Methods to Remove the Border Noise From Sentinel-1 Synthetic Aperture Radar Data: Implications and Importance For Time-Series Analysis

Iftikhar Ali, Senmao Cao, Vahid Naeimi, Christoph Paulik, and Wolfgang Wagner, *Senior Member, IEEE*

***Abstract*—The Sentinel-1 GRD (ground range detected) Level-1 product generated by the Instrument Processing Facil-ity of the European Space Agency has noise artifacts at the image borders, which are quite consistent at both left and right sides of the satellite's cross track and at the start and end of the data take along track. The Sentinel-1 border noise troubles the creation of clean and consistence time series of backscatter. Data quality control and management become very challenging tasks, when it comes to the large-scale data processing, both in terms of spatial coverage and data volume. In this paper, we evaluate three techniques for removing the Sentinel-1 border noise and compare the results with the existing "*Sentinel-1 GRD Border Noise Removal*" algorithm implemented in the Sentinel-1 toolbox of the Sentinel application platform.[1] Validation and evaluation of the newly pro-posed algorithms was done using random samples containing 1500 Sentinel-1 scenes selected from a complete Sentinel-1 archive. The newly proposed approach has successfully achieved the required level of accuracy and solved the issue of time-series anomalies due to the border noise.**

***Index Terms*—C-band synthetic aperture radar (SAR), interfer-ometric wide swath (IW), Sentinel-1, Sentinel application platform (SNAP), Sentinel-1 border noise.**

I. Introduction

WITH the advancement in technology, space industry is growing at a rapid pace. In several years, many new satellites have been launched for effective and real-time Earth observation. These advancements have resulted in the produc-tion of huge amounts of data because of wider spatial cover-age and improved spatiotemporal resolution. Management and quality insurance of large amounts of data at the scale of several hundreds of terabytes to petabyte have been becoming a challenging task.

The recently launched constellation of two C-band synthetic aperture radar satellites named Sentinel-1 (A and B) is getting immense attention from both scientists and commercial users. Due to the open access data distribution policy and improved data quality in terms of temporal and spatial resolution, the Sentinel-1 user community is growing very fast. Daily global data acquisition of Sentinel-1 A and B is more than 1.5 TB [1]. In order to handle and process such big data for regional to global scale monitoring, reliable and stable processing chains are required.

In this paper, we address the issue of image border noise in the Sentinel-1 GRDH (ground-range-detected high resolution) Level-1 product for large-scale data processing. The proposed methods are self-adaptive and are applicable to other acquisition modes, however, fine-tuning might be needed.

A. Overview: Sentinel-1 C-Band Synthetic Aperture Radar (SAR) Sensor

Sentinel-1 is the first satellite mission of the Copernicus program, consisting of a constellation of two polar-orbiting satellites, acquiring C-band synthetic aperture radar images day and night independent of weather conditions. Sentinel-1A was launched on April 3, 2014 and Sentinel-1B on April 25, 2016. Sentinel-1 is the continuity mission to the SAR instruments flown on board of European Remote-Sensing Satellite (ERS) and ENVISAT. Over land, the Sentinel-1 SAR instrument is predominantly operating in the so-called interferometric wide swath (IW) mode. The Sentinel-1 Level-1 Ground Range Detected (GRD) products consist of focused SAR data that have been detected, multilooked, and projected to ground range us-ing an Earth ellipsoid model. The Sentinel-1 IW GRDH mode has a resolution of 20 m (10-m grid spacing). Using the two satellite units, complete global coverage can be achieved within 6–8 days [1]. Over Europe and Canada, the temporal revisit time is expected to be about 1–3 days. Sentinel-1 provides SAR data acquisitions at an unprecedented spatiotemporal sampling. The magnitude of the data volumes generated by Sentinel-1 will reach hundreds of terabytes and eventually will exceed petabyte

Manuscript received July 18, 2017; revised November 8, 2017 and December 20, 2017; accepted December 23, 2017. This work was supported in part by the Austrian Research Promotion Agency (FFG) through EOP-Danube project, "Towards an Earth Observation Platform for the Greater Danube Region" [854030] and in part by the European Union's Horizon 2020 research program through I-REACT project, "Improving Resilience to Emergencies through Advanced Cyber Technologies" [700256]. *(Corresponding author: Iftikhar Ali.)*

The authors are with the Department of Geodesy and Geoinformation (Microwave Remote Sensing Research Group), Vienna University of Technology, 1040 Vienna, Austria (e-mail: Iftikhar.Ali@geo.tuwien.ac.at; Senmao.Cao@ geo.tuwien.ac.at; Vahid.Naeimi@geo.tuwien.ac.at; Christoph.Paulik@geo. tuwien.ac.at; Wolfgang.Wagner@geo.tuwien.ac.at).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

[1] http://step.esa.int/main/toolboxes/snap/

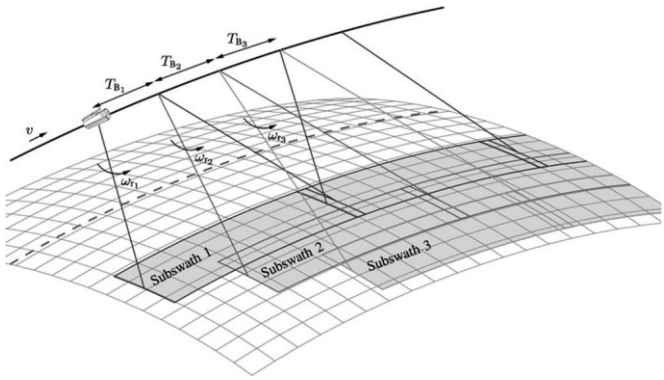


Fig. 1. Acquisition geometry of IW mode's three subswaths using TOPS (figure source: [2]).

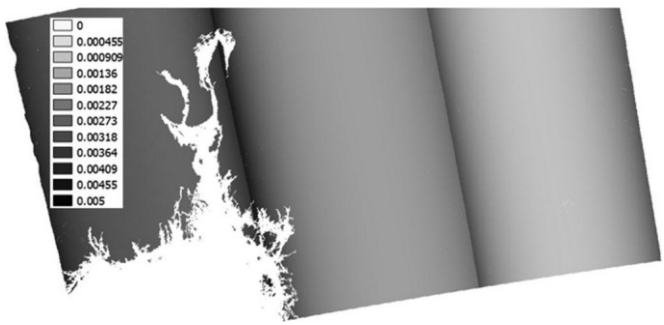


Fig. 2. Difference between SAR backscatter (σ^0 [dB]) with and without applying SNAP's *"S-1 Thermal Noise Removal"* correction module.

scale during the mission lifetime. The Sentinel-1 Level-0 and Level-1 products are segmented into "slices" of defined length along a track. Product slices make the data more manageable for the users and enable the ground segment to process slice data in parallel. The size of each Sentinel-1 Level-1 scene (25-s data acquisition scene) is about 0.8–1.7 GB depending on single or dual polarization.

B. Sentinel-1 IW Acquisition Mode and GRD Level-1 Product

Sentinel-1 operates in four acquisition modes [stripmap (SM), IW, extra-wide swath (EW), and wave mode (WV)] in order to acquire data in different spatial resolution and swath size. The IW mode is capable of acquiring data with large swath width (250 km) with a spatial resolution of 5 m × 20 m. In order to cover a large area in the azimuth direction, a so-called terrain observation with progressive scans (TOPS) technique is used. Where, for each burst B_i [where T_{Bi} is the burst duration and w_i is the steering angle rate (i = 1, 2, 3)], the beam is electronically steered from backward to forward in the azimuth direction (for details see Fig. 1).

The TOPS technique ensures homogeneous image quality through the swath, however, gradient-like thermal/system noise artifacts can occur at interswath joins of Sentinel-1 subswaths. Fig. 2 shows the scale of noise at interswath joins, where the effect of interswath noise is evident when a difference between the SAR backscatter (σ^0 [dB]) calculated with and without applying SNAP's *"S-1 Thermal Noise Removal"* correction module. Even though the difference is small [3], but in some cases, it might be very sensitive (i.e., the effect of this thermal noise was detected in surface soil moisture retrieval). The anal-ysis of noise at the interswath joins of Sentinel-1 subswaths is beyond the scope of this paper.

Level-1 GRD products are focused, multilooked and pro-jected to ground range (here phase information is lost) using an Earth ellipsoid model such as WGS84. For IW and EW GRD products, multilooking is performed on each burst individually, and finally, all subswaths are merged to form a single image per polarization channel [4].

C. Problem Description

The Sentinel-1 Level-1 and Level-2 products are produced by the Sentinel-1 Instrument Processing Facility. Several interme-diate steps are involved to generate Level-1 products from RAW Level-0 data. During this process, different types of artifacts ap-pear at the image borders of the Level-1 products. The artifacts appear mainly during the azimuth and range compression and the handling of sampling window start time (SWST—the time offset between the start time of the transmitted pulse and the start time of the current received sampling window) [5]. As a result, artifacts like "no-value" and pixels with very low values (which are mainly due to the resampling step in the SWST) are introduced in the Level-1 products. These artifacts are quite consistent both in range and azimuth direction as shown in Fig. 3. However, the patterns in range and azimuth directions are different.

1) Cross-track border noise in the range direction appears both in near and far range, however the noise pattern is slightly different. In the near range pixels on the border have generally low values, sometimes mixed with no-values/not-a-number (NaN) pixels. In the far range, there is a low-value zone, which is always followed by a no-value/not-a-number (NaN) zone pixels (see Fig. 4). Fig. 5 illustrates the cumulative number of dirty pixels on the borders of sample images separately. As it is shown in Fig. 5, the number of columns to be removed on the right side (far range) is larger than left side (near range). The size of noise affected pixels varies from scene to scene [see Fig. 6(a), (b), and (d)].
2) Along-track border-noise artifacts in azimuth direction ap-pear only in the first and last slices of a data-take (Sentinel-1 data acquisition segment of a few minutes depending on the acquisition mode). This means that the border noise in azimuth direction appears only in IW1 and IW3 sub-swaths (either in the first or last slice of a data take) as shown in Fig. 6(a)–(c).

The accuracy requirements for border noise removal are very high since only a fully adaptive, robust, and 100% accurate method can be considered as a successful solution to the prob-lem. Only in that case, a fully automatic processing chain can be setup for operational use. Sentinel-1 monthly data over Eu-rope is approximately 4 TB, so manually checking the quality of processed data for the border noise is simply not possible. If

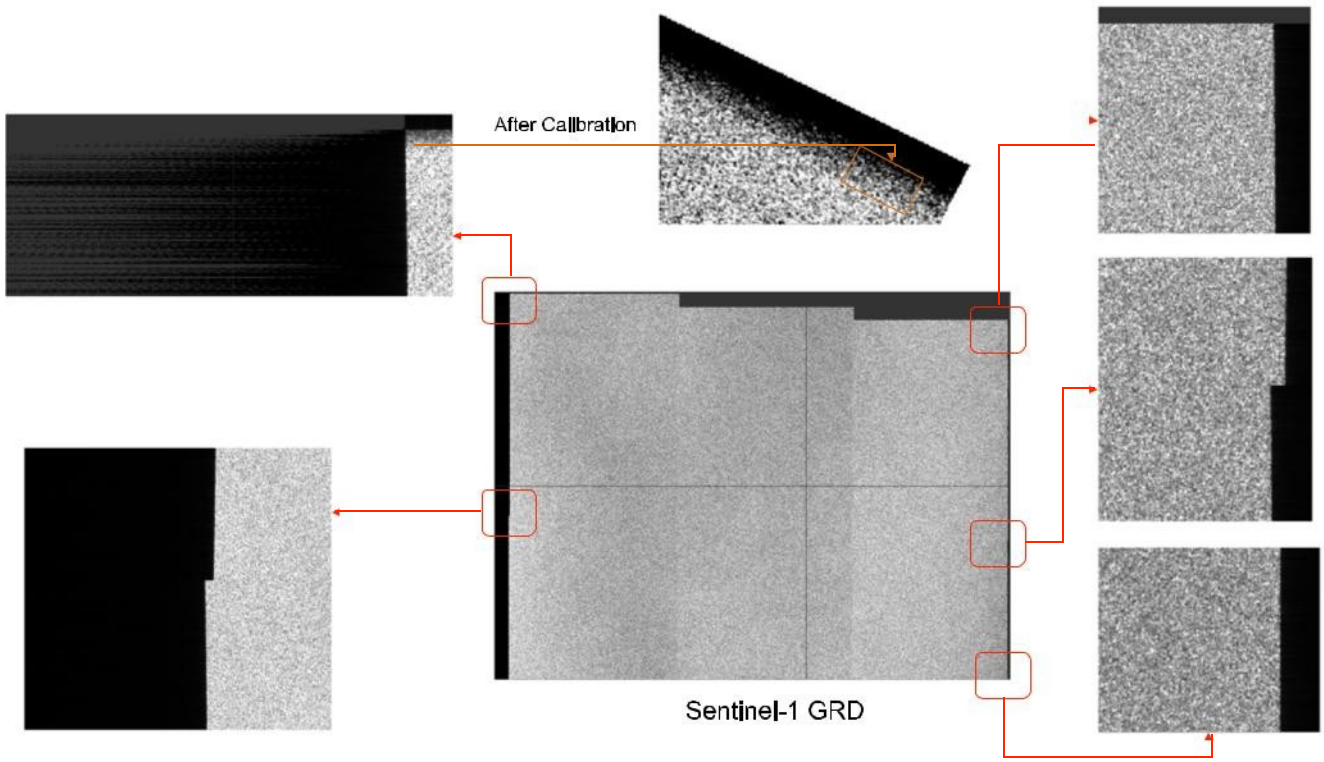


Fig. 3. Overview of "Sentinel-1 Border Noise" types, properties, and location.

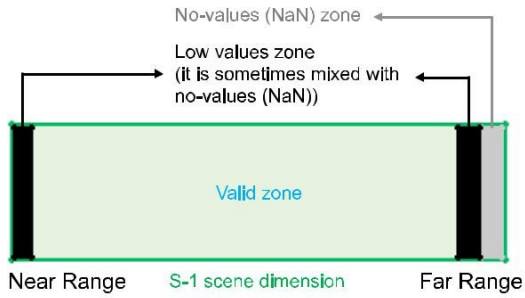


Fig. 4. Sentinel-1 border noise patterns in near- and far-range direction. Where, valid zone represents the border noise free portion of the Level-1 IW GRD product.

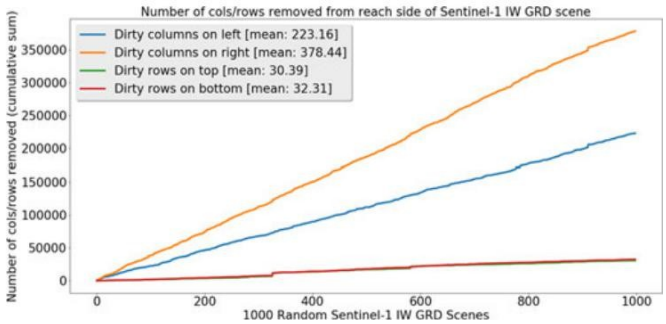


Fig. 5. Cumulative sum of affected pixel columns on left and right and affected pixel rows on top and bottom calculated from randomly selected 1000 Sentinel-1 GRD Level-1 scenes. The volume of the border noise at the top and bottom is approximately same, therefore, red and green lines are showing the same trends.

the border noise is not fully removed, it will appear in image mosaics or other higher level data products. Fig. 7(a) shows an example of Sentinel-1 mosaic of multitemporal minimum backscatter values from November 1, 2014 to July 26, 2016; the severity of the border noise problem is clearly visible in the zoomed tile.

Apart from the laborious job of checking the quality of each individual scene, optimizing the resources and computation power is another point of concern for such heavy processing. With the growing volume of data, high-performance computing and better resource utilization is a major aspect that has to be considered. Due to the large volume of the Sentinel-1 data, reprocessing of data in case of any artifact is a computationally expensive task. For the development of global or continental scale services and data production for ecosystem monitoring a highly reliable and stable data processing framework is required. The proposed approach of border noise removal will play a crucial role both in data processing [6] and applications [7]–[9].

The Sentinel-1 GRD Level-1 product file size is approximately 0.8–1.7 GB. The preprocessing of a S-1 single-scene Level-1 IW GRDH, including calibration, applying orbit files, georeferencing, terrain correction, and tilling, takes about ≈ 2.69 s per MB on a typical VSC-3[2] computing node. Fig. 7(b)

[2]Vienna Scientific Cluster (VSC-3) consists of 2020 nodes, each equipped with two processors (Intel Xeon E5-2650v2, 2.6 GHz, 8 cores Ivy Bridge-EP family) and internally connected with an Intel QDR-80 dual-link high-speed InfiniBand fabric.

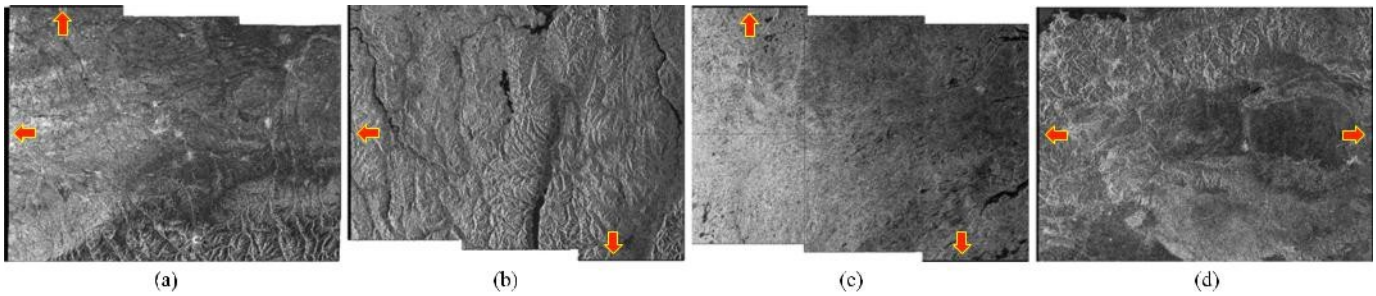


Fig. 6. Border noise examples of the Sentinel-1 GRD Level-1 product. Red arrows indicate the direction (azimuth and range) where the border noise is observable. The Sentinel-1 scenes shown in (a)–(d), respectively, are: S1A_IW_GRDH_1SDV_20141009T031532_20141009T031601_002747_003156_E57B, S1A_IW_GRDH_1SSV_20161224T013734_20161224T013759_014515_017915_0168, S1A_IW_GRDH_1SDV_20141218T050411_20141218T050441_003769_0047F3_0525, and S1A_IW_GRDH_1SDV_20141009T031716_20141009T031741_002747_003156_E3FE.

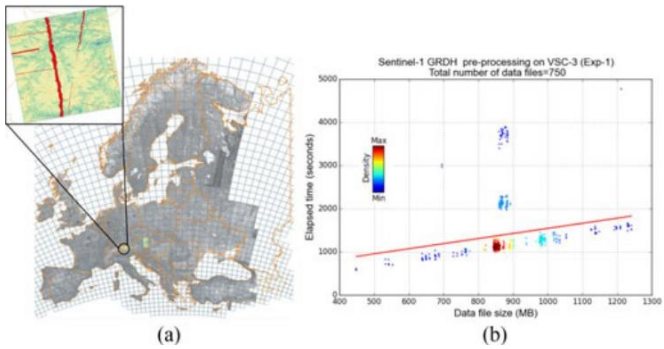


Fig. 7. Implications of the border noise for large-scale applications. (a) Sentinel-1 mosaic of multitemporal minimum backscatter values from Novem-ber 1, 2014 to 26 July 26, 2016, where zoomed inset is the single band pseu-docolor composition of the highlighted tile in (a). Red color strip-like artifacts are the result of the border noise. (b) Relation between Sentinel-1 GRD Level-1 product file size (MB) and the processing time (min) using 64-GB RAM CPU.

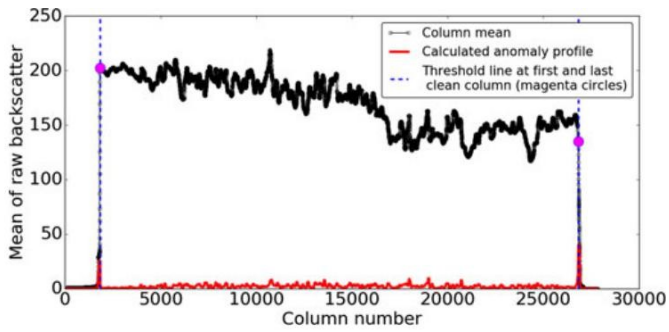


Fig. 8. Graphical abstract of the anomaly detection method (*y*-axis: raw backscatter (digital number) before radiometric calibration).

shows the relationship between the Sentinel-1 GRD file size (MB) and processing time (min) required for each file/scene. Processing time over the same latitude is more consistent, as shown in Fig. 7(b), whereas the files from the high latitude takes more processing time due to large file size caused by the map project at high latitudes. Sentinel-1 single-scene process-ing time is about 40–45 min (with Sentinel-1 Toolbox (s1tbx) version 1.1.1) that falls in the category of the computationally expensive job. Therefore, it is very important to avoid the repro-cessing of large amount of data just because of artifacts created by the border noise. This suggests that the reprocessing of such high resolution and large volume of data due to quality-control-related issues would result in an operationally unstable service.

II. Materials and Methods

This section provides a detailed description of available and suggested methods for Sentinel-1 border noise removal.

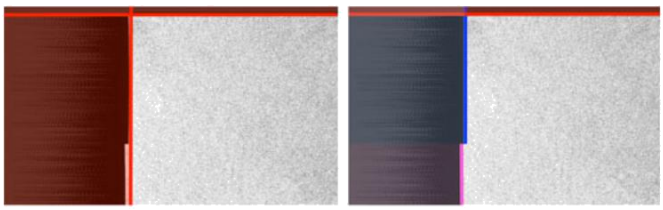
The level of signal in the copolarization channel (VV) is always higher than in the cross-polarization channel (VH) [10]. For example, over the ocean, at low wind speed, the pixel values in the cross-polarization channel are as low as the noise. Therefore, in this paper, we built a border noise removal mask based on copolarization using randomly selected samples of Sentinel-1 images. However, the border noise is equally seen in VV and VH.

A. Algorithms for "Sentinel-1 Border Noise" Correction

1) *SNAP "Sentinel-1 Remove GRD Border Noise":* The "Sentinel-1 Remove GRD Border Noise" operator imple-mented in the Sentinel-1 Toolbox in SNAP works best over land, however, over sea, it might result in some residuals. Complete algorithmic and implementation details about the "Sentinel-1 Remove GRD Border Noise" operator can be found in [11].
2) *Anomaly Detection Method:* This method is based on a statistical approach and mainly exploits the use of in-terquartile range (IQR). Before utilizing the IQR-based Tukey test [12], we performed some array manipulation and tuning of related constant values in order to formulate the required scheme. The algorithm is described in detail in Algorithm 1 and Fig. 8.
3) *Bidirectional-Sampling-Based Methods:* Here, the term "bidirectional" refers to the range and azimuth direction, the following two algorithms remove the Sentinel-1 bor-der noise both from left/right and top/bottom of the scene. In this category, we have proposed two algorithms: "bidi-rectional *k*-samples method," which creates a mask based on one cut value for each side [see Fig. 9(a)], and "bidirec-tional all-samples method" creates a mask based on one

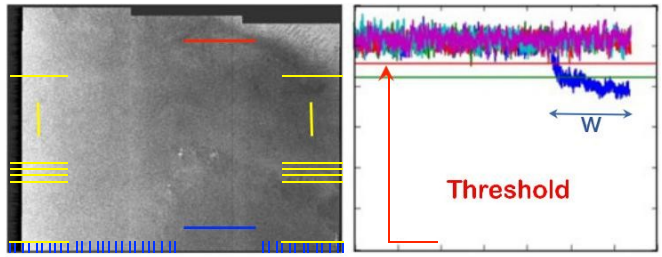
Algorithm 1: Anomaly detection method.

Data: $\mathcal{L} \leftarrow$ List of Sentinel-1 scenes each of size $m \times n$
Result: $\mathcal{I}_M \leftarrow$ Binary mask for border noise removal
for $\mathcal{I} \leftarrow 1$ *to* $\mathcal{L}$ **do**
  1. $\mathcal{I}_m \leftarrow$ Mean per column;
  2. $\mathcal{I}_s \leftarrow$ Smoothing $\mathcal{I}_m$ using $medfilt$;
  3. $\mathcal{I}_{mv} \leftarrow$ Moving average of $\mathcal{I}_s$ for k $(m = 25)$ recent values;
  4. $\mathcal{I}_a \leftarrow$ Anomaly: difference between average of k values before pixel p and average of m values after p (including the pixel p) of $\mathcal{I}_{mv}$;
  5. $\mathcal{I}_{anom} \leftarrow$ Smoothing $\mathcal{I}_a$ using $medfilt$;
  6. $\mathcal{I}_o \leftarrow$ Outlier detection
  Interquartile range based Tukey test (expression given below) was used to find the valid range outside of which values are flagged as outliers. In this case after some tests and tuning $z = 9$ was selected.
  $[(Q_1 - z(Q_3 - Q_1)), (Q_3 + z(Q_3 - Q_1))]$;
  7. $\mathcal{I}_{nc} \leftarrow$ For $\mathcal{I}_o$ set max number of noisy columns nc (where, $nc = 3000$ is selected) from each side of the image;
  8. $\mathcal{I}_{ha} \leftarrow$ Columns in an image $(\mathcal{I}_{nc})$, where anomaly is high;
  9. $\mathcal{I}_f . \mathcal{I}_l \leftarrow$ Extract the first and last pixel with valid/clean values (magenta circles in Figure 8) from $\mathcal{I}_{ha}$;
  10. $\mathcal{I}_M \leftarrow$ Create mask using $\mathcal{I}_f, \mathcal{I}_l$;
end



(a) (b)

Fig. 9. Overview sketch of Sentinel-1 border noise removal approaches. (a) Strategy for the *k*-sample method—one threshold (or cut) values for each side of the scene. (b) Strategy for *all*-sample method—one threshold (or cut) value for each step on all sides of the scene.



(a) (b)

Fig. 10. Bidirectional *k*-samples method: (a) array-based sampling framework, and (b) *W* and Threshold (*s*)-based labeling of the noise sample, which has border artifact.

Algorithm 2: Bi-directional k–samples method.

Data: $\mathcal{L} \leftarrow$ List of Sentinel-1 scenes each of size $m \times n$
Result: $\mathcal{I}_M \leftarrow$ Binary mask for border noise removal
for $\mathcal{I} \leftarrow 1$ *to* $\mathcal{L}$ **do**
  $LRBT = [\,]$;
  $\mathcal{S} \leftarrow \{S_L, S_R . S_B . S_T\}_{n=k,\ len=2000} \in \mathcal{I}$;
  for $s \leftarrow 1$ *to* $\mathcal{S}$ **do**
    $s_{fil} = [\,]$;
    if $SUM(10 \leq s \geq 0) \neq 0$ **then**
      $s_{fil}.append(s)$
    else
      Pass
    end
  end
  for $s_f \leftarrow 1$ *to* s_{fil} **do**
    $s_{cut} = [\,]$;
    $W = COUNT(10 \leq s_f \geq 0)$;
    $s_{cut}.append(W)$
  end
  $final_{cut} = max(s_{cut}) + std(s_{cut})$;
  $LRBT.append(final_{cut})$;
end
$\mathcal{I}_M \leftarrow$ create mask for $\mathcal{I}$ using four (left, right, top, bottom (LRTB)) cut values one for each side

cut value for each step on all four sides of the scene [see Fig. 9(b)].

a) *Bidirectional k-samples method:* In a bidirectional *k*-samples method, *k* number of samples (here, *k* = 3500 (represented by yellow, blue, and red transacts shown in Fig. 10), and *k* must not be larger than the respective dimension of the selected scene) where each transact/sample of length 2000 pixels were selected from all four sides of the scene as shown in Fig. 10(a). In the first step, the samples were filtered out if they do not intersect with the potential border noise region, i.e., red-line samples in Fig. 10(a). In this case, a threshold of 10 was selected to determine the width of the border noise. In the next step, the algorithm tries to find the width (W) of the border noise in each sample and saves it in an array (s_{cut}) as shown in Fig. 10(b). Finally, the width of the border noise (or number of columns and row to be removed from each side) is calculated based on the following expression: max (s_{cut}) + std (s_{cut}). The detailed algorithmic steps are given in Algorithm 2.

b) *Bidirectional all-samples method*: The bidirectional *k*-samples method was further modified in order to save more valid pixels in the image. In the bidirectional *all*-samples method, instead of tak-ing *k* samples, it exploits all samples each of length 2000 from all four sides. In this method, as it is shown in Fig. 9(b), an optimized smooth-ing procedure is used to better identify the bor-der noise, which shows systematically a kind of harmonic behavior; i.e., the smoothing was applied on backscatter differences between adjacent pixels instead of backscatter itself. The detailed algorithmic steps are given in Algorithm 3, and Fig. 11(a) and (b) shows the graphical representation of the proposed method.

Algorithm 3: Bi-directional all–samples method.

Data: $\mathcal{L} \leftarrow$ List of Sentinel-1 scenes each of size $m \times n$
Result: $\mathcal{I}_M \leftarrow$ Binary mask for border noise removal

for $\mathcal{I} \leftarrow 1$ *to* $\mathcal{L}$ **do**
  $rows,\ cols = \mathcal{I}.shape$;
  $\mathcal{S} \leftarrow \{S_L, S_R, S_B.S_T\}_{n=all,\ len=2000} \in \mathcal{I}$;
  for $s \leftarrow 1$ *to* $\mathcal{S}$ **do**
    /* Difference between adjacent array elements i and $i+1$ (see Figure 11 (A)) */
    $diff \leftarrow s[i+1] - s[i]$;
    /* Smoothing $diff$ by averaging adjacent array elements i and $i+1$ (see Figure 11 (b)) */
    $M_{diff} \leftarrow (diff[i+1] + diff[i])/2$;
    /* Get the index of M_{diff} where values are > 10 */
    $nodata\_idx \leftarrow where(M_{diff} > 10)$;
    $faulty\_idx = (\mathcal{I}[:,0] > 50)\ \&\ (nodata\_idx \neq 0)$;
    $nodata\_idx[faulty\_idx] = 0$;
    $nodata\_idx = median\_filter(nodata\_idx,\ size = 51)$;
    $idx\_differ = nodata\_idx[i+1] - nodata\_idx[i]$;
    /* Find the three biggest steps as shown in Figure 9 (b) */
    $steps = [0,\ ]$
    for $steps \leftarrow 1$ *to* $sorted(idx\_differ)[-3:]$ **do**
      $step\_idx = where(idx\_differ == step)[0]$;
      if $length(step\_idx) == 1)\ \&\ (step > 5)$ **then**
        $steps.append(step\_idx[0] + 1)$;
      else
        Pass
      end
    end
    $steps.append(rows)$;
    $steps.sort()$;
    /* Set a noise width (n_width) for each step */
    for $st \leftarrow 1$ *to* $range(length(steps) - 1)$ **do**
      $n\_width = percentile(nodata\_idx[steps[st] : steps[st+1]],\ 99.8)$;
      $n\_width = n\_width + 1$;
      if $n\_width \leq 2$ **then**
        $n\_width = 0$;
      else
        $n\_width + buffer$;
        /* Default: $buffer = 5$ */
      end
      $nodata\_idx[steps[st] : steps[st+1]] = n\_width$;
      /* $numpy.tile(A, reps)$: Construct an array by repeating A the number of times given by $reps$. */
      $index\_arr = np.tile(range(cols), (rows, 1))$;
      $step\_out = index\_arr < nodata\_idx.reshape(rows, 1)$
    end
    $\mathcal{I}_M \leftarrow$ Create binary mask for each side and based on noise width for each step given by $step\_out$
  end
end

B. Implementation

The newly developed Sentinel-1 border noise removal module has been integrated within the SAR Geophysical Retrieval Toolbox (SGRT), which is being used for the large-scale data processing and product generation. The SGRT is a software package developed by the Vienna University of Technology (TU Wien) for extracting geophysical parameters from SAR data [13]. Version 2.3.0 of the SGRT is an adaptation to Sentinel-1 of the earlier SGRT 1.0 developed for ENVISAT Advanced Synthetic Aperture Radar (ASAR) data, incorporating optimizations intended for handling the considerably higher spatial resolution and resulting explosion in data volumes foreseen of Sentinel-1 relative to ENVISAT ASAR. The processing toolbox of SGRT 2.3.0 consists of three main components, namely: preprocess-ing, analytics, and production. It was implemented in the Python programming language.

C. Evaluation Dataset

An intensive evaluation of the proposed algorithms was per-formed by applying the methods to a random sample of 1500 scenes, which was selected from the Earth Observation Data Center[3] date archive of 324 097 Sentinel-1 GRD Level-1 prod-ucts Fig. 12(a). Fig. 12(b) shows the distribution of the selected sample.

D. Performance Evaluation Criteria

The numbers used for the performance and accuracy percentage are based on the counts of *positive* and *negative* alarm;
where
positive alarm = border noise successfully removed
negative alarm = border noise partially or not removed.
In order to flag a scene as *positive* or *negative* alarm, a color composite [as shown in Fig. 12(c)] of *before* and *after* the border noise removal was created for each scene. These color compos-ites were then manually checked and flagged. Fig. 12(c) shows an example of a dataset that was used to determine the perfor-mance of each method.

III. Results and Discussion

For this study, we have evaluated existing solutions for Sentinel-1 border noise removal. Due to the inconsistent performance and lack of robustness when tested with large number of scenes, we have proposed new solution to this problem. The evaluation results of the existing and newly developed Sentinel-1 border noise removal methods are discussed in the following section.

A. SNAP "Sentinel-1 Remove GRD Border Noise"

In order to handle the border noise in SNAP, it is very important to tune the parameters of "Sentinel-1 Remove GRD Border Noise" module in SNAP, which is feasible for few scenes where parameters can be tuned for each individual scene one-by-one. However, doing so for batch processing where the number of scenes are in 100 s or 1000 s is practically impossible. Even in case of a single scene, tuning of input parameters does not work 100% especially over sea where the backscatter values are low.

White rectangle in Fig. 13(a) shows an example where the border noise is not fully removed but a gap between the clean pixels

[3] https://www.eodc.eu/

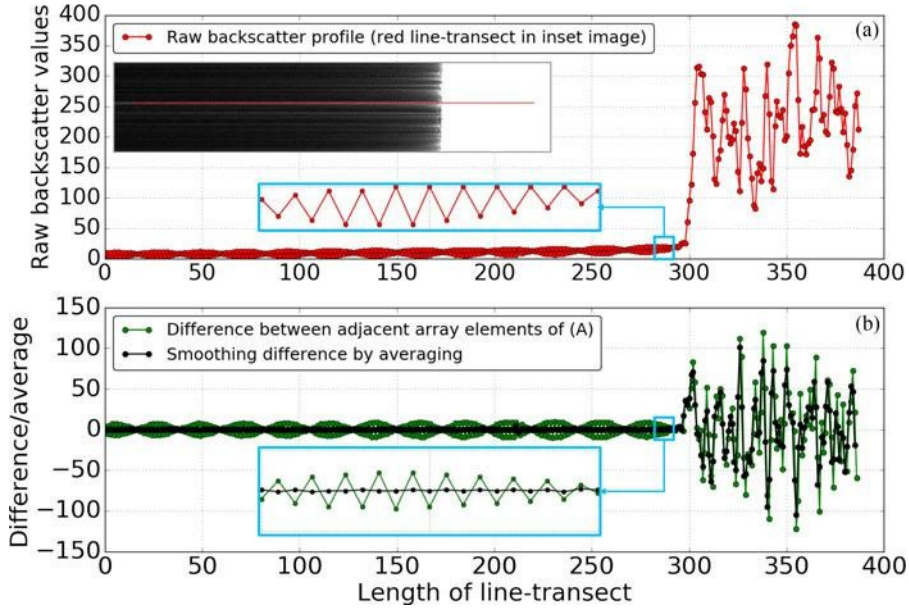


Fig. 11. Bidirectional *all*-samples method: (a) raw backscatter profile of red line-transect in inset image, and (b) difference and smoothing of aforementioned (a) profile.

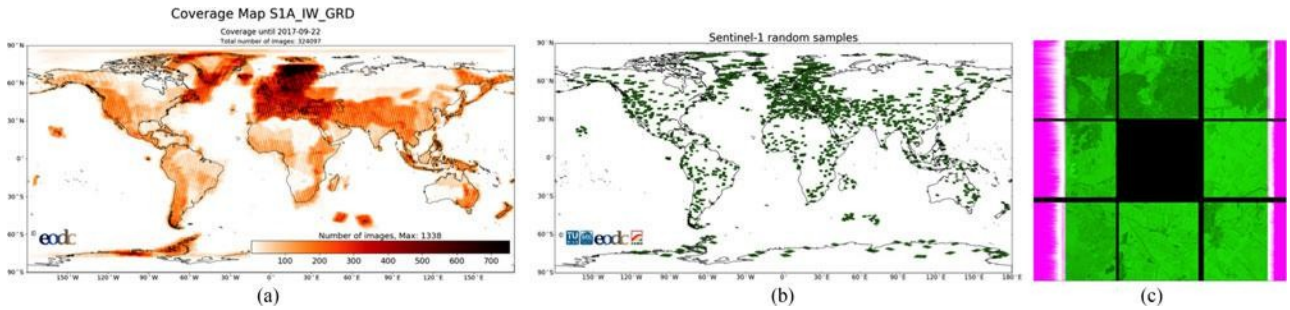


Fig. 12. Evaluation dataset: (a) Sentinel-1A IW GRD global coverage map, (b) random sample of 1500 scenes selected for the evaluation of proposed algorithms, (c) this is the type of color composite generated for the evaluation of the results of all methods for each input scene. The eight boxes (each of size 1500 × 1500 pixels) show the four corners and the middle part of each side (left, right, top, and bottom). And *green* color represents the final parts of the image after border noise removal, *magenta* shows the border noise removal mask, and *Black* color represents the background.

and the border noise margin still exists. Furthermore, over sea (or open water) area valid pixels are masked as border noise, which result in a noise granularity like effect as shown in yellow rectangle in Fig. 13(a).

In the first evaluation case, SNAP module ("Sentinel-1 Remove GRD Border Noise") was run for 500 Sentinel-1 GRD Level-1 scenes in order to fix the border noise. The computation time was ranging between 11 and 15 s. During the validation step, it was found that the output of the "Sentinel-1 Remove GRD Border Noise" module is highly unstable and inconsistent. An error rate (number of times algorithm failed to completely remove the noisy pixels) of 68.96% was found, and based on this result, further analysis of SNAP module on the remaining 1000 Sentinel-1 samples was not continued. The failure of SNAP module over ocean was in general very consistent. Following is a list of inconsistencies and failures of border noise removal observed during the evaluation:

1) border noise partially or completely missed both in range and azimuth;

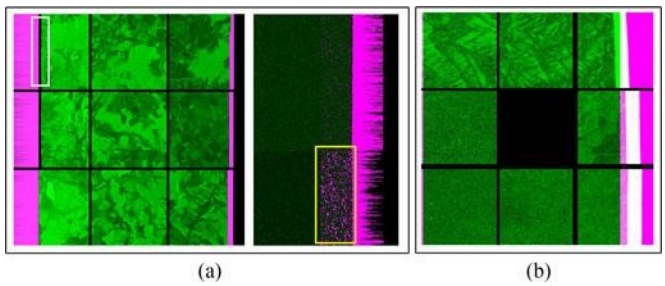


Fig. 13. (a) Example of "Sentinel-1 Remove GRD Border Noise" module in the SNAP toolbox: (left): missing noise pixels along the border (where the eight boxes (each of size 1500 × 1500 pixels) show the four samples from the corners and the middle part of each side of original image (left, right, top, and bottom), where *green* color represents the valid pixels of the image after border noise removal and *magenta* shows the border noise removal mask; (right): missclassification of valid pixels. (b) Most common scenario where the anomaly detection method failed frequently. The eight boxes (each of size 1500 × 1500 pixels) show the four corners and the middle part of each side. And *green* color represents the final image after border noise removal, *magenta* shows the border noise removal mask, and *black* color represents the background.

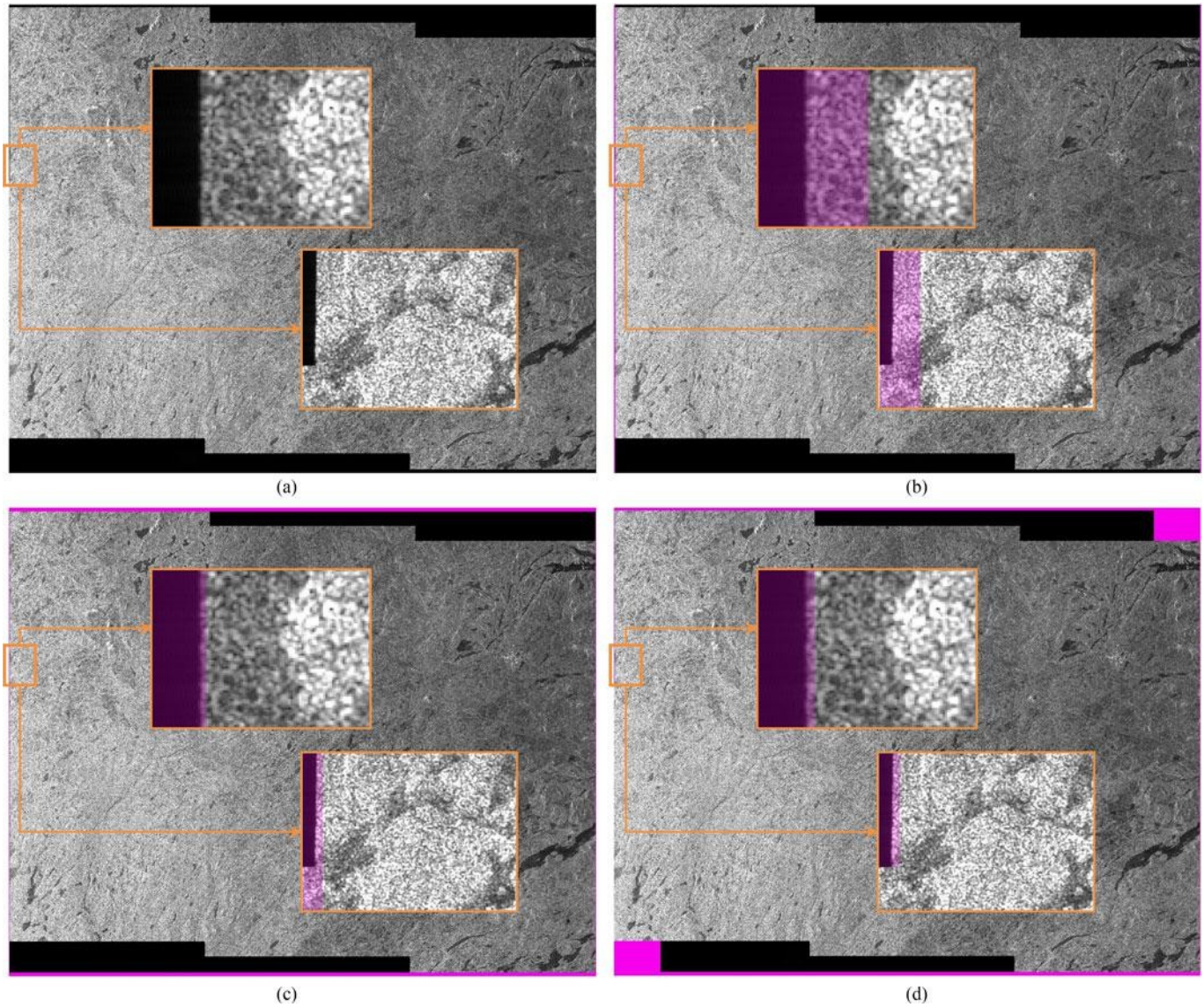


Fig. 14. Examples of the proposed algorithms for Sentinel-1 border noise removal. (a) Original Sentinel-1 GRD scene (S1A_IW_GRDH_1SDV_20141218T050411_20141218T050441_003769_0047F3_0525), (b) anomaly detection method, (c) bidirectional *k*-samples method, and (d) bidirectional *all*-samples method. *Magenta* color (with 75% transparency) is the output mask produced by each method. Zoomed part (orange box) highlights the comparison of all methods over the same portion of the scene. Here, bidirection means the border noise is removed from both left/right and top/bottom sides.

2) granular-like noise observable over ocean as shown in Fig. 13(a-right);
3) remaining noise gaps between valid pixels and the masked region as shown in Fig. 13(a-left);
4) due to the inconsistency in threshold selection in SNAP, the probability of removing more valid pixels is high.

B. Anomaly Detection

In terms of computation time, the anomaly detection method (see Table I) is significantly faster than all other methods including the SNAP's "Sentinel-1 Remove Border Noise" module. During the evaluation process, it was observed that the anomaly detection method works better with higher accuracy if the width of the border noise is consistent along the side. If there is smooth transition from wide border noise to short border noise, then this will result in a partial removal of the border noise. As shown in Fig. 13(b), where on the right side of the scene, the width and orientation of border noise is not consistent. At the right-top side the border noise width (low-values + no-values; as described in

TABLE I
PERFORMANCE COMPARISON OF ALL METHODS

Method	Sample Size	Negative Alarms	Positive Alarms	Computation Time (s)	Error (%)
Anomaly detection	1500	118	1 382	03.00	7.86
Bidirectional *k*-samples	1500	1	1 499	15.83	0.06
Bidirectional *all*-samples	1500	0	1500	14.72	0.00

Where, sample size is the number of Sentinel-1 scenes randomly selected for this experiment and positive/negative alarm refer to the case when the border noise is completely removed (positive alarm) or the border noise did not or partially removed (negative alarm).

Fig. 4) is greater than the noise width at the right middle and right bottom of the scene. This is due to the fact that this smooth curve like pattern of the border noise will also create a smooth transition curve in the calculated column mean plot. In this case, the width of the border noise using the Tukey test also depends on the value of k (see Algorithm 1) whose value is manually tuned. Therefore, these kind of ambiguities will still be problematic in some cases. Another limitation of this approach is that most of the times it creates a large mask-cut value and remove the valid pixels as well along with the border noise. For example, on the left side in Fig. 13(b), there is no border noise but it still removes a buffer of some valid (or noise free) pixels. Fig. 14(a) shows a sample Sentinel-1 IW GRD scene, which is selected for the visual demonstration of three methods. Fig. 14(b) shows an example where anomaly detection method has removed the valid pixels. The width of this border noise region (referred as *low-value zone* in Fig. 4) has a fluctuating behavior and vary from scene-to-scene and mainly depends on land-cover type.

C. Bidirectional Sampling Methods

Our bidirectional sampling methods are specifically designed to address this issue of the border noise in the Sentinel-1 GRD Level-1 product. These methods have been intensively validated using large and well-distributed samples in order to achieve the best performance and optimization especially in terms of preserving the valid pixels.

The bidirectional *k*-samples method produced a border noise mask based on one linear cut for each side, therefore, the scene after applying the mask will have the rows and columns of same length as illustrated in Figs. 9(a) and 14(c). The Bidirectional *all*-samples method follows the same rule if there is no big shift/step in the border noise pattern. And if there is/are step(s) along the border noise, then there will be one mask-cut value for each step. In this case, the rows and columns of final im-age can be of different length. The zoomed parts of Fig. 14(c) and (d) show that a bidirectional *all*-samples approach is very well optimized in terms of saving the valid pixels, where the mask precisely follow the pattern of the border noise instead of removing the valid pixels.

Both these methods are robust and adaptive to new ambigu-ous border noise patterns, and the computation time is approxi-mately similar to the "Sentinel-1 Remove Border Noise" module in SNAP. The proposed approach is not only effective over land areas where the signal level is always higher than the instrument noise floor, but also over ocean where the situation may be op-posite under low wind conditions. Table I gives a comparative performance overview of the tested algorithms.

Finally, the bidirectional all-samples approach as the best-performing algorithm has been selected and integrated in SGRT and is being used for the large-scale data processing.

IV. Conclusion

Reprocessing of satellite data because of noise removal fail-ures becomes challenging during the processing of large amount of data or the near-real-time data production with limited com-puting resources. For large-scale data processing and opera-tional product and service development, it is very crucial to have a stable and reliable preprocessing chain. In case of the Sentinel-1 Level-1 GRD product, this aspect becomes even more impor-tant due the large file size and the required processing time. In this paper, we have evaluated different methods to fix the Sentinel-1 border noise using a large sample of Sentinel-1 data. Among the proposed methods, we selected the bidirectional *all*-samples. This method shows the best performance on scenes covering the land surfaces as well as the scenes that are partially over ocean or inland water surfaces, which is one of the ma-jor limitations of the "Sentinel-1 Remove GRD Border Noise" module in the SNAP toolbox. This algorithm has been inte-grated in the SGRT preprocessing chain, which uses the SNAP toolbox at the back-end for the Sentinel-1 data preprocessing. This will enable automatic Sentinel-1 data preprocessing for real-time or near real-time services, in which large-scale high-quality artifact-free mosaics should be delivered for operational purposes.

Acknowledgment

TThe computational results presented have been achieved using the Vienna Scientific Cluster (VSC).

References

[1] P. Potin, "Sentinel-1 mission overview," in *4th Advanced Course on Radar Polarimetry*, ESA—ESRIN, Frascati , Italy, Feb. 2, 2017.

[2] J. Mittermayer *et al.*, "TOPS Sentinel-1 and terraSAR-x processor compar-ison based on simulated data," in *Proc. 8th Eur. Conf. Synthetic Aperture Radar*, Jun. 2010, pp. 1–4.

[3] J. W. Park, A. A. Korosov, M. Babiker, S. Sandven, and J. S. Won, "Ef-ficient thermal noise removal for Sentinel-1 TOPSAR cross-polarization channel," *IEEE Trans. Geosci. Remote Sens.*, vol. PP, no. 99, pp. 1–11, doi: 10.1109/TGRS.2017.2765248.

[4] Collecte Localisation Satellites, "Sentinel-1 product specification," Col-lecte Localisation Satellites, Tech. Rep. 1, CLS-2016a, Reference: MPC-0243, Jun. 2015.

[5] Collecte Localisation Satellites, "Sentinel-1 level 1 detailed algorithm definition," Collecte Localisation Satellites, Tech. Rep. 2, CLS-2016b, Document No.: SEN-TN-52-7445, Feb. 2016.

[6] N. Yague-Martinez, F. D. Zan, and P. Prats-Iraola, "Coregistration of interferometric stacks of Sentinel-1 tops data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1002–1006, Jul. 2017.

[7] F. Nunziata, A. Buono, M. Migliaccio, and G. Benassai, "Dual-polarimetric C- and X-band SAR data for coastline extraction," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 11, pp. 4921–4928, Nov. 2016.

[8] M. Hornacek *et al.*, "Potential for high resolution systematic global surface soil moisture retrieval via change detection using Sentinel-1," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1303–1311, Aug. 2012.

[9] I. Ali, V. Naeimi, S. Cao, S. Elefante, B. Bauer-Marschallinger, and W. Wagner, "Sentinel-1 data cube exploitation: Tools, products, ser-vices and quality control," in *Proc. Big Data Space*, 2017, pp. 40–43, doi: 10.2760/383579.

[10] Y. Kim and J. J. van Zyl, *Synthetic Aperture Radar Polarimetry*. New York, NY, USA: Wiley, Dec. 2011.

[11] Collecte Localisation Satellites, "Masking "no-value" pixels on GRD products generated by the Sentinel-1 ESA IPF," Collecte Localisation Satellites, Tech. Rep. 3, CLS-2015, 2016.

[12] J. W. Tukey, "Comparing individual means in the analysis of variance," *Biometrics*, vol. 5, no. 2, pp. 99–114, 1949.

[13] V. Naeimi, S. Elefante, S. Cao, W. Wagner, A. Dostalova, and B. Bauer-Marschallinger, "Geophysical parameters retrieval from Sentinel-1 SAR data: A case study for high performance computing at EODC," in *Proc. 24th High Performance Comput. Symp.*, 2016, pp. 10-1–10-8.

Iftikhar Ali received the B.S. degree in mathematics from Bahauddin Zakariya University, Multan, Pakistan, in 2008, the M.Sc. degree in geodesy and geoinformation science (major in remote sens-ing) from the Technical University of Berlin, Berlin, Germany, in 2011, and the Ph.D. degree in remote sensing from the University College Cork, Cork, Ireland, in March 2016.

From June 2011 to June 2012, he worked as a Remote Sensing Researcher with the Institute for Earth Observation, EURAC, Bolzano, Italy, and the Technical University of Berlin. He also worked as a Visiting Student Researcher with Jet Propulsion Laboratory, NASA, Pasadena, CA, USA. He is currently a Postdoctoral Researcher with the Microwave Remote Sensing Research Group, Vienna University of Technology, Wien, Austria. His research interests include rada interferometry, remote-sensing-based biophysical parameters retrieval, and application of machine learning algorithms for remote-sensing-based parame-ters retrieval.

Senmao Cao received the B.Sc. degree in surveying and mapping engineering from the Xi'an University of Science and Technology, Xi'an, China, in 2008, and the M.Sc. degree in remote sensing of environ-ment and resources from the Spatial Information Re-search Center of Fujian, Fuzhou University, Fuzhou, China, in 2011. His master study focused on feature extraction and classification for remote sensing im-agery, specializing in semisupervised learning.

From 2011 to 2013, he worked for Beijing Earth-View Image, Inc., as a Software Developer, where he was responsible for the development of remote sensing image processing tools, GIS database design, and performance optimization. Since May 2014, he has been working in the Remote Sensing Research Group, Department of Geodesy and Geoinformation, Vienna University of Technology, Vienna, Austria, con-centrating on the development of SAR processing chain. His research interests include soil moisture retrieval, flood mapping, and scientific programming. He specializes in image processing, scientific programming, and data analysis.

Vahid Naeimi received the Ph.D. degree in microwave remote sensing from the Vienna University of Technology (TU Wien), Vienna, Austria, in 2009.

He has been working with the Institute of Photogrammetry and Remote Sensing, TU Wien, since 2005, as a Research Scientist. From 2010 to 2013, he was with German Remote Sensing Data Centre, German Aerospace Centre. He is currently a University Assistant with the Department of Geodesy and Geoinformation, TU Wien. His main research interests include geophysical parameters retrieval from remote sensing data especially active sensor (SAR and Scatterometers). He has participated in several national and international projects and actively involved in scientific algorithms development, prototyping, processing chains implemen-tation, and high-performance computing.

Christoph Paulik was born in Vcklabruck, Austria, on January 24, 1985. He received the Dipl.-Ing. de-gree in geodesy and geoinformation science from the Vienna University of Technology, Vienna, Austria, in 2009.

He has six years of experience in working with microwave remote sensing data with a focus on bulk data processing and software development. He is responsible for the validation of the soil moisture products in the current Copernicus Global Land Service and also a Lead Developer of the operational near real-time processing chains. He is also the Lead System Engineer of the processing system for the C3S (C3S 312a) Lot-7 Soil Moisture project.

Wolfgang Wagner (M'98–SM'07) was born in Austria, in 1969. He received the Dipl.-Ing. degree in physics and the Dr.techn. degree in remote sens-ing from the Vienna University of Technology (TU Wien), Wien, Austria, in 1995 and 1999, respectively.

In support of his Masters and Ph.D. studies, he received fellowships to carry out research at the University of Bern, Atmospheric Environment Service Canada, NASA, ESA, and the EC Joint Research Centre. From 1999 to 2001, he was with German Aerospace Centre. In 2001, he joined the TU Wien, where since 2012, he has been the Head of the Department of Geodesy and Geoinformation. He is the Cofounder and Head of the Science of the EODC Earth Observation Data Centre. His main research interest include gain physi-cal understanding of the mechanisms driving the interaction of electromagnetic waves with the land surface. Based on this understanding, he has developed models for retrieving soil moisture, biomass, and other land surface variables from scatterometer, SAR, and full-waveform lidar observations.

Dr. Wagner is a Member of the EUMETSAT/ESA Science Advisory Group for METOP-SG SCA, and since 2016, has been the Chair of the GCOS/WCRP Terrestrial Observation Panel for Climate. From 2008 to 2012, he was the ISPRS Commission VII President, and from 2009 to 2011, the Editor-in-Chief of the Open Access Journal *Remote Sensing*. He was the recipient of the ISPRS Fred-erick J. Doyle Award for his scientific contributions in active remote sensing.