Infoterra. Radiometric calibration of TerraSAR-X data. [Online]. Available: http://www2.astrium-geo.com/files/pmedia/public/r465_9_tsxx-itd-tn-0049-radiometric_calculations_i1.00.pdf.

[33] G. F. De Grandi, M. Leysen, J.-S. Lee, and D. Schuler, "Radar reflectivity estimation using multiple SAR scenes of the same target: technique and applications," in *IEEE Int. Geosci. Remote Sens. Symp.*, 1997, pp. 1047–1050.

[34] L. Bruzzone, M. Marconcini, U. Wegmuller, and A. Wiesmann, "An Advanced System for the Automatic Classification of Multitemporal SAR Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1321–1334, 2004.

[35] A. Khellaf, A. Beghdadi, and H. Dupoisot, "Entropic Contrast Enhancement," *IEEE Trans. Med. Im.*, vol. 10, no. 4, pp. 589–592, 1991.

[36] C. Healey, K. S. Booth, and J. Enns, "Visualizing real-time multivariate data using preattentive processing," *ACM Trans. Model. Comput. Simul.*, vol. 5, no. 3, pp. 190–221, 1995.

[37] S. G. Dellepiane and E. Angiati, "A New Method for Cross-Normalization and Multitemporal Visualization of SAR Images for the Detection of Flooded Areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2765–2779, 2012.

[38] M. Bossard, J. Feranec, and J. Otahel, "CORINE land cover technical guide: Addendum 2000," European Environment Agency Copenhagen, Tech. Rep., 2000.
[39] A. K. Fung, "Scattering from a Vegetation Layer," *IEEE Trans. Geosci. Elect.*, vol. 17, no. 1, pp. 1–6, 1979.
[40] Z. Malenovský, H. Rott, J. Cihlar, M. E. Schaepman, G. García-Santos, R. Fernandes, and M. Berger, "Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land," *Remote Sens. Environ.*, vol. 120, pp. 91–101, 2012.
[41] D. Amitrano, G. Di Martino, A. Iodice, F. Mitidieri, M. N. Papa, D. Riccio, and G. Ruello, "Sentinel-1 for Monitoring Reservoirs: A Performance Analysis," *Remote Sens.*, vol. 6, pp. 10 676–10 693, 2014.
[42] F. De Zan and A. Monti Guarnieri, "TOPSAR: Terrain Observation by Progressive Scans," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2352–2360, 2006.
[43] D. Amitrano, G. Di Martino, A. Iodice, D. Riccio, G. Ruello, M. N. Papa, F. Ciervo, and Y. Koussoube, "Effectiveness of high-resolution SAR for water resource management in low-income semi-arid countries," *Int. J. Remote Sens.*, vol. 35, no. 1, pp. 70–88, 2014.
[44] D. Amitrano, F. Ciervo, G. Di Martino, M. N. Papa, A. Iodice, Y. Koussoube, F. Mitidieri, D. Riccio, and G. Ruello, "Modeling Watershed Response in Semiarid Regions with High Resolution Synthetic Aperture Radars," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 7, no. 7, pp. 2732–2745, 2014.
[45] D. Amitrano, G. Di Martino, A. Iodice, D. Riccio, G. Ruello, F. Ciervo, M. N. Papa, and Y. Koussoube, "High resolution SAR for the monitoring of reservoirs sedimentation and soil erosion in semi arid regions," in *IEEE Int. Geosci. Remote Sens. Symp.*, 2013, pp. 911–914.
[46] D. Amitrano, G. Di Martino, A. Iodice, D. Riccio, and G. Ruello, "An end-user-oriented framework for the classification of multitemporal SAR images," *Int. J. Remote Sens.*, vol. 37, no. 1, pp. 248–261, 2016.
[47] T. Kohonen, *Self-Organizing Maps*. Berlin, Heidelberg: Springer-Verlag, 2001.
[48] P. Törönen, M. Kolehmainen, G. Wong, and E. Castrén, "Analysis of gene expression data using self-organizing maps," *FEBS Lett.*, vol. 45, no. 2, pp. 142–146, 1999.
[49] C. Serrano-Cinca, "Self organizing neural networks for financial diagnosis," *Decis. Support Syst.*, vol. 17, no. 3, pp. 227–238, 1996.
[50] B. C. Hewitson and R. G. Crane, "Self-organizing maps: applications to synoptic climatology," *Climate Res.*, vol. 22, pp. 13–26, 2002.
[51] S. Dollhopf, S. Hashsham, and J. Tiedje, "Interpreting 16S rDNA T-RFLP Data: Application of Self-Organizing Maps and Principal Component Analysis to Describe Community Dynamics and Convergence," *Microbial Ecol.*, vol. 42, no. 4, pp. 823–836, 2013.
[52] K. Kiviluoto, "Predicting bankruptcies with the self-organizing map," *Neurocomputing*, vol. 21, no. 1-3, pp. 191–201, 1998.
[53] P. T. Pearson and C. I. Cooper, "Using Self Organizing Maps to Analyze Demographics and Swing State Voting in the 2008 U.S. Presidential Election," in *Artificial Neural Networks in Pattern Recognition*, ser. Lecture Notes in Computer Science, N. Mana, F. Schwenker, and E. Trentin, Eds. Springer Berlin Heidelberg, 2012, vol. 7477, pp. 201–212.
[54] P. M. Atkinson and R. L. Tatnall, "Introduction Neural networks in remote sensing," *Int. J. Remote Sens.*, vol. 18, no. 4, pp. 699–709, 1997.
[55] M. M. Kamal, P. J. Pasmore, and I. D. H. Shepherd, "Integration of geographic information system and RADARSAT synthetic aperture radar data using a self-organizing map network as compensation for real-time ground data in automatic image classification," *Journal Appl. Remote Sens.*, vol. 4, no. 1, pp. 1–13, 2010.
[56] S. Patra and L. Bruzzone, "A Novel SOM-SVM-Based Active Learning Technique for Remote Sensing Image Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 606–616, 2014.
[57] M. L. Gonçalves, M. L. A. Netto, J. A. F. Costa, and J. Zulio Júnior, "An unsupervised method of classifying remotely sensed images using Kohonen self-organizing maps and agglomerative hierarchical clustering methods," *Int. J. Remote Sens.*, vol. 29, no. 11, pp. 3171–3207, 2008.
[58] Y. Ito and S. Omatu, "Polarimetric SAR data classification using competitive neural networks," *Int. J. Remote Sens.*, vol. 19, no. 14, pp. 2265–2684, 1998.
[59] V.-E. Neagoe, R.-M. Stoica, A.-I. Ciurea, L. Bruzzone, and F. Bovolo, "Concurrent Self-Organizing Maps for Supervised/Unsupervised Change Detection in Remote Sensing Images," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 7, no. 8, pp. 3525–3533, 2014.
[60] S. Skakun, "A Neural Network Approach to Flood Mapping Using Satellite Imagery," *Computing and Informatics*, vol. 29, pp. 1013–1024, 2010.
[61] M. L. Gonçalves, J. A. F. Costa, and M. L. A. Netto, "Land-Cover Classification Using Self-Organizing Maps Clustered with Spectral and Spatial Information," in *Self Organizing Maps - Application and Novel Algorithm Design*, J. I. Mwasiagi, Ed. InTech, 2011, pp. 299–322.
[62] R. Kruse, C. Borgelt, F. Klawonn, C. Moewes, M. Steinbrecher, and P. Held, *Computational Intelligence: A Methodological Introduction*. Berlin, Heidelberg: Springer-Verlag, 2013.
[63] J. H. J. Ward, "Hierarchical Grouping to Optimize an Objective function," *J. Am. Stat. Assoc.*, vol. 58, no. 30, pp. 236–244, 1963.
[64] L. Shapiro and G. Stockman, *Computer Vision*. Upper Saddle River, NJ: Prentice Hall, 2002.
[65] A. K. Shackelford and C. H. Davies, "A Combined Fuzzy Pixel-Based and Object-Based Approach for Classification of High-Resolution Multispectral Data Over Urban Areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 10, pp. 2354–2363, 2003.
[66] A. Baraldi, "Satellite Image Automatic Mapper (SIAM) - A Turnkey Software Executable for Automatic Near Real-Time Multi-Sensor Multi-Resolution Spectral Rule-Based Preliminary Classification of Spaceborne Multi-Spectral Images," *Recent Patents on Space Technology*, vol. 1, no. 2, pp. 81–106, 2011.
[67] G. Franceschetti, M. Migliaccio, D. Riccio, and G. Schirinzi, "SARAS: A Synthetic Aperture Radar (SAR) Raw Signal Simulator," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 1, pp. 110–123, 1992.
[68] G. Franceschetti and D. Riccio, *Scattering, Natural Surfaces and Fractals*. Burlington, MA: Academic, 2007.
[69] R. Guida, A. Iodice, and D. Riccio, "Height Retrieval of Isolated Buildings from Single High-Resolution SAR Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 2967–2979, 2010.
[70] G. Franceschetti, A. Iodice, and D. Riccio, "A canonical problem in electromagnetic backscattering from buildings," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1787–1801, 2002.