Mosaicking Copernicus Sentinel-1 Data at Global Scale

Vasileios Syrris ⬤, Christina Corbane ⬤, Martino Pesaresi, and Pierre Soille

Abstract—This paper presents a processing chain for handling big volume of remotely sensing data for generating wide extent mosaics. More specifically, the data under consideration are level-1 ground range detected Sentinel-1 products with dual polarisation (VV+VH or HH+HV). Two approaches for a) distribution discretization accompanied by false color composition and b) image rendering and mosaicking are proposed. While these two components are necessary constituents of the presented mosaicking workflow, they can operate independently of each other. The design of the processing chain satisfies three objectives: i) contrasting derivative products of the input Sentinel-1 imagery such as the Global Human Settlement Layer, ii) adapting on a high-throughput computing system for fast execution, and iii) allowing potential extensions to more complex applications such as the image classification. Fast processing, process automation, incremental adjustment and information distinction are the main advantages of the proposed method. Elaboration and focus on these features are carried out during the presentation of the results.

Index Terms—Big data, remote sensing, mosaic, Sentinel-1, SAR, Copernicus, image compositing, histogram, discretization, rendering, image classification

---

1 INTRODUCTION

LARGE scale mosaics of satellite data are crucial for several applications that involve geo-analysis with remote sensing data such as mapping large natural hazard areas [1], landcover classification [2] and for guiding field investigations [3]. Satellite image mosaicking consists in splicing multiple orthoimages which have overlaps between each other into a seamless complex one. Several methods for image mosaicking have been developed with variations related either to the type of sensor (optical or radar), to the platform (aerial, satellite) or to the type of application (for quantitative or qualitative assessments).

Sensor based methods can be split into two main categories: optical and Synthetic Aperture Radar (SAR) imagery. In case of optical imagery, the most critical issues to be addressed are the inhomogeneous radiometry related to different illumination conditions across overlapping scenes, different content such as changed agriculture or forestry and undesired objects such as clouds and shadows in the images [4]. For SAR data, several factors affect the geometric and radiometric quality of the data like terrain distortions, changes in look angles or in look directions. Hence, they require adequate methods for compensating those effects when mosaicking the scenes [5].

Platform based methods follow also two trends: for aerial and Unmanned Aerial Vehicle (UAV), non nadir views, parallax effects, lense distortions require tailored image registration and colour equalization methods for image compositing (e.g., the network-based color equalization approach and the least squares adjustment). For satellite images, changes in the intensity and direction of solar illumination combined with differences in atmospheric conditions call for specific colour balancing approaches usually based on state-of-the quadratic programming methods [6] or using histogram matching techniques successfully applied to very high resolution satellite images at continental scale [7].

Application based approaches depend on the purpose of the remote sensing image mosaic: if the mosaic is to become a viable quantitative product for deriving spectral indices or for classification, ensuring radiometric fidelity is a necessity [8]. When the purpose is to produce a seamless mosaic for illustrative use or as a geographic baseline layer for visual inspection and products updating [9], then the choice of the false colour composition and colour harmonization becomes essential.

The recent launch of Sentinel-1 (SAR) and Sentinel-2 (multispectral) missions of the Copernicus program of the European Space Station offers new capabilities for continuous generation and update of large scale mosaics in support to various fields and applications. In particular, a single Sentinel-1 satellite is potentially able to map the global landmasses in the Interferometric Wide swath mode once every 12 days, in a single pass (ascending or descending). The two-satellite constellation (Sentinel-1 A and B) offers a 6 day exact repeat cycle. These characteristics together with the ability to penetrate clouds and darkness make Sentinel-1 ideal for effective generation of valuable base maps over large areas.

● *V. Syrris and P. Soille are with the European Commission, Joint Research Centre (JRC), Directorate I. Competences, Unit I.3 Text and Data Mining, via E. Fermi 2749, Ispra , (VA) I-21027, Italy. E-mail: {vasileios.syrris, pierre.soille}@ec.europa.eu.*
● *C. Corbane and M. Pesaresi are with the European Commission, Joint Research Centre (JRC), Directorate for Space, Security & Migration, Unit E.1 Disaster Risk Management, via E. Fermi 2749, Ispra, (VA) I-21027, Italy. E-mail: {Christina.CORBAN, Martino.PESARESI}@ec.europa.eu.*

In 2016, an experiment exploiting Sentinel-1A data at global scale was performed by [10] with the aim of producing up-to-date global information on the status and evolution of built-up areas in the context of the Global Human Settlement Layer initiative [11]. For the purpose of the GHSL experiment, a dedicated selection of 5,026 Sentinel-1A products aimed at covering global landmass, with a minimum repetition. Besides being useful for the automatic extraction of information on built-up areas, the Sentinel-1A collection could be also used for generating an RGB false-colour mosaic in support to visual assessment of the results of the built-up detection. The provision of such a base map offering a user-friendly representation of the physical information is a necessary condition for the full exploitation of the Sentinel-1 sensor. None of the recently produced national and regional mosaics of Sentinel-1 data (Romania, Germany, Europe [12], etc.), covers the needs for: i) enhanced interpretability of built-up areas, ii) fully automated processing chain suitable for being executed in high-throughput computing facility, and the most important iii) global coverage.

To answer those needs, an algorithmic workflow is proposed for building and mosaicking false colour composites based on a global coverage of Sentinel-1 data. The workflow follows a fully automated and non-parametric approach, adapted for high-throughput computing systems and customized for enhancing the interpretability of Sentinel-1 data while emphasizing the presence of built-up areas. Given the type of input platform/sensor and the purpose of the mosaic, the approach developed in this paper is tailored to spaceborne SAR imagery and designed for supporting the visual assessment of automatically detected built-up areas. Potential extension of its application to image classification is discussed at the last section. Furthermore, due to the large volume, complexity and heterogeneity of the Sentinel-1 dataset involved in the experiment, the proposed workflow can be considered as a demonstration of advanced capacity for handling big earth observation data. Preliminary results were presented at the 2017 Big Data from Space conference [13].

2   COLLECTION OF SENTINEL-1 DATA

Two selections of Sentinel-1 data were used for the purpose of producing a new generation of human settlement grids in the framework of the Global Human Settlement Layer (GHSL) [14]. The first collection which consisted of Sentinel-1A images only was designed in order to optimize the data storage and download time [13]. The rules for data selection addressed: the processing level (Level-1), swath mode of data acquisition (Interferometric Wide, IW), product type (Ground Range Detected, GRD), acquisition date (between December 2015 and October 2016), dual polarization of type VV+VH mainly and less products of type HH+HV. Additionally, the priority of the product selection aimed at the land mass, presence of built-up areas (guided by previously available datasets on built-up areas), and regional consistency in acquisition orbit. The resulted collection covering the entire globe with some gaps consists of 5,026 Sentinel-1A with 10×10 m of pixel spacing and 20×22 m of spatial resolution (in range and azimuth respectively). Beginning of 2018 and after the availability and calibration of Sentinel-1B, a second selection has been formed with the aim of filling the gaps of the global map and giving the opportunity

to us to test the workflow on the new sensor and verify better whether the derived mosaic fulfilled the intended purpose. This new list consists of 2,396 products (1,732 Sentinel-1A plus 664 Sentinel-1B). The total volume of Sentinel-1 data reaches the number of 15 terabytes after terrain correction, calibration and warping to EPSG:3857, a standard for web mapping applications.

3   THE JEODPP PLATFORM

For the analysis of Sentinel-1 data, the Joint Research Centre Earth Observation Data and Processing Platform (JEODPP) was used. It is a computational platform based on both commodity hardware and open source software, adequate to support file storage at petabyte-scale. It enables information extraction from large image datasets by users originating from different application domains and having diverse data and software requirements [15]. A concise presentation of the JEODPP main components follows below:

3.1   Hardware

- An extensible computer cluster consisting of a considerable number of processing servers (at the time of writing, 47 nodes associating to 1,432 cores). They compose a high throughput commodity system allowing computation-intensive scientific calculations.
- The EOS storage system which is based on Just a Bunch of Disks (JBODs) schema. It is an open source distributed disk storage system supporting multi petabyte storage installations, developed and maintained by the European Organization for Nuclear Research (CERN) [16].
- A scalable networking system providing connection among storage, processing and meta-data servers through a (single or double) bonded network configuration.

3.2   Software and Services

- Batch processing coordinated by a flexible job scheduler (HTCondor [17]) which undertakes the task of distributing the workload over the processing servers;
- Process virtualisation aiming at resolving the issues of software versioning, code portability, and libraries and packages dependency. It is based on the lightweight Docker containerization [18] permitting flexible management of hardware resources and processing environments;
- Accessibility of data & processing services: two web-based modes are provided to the user for fast prototyping and data analytics via i) remote desktop supported by the Apache Guacamole gateway, and ii) interactive visualization and on the fly processing through Jupyter notebooks equipped with in-house hard-coded libraries.

4   PROCESSING WORKFLOW

Although the Sentinel-1 images used for the GHSL and consequently for the mosaic have been selected having as criterion the consistency in orbit acquisitions, discontinuities between
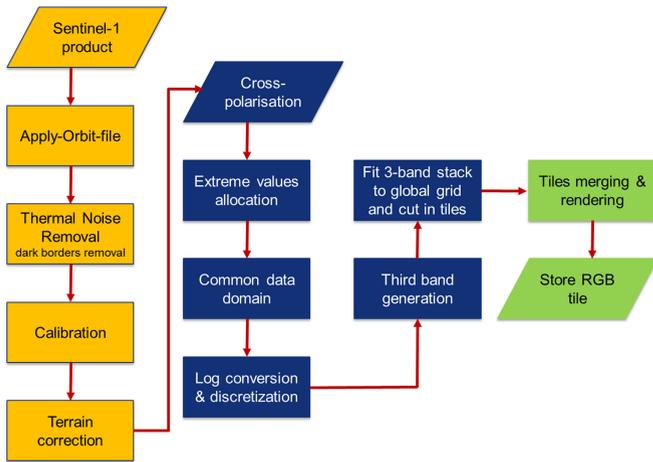
Fig. 1. The three phases of the processing workflow which correspond to Sentinel-1 product pre-processing (orange), RGB false color composition (blue) and tiles merging & rendering (green). The input images in the second phase are the two polarizations (VV+VH or HH+HV) after their calibration, terrain correction and warping to EPSG:3857.

adjacent scenes is common place reflecting i) seasonal changes in the intrinsic scattering properties of the surface features (this becomes more intense due to the wide time-frame of image acquisitions in the two collections), ii) differences in viewing geometries and/or incidence angles, and iii) topographic induced radiometric distortions. To overcome these problems, a three-stage process is proposed which involves: 1) pre-processing, 2) generation of an RBG composite and 3) mosaicking and rendering at a custom-made tile level (Fig. 1). These three phases can be executed independently of each other, permitting in that way the continuous update or retrospection of the process output and the experimentation with different image compositing and blending approaches according to evolving user requirements.

4.1 Pre-Processing of Sentinel-1 GRD Data

The Sentinel-1 Level-1 GRD images automatically downloaded and ingested in the JEODPP correspond to SAR data detected, multi-looked and projected to ground range using an Earth ellipsoid model. Geometric and radiometric calibrations and border noise removal are required prior to the generation of the false color RGB mosaic. The pre-processing chain of Sentinel-1 data, though not novel and extensively addressed in several papers is an integral part of the mosaicking workflow. The interpretive and mosaicking problems of side-looking radar images are well known and were described by [19]. In summary, geometric distortions due to the earth rotation, earth curvature, and topographic effects (layover, foreshortening, and radar shadows) generate discontinuities and variations between adjoining scenes and can result in the misinterpretation of surface features, hence limiting the use of the mosaic especially when integrated with other sources [5]. In addition to geometric corrections, a further step may be taken at the user level to radiometrically calibrate and correct relief induced radiometric distortions such as those related to local incidence angle variations. Most Sentinel-1 images have a dark border of invalid pixels the values of which are low but not necessarily 0 that impedes the calculation of accurate image statistics. Therefore, it is essential to cut-off

the dark borders without impacting the valid pixel values. The pre-processing step is hence meant to identify the above reported problems and mitigate them in the orthorectified mosaic to be produced.

A fully automated pre-processing chain, building on both the functionalities embedded in the ESA Sentinel-1 toolboxes (S1TBX) ver.2.0.2 and ver.6, and on in-house developed libraries, was executed over the JEODPP processing infrastructure using the concept of virtualization/containerization. For the removal of dark strips at the edge of the scenes of the first collection which were processed by the S1TBX ver.2.0.2, an algorithm based on connected components was applied on each scene prior to the radiometric and geometric corrections. This was then followed by thermal noise removal and radiometric calibration to sigma nought (dB). Finally a Range-Doppler terrain correction based on the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) in 1 arc-second resolution (approximately 30 m at the equator) was applied in order to orthorectfiy the SAR scenes. For the scenes falling outside the area covered by SRTM data (1,048 out of 7,422 Sentinel-1 products), terrain corrections were performed using the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 2 (GDEM V2) with a global resolution of approximately 30 m. The output of the pre-processing stage per each S1 product is composed by two orthorectified scenes resampled to 19.11 meters (corresponding to a zoom level of 13 in the Tile Map Service) with geometric, topographic and radiometric distortions removed.

4.2 False Colour Composition

In this work, the purpose is to increase the visual interpretability of the Sentinel-1 mosaic while enhancing the presence of built-up areas. This addresses the contrast adjustment of pixels intensities that correspond to different surface features. The ability to visually differentiate built-up areas from other natural or man-made features depends on the optimal combination of bands that provides the maximum separability between those different features. Generating false colour composition of SAR data is a well-known practice aimed at lowering the expertise to manage the data while enhancing the user experience/interaction with the enriched data compared to the original grayscale images. Most of the works in the literature on RGB composition of SAR imagery have been designed for multitemporal datasets and often exploit interferometric coherence and backscatter intensity change to visually distinguish different surfaces features [20], [21], [22]. In case of monotemporal dual polarization SAR data, such as the Sentinel-1 data used in this work, the RGB compositions proposed in the literature have been tailored to specific applications or use cases such as visual assessment of sea ice signature [23] or forested areas [24] or for the discrimination of several vegetation classes based on the structure of the canopy and water content of the leaves [25].

In the context of urban environments, bright features in a single SAR image indicate built-up areas and significant facilities such as railways and bridges. Co-polarized (HH or VV) and cross-polarized (HV or VH) SAR images may demonstrate distinct characteristics of the targets within the urban environment. In [26], it was demonstrated that using images
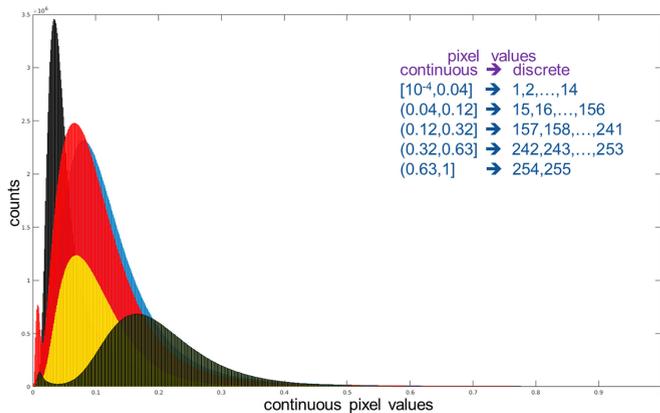
Fig. 2. The figure displays five indicative statistical distributions referring to five different VV images. The mapping shows how the critical points employed for the histograms discretization have been derived in the case of positively skewed distributions (based on the shape and the percentiles).

with dual polarization can increase the chance of successful recognition of urban targets. It has been also shown [27] that integration of co-polarized and cross-polarized images can help to determine some interesting features of urban areas, such as a pattern map of building blocks. Furthermore, the average intensity levels of the co-polarized and cross-polarized channels showed that built-up areas are distinct in radiometry from other types of landcover and water.

Taking into consideration all the above findings and after intensive testing, the following band combination has been proposed to cover sufficiently the requirements of the global Sentinel-1 mosaic:

- VH or HV band → Red channel;
- a linear combination of Red and Blue → Green channel;
- VV or HH band → Blue channel

Accordingly, the generation of the RGB false color composition is detailed below:

1) Saturate the extreme values: Given that image $I_{X_{i,i=1,2}}$ denotes one of the two $\{VH, VV\}$ or $\{HV, HH\}$ polarizations with V: Vertical, H: Horizontal, apply the function $I_{X_i}(v) = \begin{cases} 1, v > 1 \\ 10^{-4}, -1 < v \leq 10^{-4} \\ v, \text{otherwise} \end{cases}$

   where -1 has been set as no data value. Even if the range of values after calibration/terrain correction is expected to lay out between 0 and 1, few negative values may be produced due to wrong operation of the thermal noise removal which is intended for scenes over land and not over large water bodies. In the other side of the range, few values greater than 1 may be attributed to local strong scatterers. Without affecting the outcome, the function above sets all the values to the expected range;

2) Compute the common data domain for the two polarisations by applying morphological operators [28] to clean further the noisy/low value borders: $D = \bigcap_{X_{i,i=1,2}} \varepsilon_{N_{7\times7}} \Big(\varepsilon_{N_{5\times5}} \big(\delta_{N_{5\times5}} (D_{X_i}) \big) \Big)$, where $N_{7\times7}$ and $N_{5\times5}$ are two structuring elements selected for making the area compact and for cropping the image borders sufficiently; the functions δ and ε correspond

to the morphological operations of *dilation* and *erosion*. The binary images D_{X_i} have value 1 when for the corresponding values of $I_{X_i}(v)$ it holds $v > -1$ and 0 otherwise. Subsequently, the no data values of every I_{X_i} are being updated according to D;

3) Convert each updated I_{X_i} to its logarithmic counterpart and discretize its values in 8-bits by binning appropriately the image values distribution: Even though the standard practice is to transform to *db* via the function $10 \cdot \log_{10}(\cdot)$, in this case any kind of logarithmic function is producing equivalent result. The natural logarithm spreads intuitively the values without the multiplication by an extra factor; hence, $I^l_{X_i} = \ln(I_{X_i})$. By analysing the statistical distributions (SD) (Fig. 2) of the continuous values of all the $I^l_{X_i}$, we estimated the following critical ranges and values: for cross polarizations, i) $h = [8, 124, 107, 14, 2]$, $r = \ln([10^{-4}, 10^{-2}, 0.035, 0.06, 0.12, 1])$ when SD is close to normal, ii) $h = [8, 144, 87, 14, 2]$, $r = \ln([10^{-4}, 10^{-2}, 0.025, 0.06, 0.12, 1])$ when SD is positively skewed, and for co-polarizations, iii) $h = [14, 122, 105, 12, 2]$, $r = \ln([10^{-4}, 0.04, 0.14, 0.32, 0.63, 1])$ when SD is close to normal, iv) $h = [14, 142, 85, 12, 2]$, $r = \ln([10^{-4}, 0.04, 0.12, 0.32, 0.63, 1])$ when SD is positively skewed. For each of the four cases, the respective h and r vectors steer the recursive construction of a vector C which contains $|C| = 255$ values: $C = \bigcup_{k=1,\dots,5} [r_{1,k}, r_{j+1,k} = r_{j,k} + \frac{r_{1,k+1} - r_{1,k}}{h_k - 1} \mid j = 1, \dots, h_k - 1]$. Finally, by utilizing the C vector as being reversely ordered by the maximum to the minimum value, the data binning is being carried out as follows: $\forall v$ of $I^l_{X_i} : v < c_d \in C \Rightarrow B_{X_i}(v) = d$, where $d \in \mathbb{Z}^+ : 1 \leq d \leq 255$; c_d are the values of C used as thresholds over $I^l_{X_i}(v)$. $B_{X_i}(v)$ denotes the discrete representation of $I^l_{X_i}(v)$;

(4) Finally the green channel is generated by averaging the red and the blue bands and scaled by applying a gain and a bias: $\lceil \frac{Red+Blue}{2} \cdot 1.1 + 30 \rceil$. The gain and the bias inside the ceiling function $\lceil \ \rceil$ aim at shifting the low to medium values (which empirically appear to correspond in many cases to green areas) of the Red and Blue bands to higher levels of the green scale.

4.3 Mosaicking and Rendering

For the generation of the global Sentinel-1 mosaic within the JEODPP platform, we propose a fast and efficient approach for stitching image titles and rendering the RGB composition that allows task parallelization, modularity and efficient I/O handling. The approach is based on the merging of tiled Sentinel-1 colour compositions. This is performed by fitting the Sentinel-1 scenes into a regular grid and cropping them into tiles of 12,288 x 12,288 pixels (Fig. 3). The specific tile size was considered suitable in terms of I/O operations and file storage, while keeping in the same time enough information (samples) from the source image.

The tiles merging method uses a dual distance-weighted algorithm to determine the value of overlapping pixels. The output cell value of the overlapping areas will be a blend of values that overlap; the blending is based on an algorithm that is weight-based and is dependent on the distance from
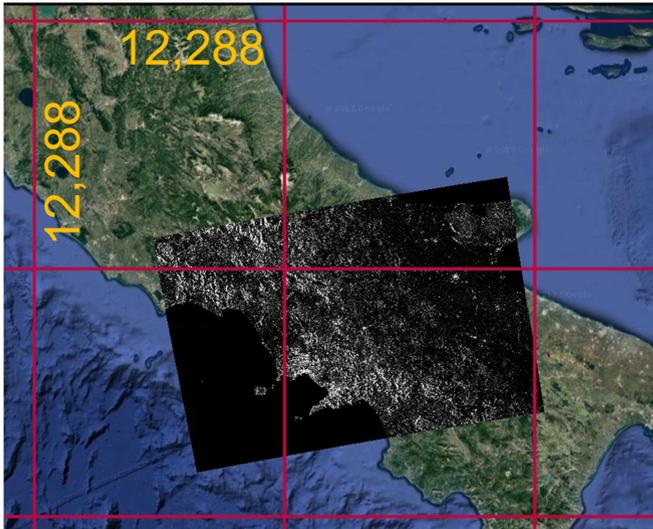
Fig. 3. Visual representation of the Sentinel-1 scene placement into a regular grid of size $12,288 \times 12,288$ pixels.

the pixel to the edge within the overlapping area. This approach has the advantage of reducing artefacts while achieving a smooth transition across two images.

Practically speaking, the process of tiles merging starts with the formation of an ordered list of overlapping tiles having as criterion the data domain size (in descending order). The first chosen tile (T_1 with data domain D_1) constitutes the canvas upon which the remaining tiles will be positioned. Next, the second tile (T_2 with data domain D_2) in the list is being evaluated.

Then, if $D_1 \cap D_2 = \emptyset \Rightarrow T_1^u = T_1 \cup T_2$, where T_1^u denotes the updated version of the base tile T_1. Otherwise, two euclidean distance transforms $DT_1(D_c)$ and $DT_2(D_c)$ are computed over the binary array $D_c = D_1 \cap D_2$. Next, the updated tile T_1^u is composed as a weighted sum as follows:

$$T_1^u = \frac{T_1(D_c) \odot DT_1(D_c) + T_2(D_c) \odot DT_2(D_c)}{DT_1(D_c) + DT_2(D_c)}, \qquad (1)$$

where the operator $\odot$ signifies the pixel-wise multiplication. For the domain $D_n = \neg D_1 \cap D_2$, it holds $T_1^u(D_n) = T_2(D_n)$.
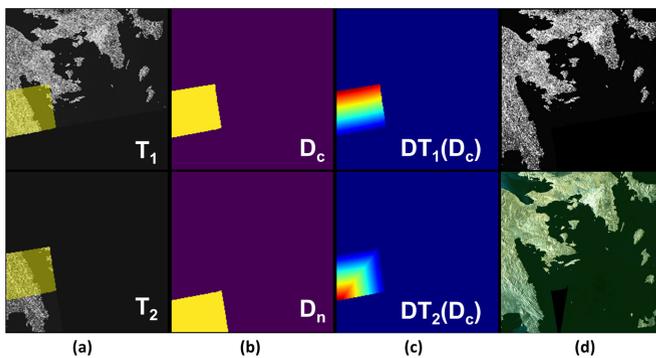

Fig. 4. An indicative example showing how the gradual tiles merging takes place. (a) Two candidate tiles T_1 and T_2 for merging; the yellow box highlights the overlapped area. (b) D_c is the datamask corresponding to the overlapped area and D_n is the area of T_2 which does not belong to T_1, yet will be added as it is on it. (c) The distance transform of the common area associated to each of the T_1 and T_2. (d) top: the result T_1^u of the tiles merging; bottom: the final tile after the gradual merging of four tiles for all of the three bands.
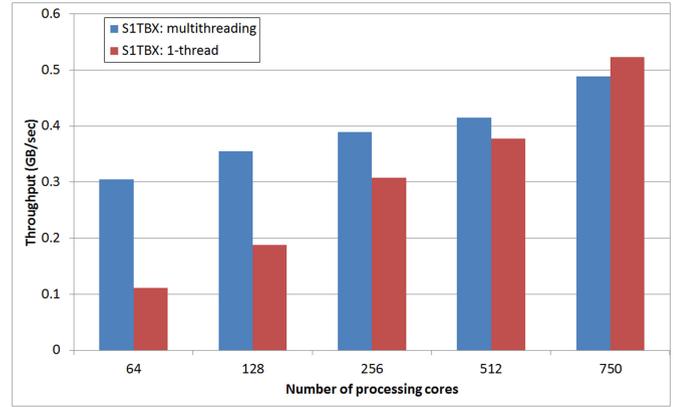

Fig. 5. The JEODPP scalability while the S1 pre-processing workflow is running (including radiometric and geometric corrections) with multi-threading enabled (left bars) and in single-threading mode (right bars). For the informed user, we mention the use of the option -q of the *gpt* command for setting 1 thread per core.

An illustrative example of the sequential merging of overlapping tiles is given in Fig. 4. Having only two tiles to process in the memory, this progressive operation turns to perform efficiently in terms of high-throughput computing, whilst allowing the incorporation of potentially new tiles without the need of re-processing other parts of the mosaic.

5 RESULTS

5.1 Processing on the JEODPP

In its current design status, the workflow for generating a mosaic from Sentinel-1 data builds on a fully automated processing chain suitable for being executed on a high-throughput computing facility. It has been optimized according to the configuration of the JEODPP platform. Fig. 5 demonstrates the scalability of JEODPP while handling the pre-processing of Sentinel-1 dataset with multi- and single-threading option enabled; the capacity is measured on how much input data can be read, processed and stored back to EOS storage system per second.

Fig. 6 shows the total elapsed time for both processes of false colour composition (652 concurrent jobs) and tiles merging and rendering (800 concurrent jobs). It is worth mentioning that both processes can be executed simultaneously on
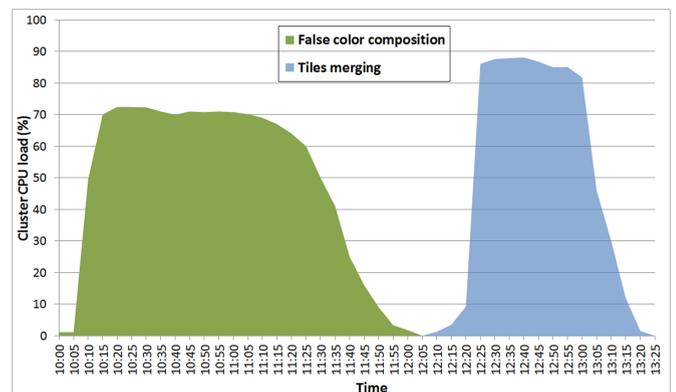

Fig. 6. The total execution time for the two stages of the main processing displayed against the cluster CPU load in a total of 912 available processing cores (652 concurrent jobs for false colour composition and 800 concurrent jobs for tiles merging and rendering).
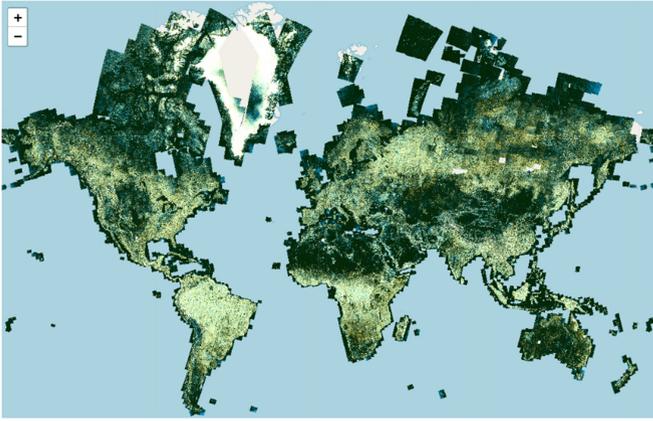
Fig. 7. The global mosaic from Copernicus Sentinel-1A and 1B data (EPSG:3857) with 19.11m spatial resolution. Web viewer: https://cidportal. jrc.ec.europa.eu/services/webview/jeodpp/databrowser/? default=jeodppS1Mosaic2016, JRC data catalogue: [26].

the cluster by setting appropriate job priorities, shortening thereby significantly the total elapsed time.

5.2 The Global Sentinel-1 RGB Mosaic

The workflow produces a dual polarization global mosaic at a spatial resolution of 19.11 m. Fig. 7 shows an overview of the landmass covered by Sentinel-1A and 1B data. This mosaic has been produced with the objective to facilitate visual assessment of built-up areas automatically extracted from Sentinel-1 data in the framework of the GHSL. This composite highlights the main landcover features. Water, flat bare soils and sandy deserts with low backscatter intensities in VV-VH and in the average of the two bands, appear in dark, almost black color. Densely vegetated areas, especially forested areas give relatively high backscatter intensities in both polarizations, especially in VH, and appear as yellowish/light green.

Built-up areas tend to have the highest intensities in both polarizations, because of multiple reflections between the ground and building walls. However, they are prone to variations in the intensity depending on the sensor-target orientation. For instance, the VH polarization is less sensitive to Bragg scattering since roof structure with a periodic pattern tends to reflect the major part of the VV signal [30]. As a result of that, built-up areas appear the brightest in the mosaic. Whenever the intensity in co-polarization (mainly in VV) is more prominent because of the sensitivity to dihedral reflection and Bragg scattering in the direction normal to the sensor track, urban areas will appear in light bluish.

Fig. 8 provides a close view of the false colour mosaic of Sentinel-1 over Madrid with urban areas easily identifiable thanks to their bright colours provided by the selected band combination. When over-imposing the built-up areas (in red) [31] on the mosaic, the qualitative assessment of the output of automatic information extraction becomes easier with the respect to the original grayscale images. The delineated built-up areas seem to match the bright targets observed in the RGB composition, but some under detections can be noticed in the dense city centre. These typical omission errors in SAR imagery occur when built-up areas are composed of houses with flat roofs that are located very close to each other so that the typical double bounce
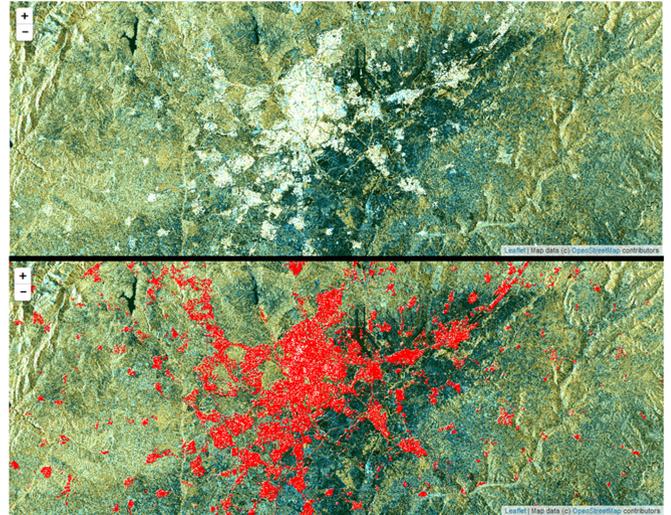

Fig. 8. Close view of Sentinel-1-mosaic over Madrid (Spain) with built-up areas visible in white and bright cyan (top). Automatically extracted built-up areas in the context of GHSL displayed in red and overlaying the Sentinel-1 mosaic (bottom). These two snapshots have been captured by the interactive, jupyter-based interface provided by the JEODPP as service.

reflection does not appear [32] - in fact the walls are not illuminated by the radar beam.

Overall, the generated mosaic provides a satisfying base layer for the purpose of the visual assessment of automatic built-up extraction. In general this product allows also a simpler separation (even at visual level) between natural and man-made features in comparison to grayscale intensity images. However, a zoom over certain areas shows that radiometric discontinuities are still visible at the edges between two scenes despite the merging and rendering approach applied for image compositing.

These discontinuities are a direct result of the sub-optimal selection of the Sentinel-1 data for mosaicking purposes. The Sentinel-1 collection used for generating the global mosaic was originally intended for built-up areas extraction with a minimum overlap to optimize the data storage and download time while covering a maximum area of the total landmass. Considerations like differences in seasonality, or differences in orbit directions in overlapping or adjacent scenes were not considered. The most frequent challenges encountered in the Sentinel-1 collection were due either to:

- Seasonal changes that induce large radiometric differences especially in mountainous areas or agricultural landscapes. Fig. 9 illustrates such an example in an area covered by two adjacent scenes acquired in January (top) and May 2016 (bottom). It can be seen that large radiometric differences exist between those two scenes with very low intensities in the image acquired in May due to eventual changes in vegetation phenological stages as well as variations of the crop type over time or changes in the soil moisture in non-vegetated areas;
- Areas imaged with opposite look directions which results in shadows becoming oppositely oriented in adjacent scenes;
- Large variations in the intensity values among the different areas of the same scene due to suboptimal

Fig. 9. Example of two adjacent scenes in VV polarization acquired in two different seasons: S1A_IW_GRDH_1SDV_20160114T150854_20160114T150923_009492_00DC6A_55B5.SAFE January (top) and S1A_IW_GRDH_1SDV_20160530T151713_20160530T151738_011490_01183A_B414.SAFE May 2016 (bottom).

de-bursting and merging of the sub-swaths. In such situation, both VV and VH images are dominated by the large scale range drop-off in intensity, making it difficult to calculate accurate image statistics for the entire scene (Fig. 10);

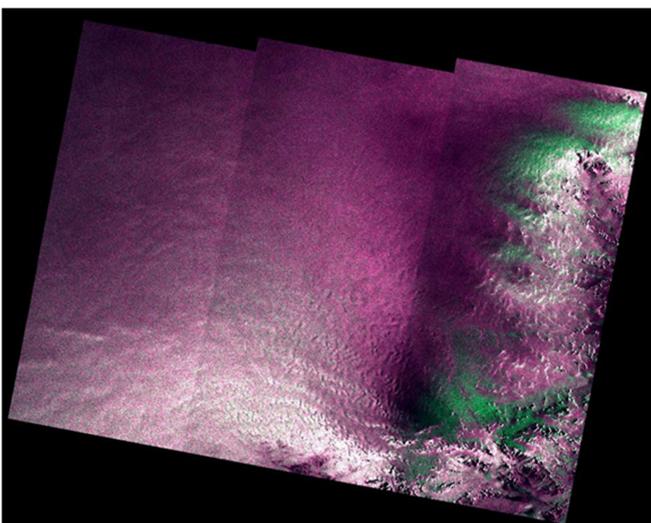- Signal saturation, especially over land which also impedes the calculation of accurate statistics.

Fig. 10. Indicative example showing radiometric artefacts due to suboptimal de-bursting and merging of the sub-swaths (scene ID: S1A_IW_GRDH_1SDH_20160331T090246_20160331T090315_010611_00FCAB_3F07.SAFE).
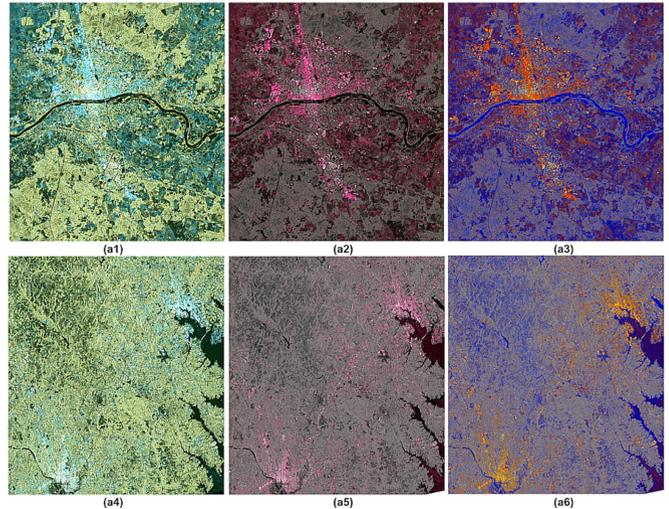

Fig. 11. Three different false color compositions: (a1-3) city of Orléans displayed at scale 1:1000000; (a4-6) the cities of Washington DC and Baltimore displayed at scale 1:5000000. First column: the proposed RGB; medium column: Red (VV), Green (VH), Blue ($0.5 \cdot (Red + Green)$); right column: Red (VV), Green (VH), Blue ($Red/Green$).

6 EXTENSIONS OF MOSAIC APPLICABILITY

Apart from the main objective of the produced Sentinel-1 mosaic to play the role of a base layer suitable for visual inspection and validation of other derived layers, in this section we investigate the possibility whether such a layer might aid to a rapid (in the prototyping sense) image classification. In such a way, the user can obtain a first informed view of what she would expect by applying more complex classification algorithms.

As we mentioned before, any of the three components of the proposed workflow can be executed independently. Regarding the RGB composite, we carried out small-scale studies with two other false color compositions which have been used extensively in the SAR literature. In Fig. 11, the middle composition has been computed by assigning VV as Red channel, VH as Green and their average as Blue. The composition on the right column has been derived by considering VV as Red channel, VH as Green and the ratio VV by VH as Blue channel. In both cases, the σ^0 values have been transferred to dB ($10 \cdot \log_{10}(\cdot)$) and then discretized using a uniform binarization after cutting the 1st and 99th percentile. Visually, at the scale of appearance, more or less they highlight the same type of information, that is, the water, the built-up and the vegetation. However, speaking about discriminative strength conveyed by the channels, experiments showed there is a clear difference among them. For instance, Fig. 12 demonstrates a classification between built-up and non built-up by applying directly a threshold over the discretized ratio channel in the ratio-based false color composition ((a3)-upper row) and the proposed one ((a3)-lower row). While the method of thresholding results approximately in the same area under curve when applied to VV or VH channels, the difference becomes apparent when considering the artificially generated third channel. For reference, in the rightmost column of Fig. 12, we show the respective area as shown in the GlobCover2009 land cover map [33], where white color denoting the "artificial surfaces and associated areas (urban areas $> 50\%$)", black color signifying "water bodies" and gray color "anything
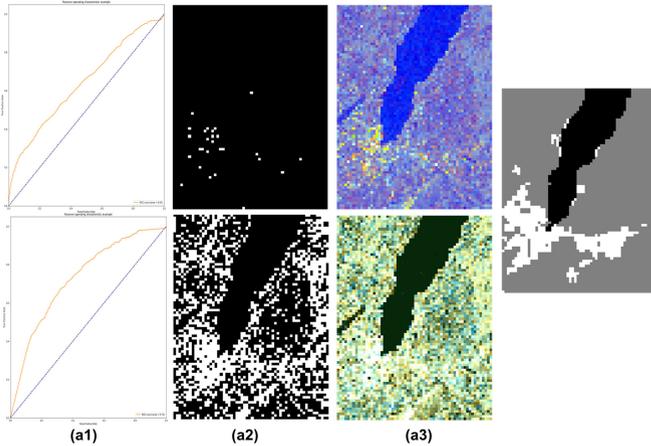
Fig. 12. Simple image classification by thresholding the discretized ratio channel with a value of 59 (a2-up) and the proposed linear combination with a value of 227 (a2-bottom). The three images (a1), (a2) and (a3) of the upper row refer to the ratio-based false color composition and the respective images in the lower row to the proposed composition. The specific thresholds are the points where the sum of sensitivity and specificity is the highest. The area under curve is 0.64 and 0.73 for (a1-top) and (a1-bottom) respectively. At the rightmost column, there is a three-class version of the specific area derived by the GlobCover2009 land cover map: white for built-up, black for water and gray for anything else.

else". Hence, apart from the natural color (green/brown for the physical environment, dark color for the water bodies and bright white/cyan color for built-up areas), the proposed RGB channel selection could accommodate a rapid, very simplistic image classification, since the experimentation demonstrates that the user is able to discriminate targets like built-up and water without effort.

Subsequently, we would like to check whether an RGB composition might be used and contribute to more sophisticated image classification. Utilizing the GlobCover2009 as reference layer in the area shown in Fig. 13, reclassified to the three classes mentioned above (built-up, water and anything else), we trained a Random Forest classifier [34] by providing the training examples displayed in Fig. 14. The aim of this experiment is not to find the best model/
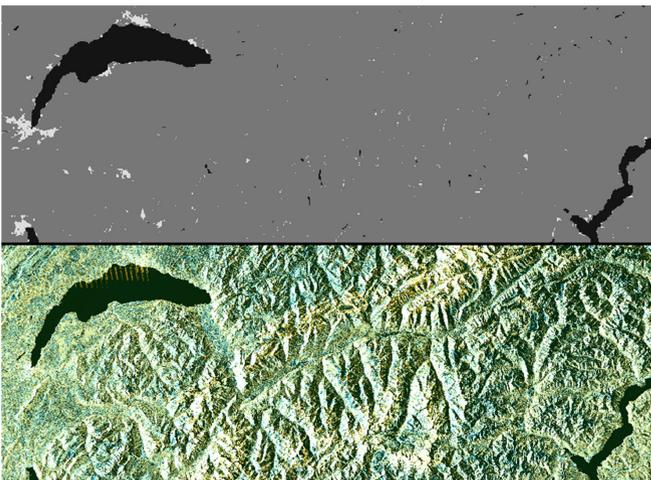

Fig. 13. The GlobCover2009 reclassified is shown in the first row (white for built-up, black for water and gray for anything else); it is used as reference layer for the image classification problem. In the second row, the same area depicted according to the proposed false color composition and rendering.
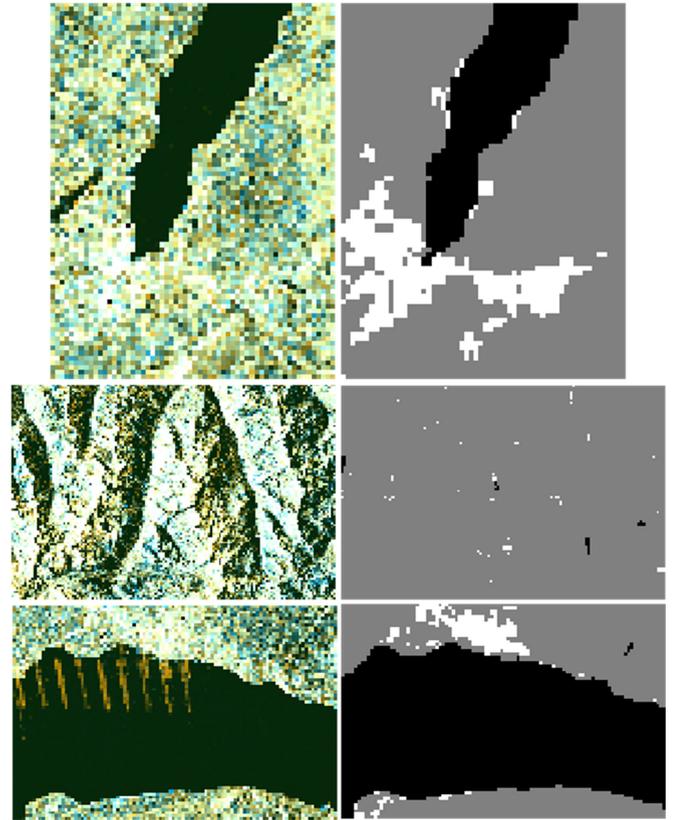

Fig. 14. Three different subareas of the area depicted in Fig. 13 (Glob-Cover2009 in grayscale and the respective false color layer). The grayscale blocks have been used as 3-class training examples to a Random Forest classifier.

classifier for the task at hand, but to test the discriminative power of the computed channels.

Fig. 15 displays four different results corresponding to four different feature selections:

(a) The VH and VV channels as two features;
(b) The VH, VV and their linear combination as three features;
(c) The VH, VV and the power to mean ratio σ/μ of VH and VV as four features, where the bold σ is the local standard deviation and μ is the local mean of the backscatter values belonging in a block of size 13×13;
(d) The same as in (c) having two more features (in total six) by including the linear combination and its power to mean ratio.

It is apparent there is an increasing improvement on the classification results as the number of features is augmenting. There is a slight yet actual improvement when using VV+VH and their linear combination instead of using only the two channels. The classifier starts to discriminate the low values of water in relationship to the low values of the mountain shadows after the use of a local statistic computed over a wide area.

The last test provides evidence about the suitability of the RGB false color composition and rendering directly to image classification. In that way, the user has a quick view of what to expect as result when proceeding towards more sophisticated modelling.
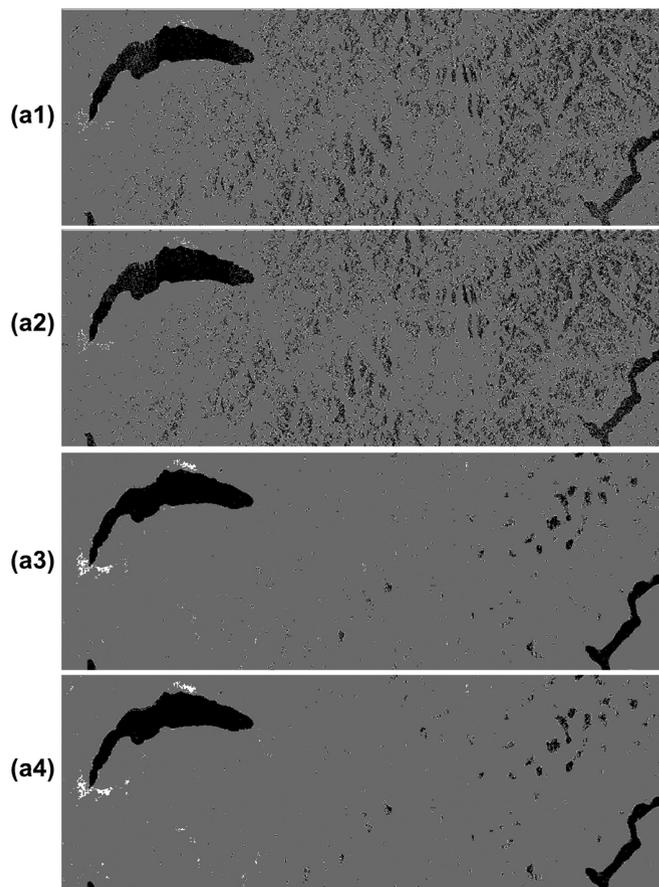
Fig. 15. Four classification results that correspond to the inclusion of an increasing number of features in the training set: (a1) VV+VH, (a2) VV +VH and their linear combination, (a3) VV+VH and their respective power to mean ratio, (a4) VV+VH and their linear combination together with their respective power to mean ratio. Classes: white for built-up, black for water and gray for anything else.

7 CONCLUSION

In this paper, we described an algorithmic workflow for the generation of an RGB false color mosaic from a global coverage of Sentinel-1. The workflow is driven by two main criteria: i) the use of the mosaic intended for qualitative purposes only, and ii) the need for a false colour composition that maximizes the contrast between built-up areas and other natural or man-made features. The workflow has the advantage of being fully automated, parameter-free and sequential. It has been designed to run on the JEODPP, a high-throughput platform at petabyte-scale, addressing big geospatial data processing and analysis needs. The functional workflow though being initially developed for Sentinel-1 data can be easily adapted to any type of high resolution satellite sensor such as Sentinel-2, assuming an optimal image selection is provided with the least cloudy-scenes [35].

The application spectrum of the Sentinel-1 mosaic is manifold. The initial purpose of the global mosaic was to assist the visual assessment of automatically extracted built-up areas in the context of the GHSL. However, the designed false colour composition allows for interactive extraction of built-up areas and water bodies using band-thresholding. These functionalities are enabled by the web-based interactive visualization and analysis component of the JEODPP [15]. The produced global mosaic provides

reference information for many users for future change detection applications with both regional and global focus. This can be achieved through yearly updates of the mosaic with global coverages of Sentinel-1A and Sentinel-1B acquisitions. The versatility of the workflow, its fast execution (less than 6 hours are currently required for the generation of a global mosaic) and the possibility for task parallelization are key features for big earth data mosaicking and for frequent updates of the mosaics.

Despite the overall satisfying visual results, the objective of reaching a fully homogeneous and seamless mosaic could not be achieved due to sub-optimal data selection and to the issues described in Section 5.2. By increasing the number of overlapping scenes while ensuring consistency in orbit direction and avoiding large temporal differences between adjacent scenes, it is possible to reduce the radiometric heterogeneities and hence improve the visual quality of the mosaic.

As a way forward, with the interactive processing and visualization components of the JEODPP, we intend to test the option of generating on the fly false colour compositions addressing different thematic applications such as change detection, monitoring of cropland, mapping of forest fires, detection of deforestation, etc. As alternative to the bands-average currently used for the green channel, it would be interesting to examine at a global scale the use of the ratio or the difference between the VV and VH bands in the blue channel as suggested in [24] and hence enhance the vegetated areas (which would then appear in green).

Multitemporal processing of Sentinel-1A and Sentinel-1B is a promising area of investigation. Some of the possible false colour compositions exploiting multitemporal SAR data were already investigated by [20], [36] at which they proposed an RGB composite based on bi-temporal images, particularly oriented toward change-detection applications. The idea would be to combine two intensity images. The third channel is reserved to the interferometric coherence, obviously computed by exploiting complex data.

The proposed workflow is robust and at the same time flexible enough to be replicated with even larger volumes of satellite datasets. This is certainly enabled and facilitated through the JEODPP architecture including the data storage, the batch processing and finally the interactive visualization and display of the results. Regarding the latter, the interactive geospatial analysis that JEODPP provides would help to assess/test/adjust on the fly different methods and rules for false colour composition and image mosaicking. The reproducibility of the method, its full automation and adaptability make it attractive for big earth data mosaicking purposes and for the generation of SAR false composites tailored to different applications and end-user communities.

REFERENCES

[1] Y. Xin, J. Li, and Q. Cheng, *Automatic Generation of Remote Sensing Image Mosaics for Mapping Large Natural Hazards Areas*. Berlin, Germany: Springer, 2007, pp. 61–73, doi: 10.1007/978-3-540-72108-6_5.

[2] S. Abdikan and C. Bayik, "Assessment of ALOS PALSAR 25-m mosaic data for land cover mapping," in *Proc. 9th Int. Workshop Anal. Multitemporal Remote Sensing Images*, Jun. 2017, pp. 1–4, doi: 10.1109/Multi-Temp.2017.8035251.

[3] H. F. Del Valle, P. D. Blanco, L. A. Hardtke, G. Metternicht, P. J. Bouza, A. Bisigato, and C. M. Rostagno, *Contribution of Open Access Global SAR Mosaics to Soil Survey Programs at Regional Level: A Case Study in North-Eastern Patagonia*. Berlin, Germany: Springer, 2016, pp. 321–346, doi: 10.1007/978-3-319-19159-1_19.

[4] F. Bignalet-Cazalet , S. Baillarin, and C. Panem, "Automatic and generic mosaicing of multisensor images: An application to pleiades HR," *Int. Archives Photogrammetry Remote Sens. Spatial Inf. Sci.*, vol. XXXIX-B1, pp. 509–512, 2012, doi: 10.5194/isprsarchives-XXXIX-B1-509-2012.

[5] Y. Sheng and D. E. Alsdorf, "Automated georeferencing and orthorectification of Amazon basin-wide SAR mosaics using SRTM DEM data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 8, pp. 1929–1940, Aug. 2005, doi: 10.1109/TGRS.2005.852160.

[6] R. Cresson and N. Saint-Geours, "Natural color satellite image mosaicking using quadratic programming in decorrelated color space," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 4151–4162, Aug. 2015, doi: 10.1109/JSTARS.2015.2449233.

[7] P. Soille, "Seamless mosaicing of very high resolution satellite data at continental scale," in *Proc. Conf. Big Data Space*, Nov. 2014, pp. 222–223,. [Online]. Available: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC92135/soille2014bids.pdf

[8] B. Guindon, "Assessing the radiometric fidelity of high resolution satellite image mosaics," *ISPRS J. Photogrammetry Remote Sens.*, vol. 52, pp. 229–243, 1997, doi: 10.1016/S0924-2716(97)00016-6.

[9] J. Liu, H. T. Li, and H. Y. Gu, "A color balancing method for wide range remote sensing imagery based on regionalization," *Int. Archives Photogrammetry Remote Sens. Spatial Inf. Sci.*, vol. XL-7/W4, pp. 103–108, 2015, doi: 10.5194/isprsarchives-XL-7-W4-103-2015.

[10] C. Corbane, G. Lemoine, M. Pesaresi, T. Kemper, F. Sabo, S. Ferri, and V. Syrris, "Enhanced automatic detection of human settlements using Sentinel-1 interferometric coherence," *Int. J. Remote Sens.*, vol. 39, no. 3, pp. 842–853, 2017, doi: 10.1080/01431161.2017.1392642.

[11] M. Pesaresi, G. Huadong, X. Blaes, D. Ehrlich, S. Ferri, L. Gueguen, M. Halkia, M. Kauffmann, T. Kemper, L. Lu, M. A. Marin-Herrera, G. K. Ouzounis, M. Scavazzon, P. Soille, V. Syrris, and L. Zanchetta, "A global human settlement layer from optical HR/VHR RS data: Concept and first results," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 5, pp. 2102–2131, Oct. 2013, doi: 10.1109/JSTARS.2013.2271445.

[12] Sentinel-1A mosaic of Europe. [Online]. Available: www.esa.int/spaceinimages/Images/2015/10/Sentinel-1A_mosaic_of_Europe, Accessed on: Jun. 30, 2017.

[13] V. Syrris, C. Corbane, and P. Soille, "A global mosaic from Copernicus Sentinel-1 data," in *Proc. Big Data Space*, 2017, pp. 267–270, doi: 10.2760/383579.

[14] C. Corbane, M. Pesaresi, P. Politis, V. Syrris, A. J. Florczyk, P. Soille, L. Maffenini, A. Burger, V. Vasilev, D. Rodriguez, F. Sabo, L. Dijkstra, and T. Kemper, "Big earth data analytics on Sentinel-1 and landsat imagery in support to global human settlements mapping," *Big Earth Data*, vol. 1, no. 1/2, pp. 118–144, 2017.

[15] P. Soille, A. Burger, D. D. Marchi, P. Kempeneers, D. Rodriguez, V. Syrris, and V. Vasilev, "A versatile data-intensive computing platform for information retrieval from big geospatial data," *Future Generation Comput. Syst.*, vol. 81, pp. 30–40, 2018, doi: 10.1016/j.future.2017.11.007.

[16] G. Adde, B. Chan, D. Duellmann, X. Espinal, A. Fiorot, J. Iven, L. Janyst, M. Lamanna, L. Mascetti, J. Pereira Rocha, A. Peters, and E. Sindrilaru, "Latest evolution of EOS filesystem," *J. Phys.: Conf. Ser.*, vol. 608, 2015.

[17] D. Thain, T. Tannenbaum, and M. Livny, "Distributed computing in practice: The Condor experience," *Concurrency and Comput.: Pract. and Experience*, vol. 17, pp. 2–4, 2005.

[18] D. Merkel, "Docker: Lightweight Linux containers for consistent development and deployment," *Linux J.*, vol. 214, no. 239, Mar. 2014. [Online]. Available: http://dl.acm.org/citation.cfm?id=2600239.2600241.

[19] A. J. Lewis and H. C. Macdonald, "Interpretive and mosaicking problems of SLAR imagery," *Remote Sens. Environ.*, vol. 1, no. 4, pp. 231–236, 1970, doi: 10.1016/S0034-4257(70)80004-9.

[20] D. Amitrano, G. D. 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, Jan. 2015, doi: 10.1109/TGRS.2014.2318997.

[21] U. Wegmuller and C. Werner, "Retrieval of vegetation parameters with SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 18–24, Jan. 1997, doi: 10.1109/36.551930.

[22] F. Siegert and A. A. Hoffmann, "The 1998 forest fires in East Kalimantan (Indonesia)," *Remote Sens. Environ.*, vol. 72, pp. 64–77, Apr. 2000, doi: 10.1016/S0034-4257(99)00092-9.

[23] R. D. Abreu, D. Flett, B. Scheuchl, and B. Ramsay, "Operational sea ice monitoring with RADARSAT-2 - A glimpse into the future," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2003, pp. 1308–1310, doi: 10.1109/IGARSS.2003.1294092.

[24] A. Dostálová, M. Hollaus, M. Milenković, and W. Wagner, "Forest area derivation from Sentinel-1 data," *ISPRS Ann. Photogrammetry Remote Sens. Spatial Inf. Sci.*, vol. III-7, pp. 227–233, 2016, doi: 10.5194/isprs-annals-III-7-227-2016.

[25] F. Siegert and S. Kuntz, "Monitoring of deforestation and land use in Indonesia with multitemporal ERS data," *Int. J. Remote Sens.*, vol. 20, pp. 2835–2853, Sep. 1999, doi: 10.1080/014311699211822.

[26] D. Perissin and A. Ferretti, "Urban-target recognition by means of repeated spaceborne SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 4043–4058, Jan. 2008.

[27] P.-H. Chen and I. Dowman, "Monitoring urban landscape of Taiwan using PALSAR data," in *Proc. 1st Joint PI Symp. ALOS Data Nodes ALOS Sci. Program Kyoto*, 2007.

[28] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Berlin, Germany: Springer, 2003.

[29] V. Syrris, C. Corbane, M. Pesaresi, and P. Soille, "Mosaic of Copernicus Sentinel-1 data at global scale," PID: 2017. [Online]. Available: http://data.europa.eu/89h/jrc-bigdataeoss-s1-mosaic

[30] H. Kimura, "Radar polarization orientation shifts in built-up areas," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 217–221, Apr. 2008, doi: 10.1109/LGRS.2008.915737.

[31] C. Corbane, M. Pesaresi, P. Politis, V. Syrris, A. Florczyk, P. Soille, L. Maffenini, A. Burger, V. Vasilev, D. R. Aseretto, F. Sabo, L. Dijkstra, and T. Kemper, "Mass processing of Sentinel-1 and Landsat data for mapping human settlements at global level," in *Proc. Conf. Big Data Space*, 2017, pp. 52–55, doi: 10.2760/383579.

[32] T. Esch, W. Heldens, A. Hirner, M. Keil, M. Marconcini, A. Roth, J. Zeidler, S. Dech, and E. Strano, "Breaking new ground in mapping human settlements from space-the global urban footprint," *ISPRS J. Photogrammetry Remote Sens.*, vol. 134, pp. 30–42, 2017.

[33] GlobCover 2009. [Online]. Available: http://due.esrin.esa.int/files/GLOBCOVER2009_Validation_Report_2.2.pdf, Accessed on: Jun. 30, 2017.

[34] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[35] P. Kempeneers and P. Soille, "Optimizing Sentinel-2 image selection in a big data context," *Big Earth Data*, vol. 1, no. 1/2, pp. 145–158, 2017, doi: 10.1080/20964471.2017.1407489.

[36] D. Amitrano, G. D. Martino, A. Iodice, D. Riccio, and G. Ruello, "RGB SAR products: Methods and applications," *Eur. J. Remote Sens.*, vol. 49, no. 1, pp. 777–793, 2016, doi: 10.5721/EuJRS20164941.

Vasileios Syrris received the BMath degree in mathematics from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1995, the MSc degree in artificial intelligence: knowledge-based systems from the University of Edinburgh, Edinburgh, Scotland, United Kingdom, in 1996, and the PhD degree in computational intelligence from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2010. He worked as an assistant lecturer/tutor with Automation and Robotics Laboratory, Aristotle University of Thessaloniki and with the Informatics and Electronics Department, Technological Educational Institute of Thessaloniki, Thessaloniki, Greece. He has worked as a postdoctoral researcher with Joint Research Centre of European Commission, Institute for the Protection and Security of the Citizen, Global Security and Crisis Management Unit, Ispra, Italy. Currently he is working as a scientific/technical project officer with Joint Research Centre of European Commission, Directorate I. Competences, Unit I.3 Text and Data Mining, Ispra, Italy. His research interests include machine learning, high-performance computing, automation, computer vision, remote sensing, robotics, control engineering, statistics and big data analysis.

Christina Corbane received the MSc degree in modelling of snow water equivalent using radar imagery from Université Saint Joseph Lebanon in collaboration with INRS-ETE, Québec, in 2001. She received the PhD degree in remote sensing applied to earth sciences and environmental sciences from the Université Montpellier II, Montpellier, France, in 2006. Her academic and professional background is in remote sensing and spatial analysis applied to the study of earth's environment and to key disaster management and security issues. In particular, she worked on the development of a prototype for automatic ship detection from remote sensing imagery and on a rapid urban mapping model between 2007 and 2009 as part of her appointment at the Institut de Recherche pour le Développement (IRD). From 2009 to 2012, she worked as a scientific and technical officer with the Joint Research Centre of the European Commission (JRC) and was responsible for the development of methodologies for the validation of rapid geo-information for disaster management. From 2012 to 2014, she joined the National Research Institute of Science and Technology for Environment and Agriculture (Irstea), Montpellier, France, to conduct research in mapping and monitoring of natural habitats with earth observation data. Between 2014 and 2016, she was involved in research activities at the JRC related to disaster risk reduction and Pan-European risk assessment. Her current responsibilities involve analysing the fitness for purpose of Sentinel-1 and Sentinel-2 for global human settlement and land cover mapping.

Martino Pesaresi works at the European Commission, Joint Research Center (JRC), contributing to programs dealing with the use of space technologies for automatic image information retrieval and decision support systems in the areas of post-natural-disaster and post-conflict damage assessment, conflict-related resource monitoring, risk and exposure mapping. From 2014 to 2016, he was initializing and leading the Global Human Settlement Layer (GHSL) project activities of the JRC, establishing new geospatial analytics technologies for assessing the human presence on the planet supporting post-2015 international frameworks. He is the chair of the "Human Planet" international initiative in the frame of the Group of Earth Observation (GEO) work program 2017-2019. From 2007 to 2013, he was the leader of the team "Information Support for Effective and Rapid External Action" (ISFEREA) of the JRC, developing remote sensing applications. In 2005-2007 he was the scientific coordinator of the Global Monitoring for Security and Stability (GMOSS) Network of Excellence.

Pierre Soille received the engineering degree from the University of Louvain, Louvain-la-Neuve, Belgium, in 1988, the PhD degree in agronomical sciences from the same university in collaboration with the Centre for Mathematical Morphology of the Ecole des Mines de Paris, France, in 1992, and the habilitation degree in computer science from Montpellier II University, Montpellier, France, in 1997. After completion of his PhD on mathematical morphology applied to digital elevation models and satellite image analysis, he pursued research on image analysis and mathematical morphology at the CSIRO Mathematical and Information Sciences in Sydney, Australia, the Centre for Mathematical Morphology of the Ecole des Mines de Paris, France, and the Pattern Recognition Department of the Fraunhofer Institut IPK in Berlin, Germany. During the period 1995-1998 he was assistant professor with the Ecole de Mines d'Alès and EERIE, Nîmes, France. In 1999, he was appointed as senior researcher with the Biotechnology and Biological Sciences Research Council, United Kingdom. He has been with the Joint Research Centre (JRC) of the European Commission, Ispra, Italy, since December 1999. He is leading the JRC big data pilot project on Earth Observation and Social Sensing that focuses on the development of big data technologies in the broad context of geospatial data with emphasis on the exploitation of the data streams originating from the European Union Copernicus programme and its fleet of Sentinel satellites. He has authored and co-authored numerous and highly cited image analysis and pattern recognition algorithms and methods published in peer-reviewed journal and conference papers as well as book chapters. He has also edited various books and is author of a reference monograph on Morphological Image Analysis.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/csdl.