TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Valentin Stennikov 164428IAPB

A TOOL FOR ASSESSMENT OF FITNESS OF SENTINEL-1 DATA FOR INSAR ANALYSIS OF USER SELECTABLE AREAS OF INTEREST

Bachelor's thesis

Supervisor: Andreas Kiik (MSc)
Co-supervisor: Juhan-Peep Ernits (PhD)

Tallinn 2019

Author's declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Valentin Stennikov

26.08.2019

Abstract

Presented work is based on a problem of low signal intensity in different areas around the world, which are being processed using InSAR methodology. In order to lower the amount of work done in regions, where results can contain low amount of data about ground displacement, or the displacement points are of low coherence, a way to assess signal intensity in areas of interest is required.

The result of this work is a program, which grants ability to select a location on a map, for which signal intensity analysis must be performed. Created tool automatically finds imagery, covering an area of interest, and processes image pixel values in the region. Signal intensity statistics is presented on a time scale, allowing to see seasonal changes in reflected signal strength as well as general amount of data available for the geographical location.

This thesis is written in English and is 17 pages long, including 5 chapters, 7 figures and 0 tables.

Annotatsioon

Sentinel-1 andmete InSAR analüüsi jaoks sobivuse hindamise tööriist kasutaja valitavatele aladele.

Käesolev töö põhineb madala signaali intensiivsuse probleemil, mis tekkib erinevatel maakera regioonitel, mille püüakse analüüsida kasutades InSAR metodoloogiat. Selleks, et väheneda töömahu suunatud maaala objektide analüüsimiseks, kus analüüsi tulemuseks tekib väike arv maa liikumisest või töötlemise tulemused on madala koherentsusega, on vaja viisi, kuidas hinnata peegeldatud radari signaali intensiivsuse huvipakkuvas alas.

Töö tulemuseks on programm, mis annab võimaluse valida maaala kaardi peal, mille signaali intensiivsuse hakatakse uurima. Loodud tarkvara automaatselt kogub satelliidi andmeid, mis kattuvad valitud regiooni, ja töötleb piltide pikslite väärtused selles alas. Signaali intensiivsuse statistikat on visualiseeritud aja skaala peal, kus saab vaadata peegeldatud signaali tugevuse muutmist aastade jooksul ning üldine info andmete kättesaadavusest analüüsitud piirkonnas.

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 17 leheküljel, 5 peatükki, 7 joonist, 0 tabelit.

List of abbreviations and terms

ESA	*European Space Agency*
SAR	*Synthetic-aperture radar*
InSAR	*Interferometric synthetic-aperture radar*
GRD	*Ground Range Detected*
SLC	*Single Look Complex*
GeoTIFF	*Geo Tagged Image File Format*
SAFE	*Standard Archive Format for Europe*
CSS	*Cascading Style Sheets*
MVC	*Model-View-Controller*
OSM	*OpenStreetMaps*
SciHub	*Copernicus Open Access Hub*
IW	*Interferometric Wide swath*
EW	*Extra Wide swath*
VV	*Vertical-Vertical polarisation*
VH	*Vertical-Horizontal polarisation*
GCP	*Ground Control Point*
WGS84	*World Geodetic System 1984*
GDAL	*Geospatial Data Abstraction Library*
GB	*Gigabyte*
MB	*Megabyte*

Table of contents

List of figures

1 Introduction

SAR (*Synthetic-aperture radar*) images contain satellite radar signal intensity values and phase information. Combining these with other data, such as time required for radar signal to return, as well as comparing one image data with other images covering the same ground area allows to detect ground movement, potentially providing insight into different events which have occurred, are occurring or might occur in the future in the analysed region.

1.1 Problem

Due to different reasons, extracting a trustworthy ground displacement information from a series of SAR images can be problematic [1]. This can potentially make InSAR (*Interferometric Synthetic-aperture radar)* analysis results unsatisfying, either requiring more resources to be spent in order to achieve a better result or dismissing initial results and wasting work done creating analysis. One of those reasons is low signal intensity in the area of interest, in case signal from SAR-enabled satellite fails to reflect back to it, resulting in lower amount of produced displacement points, which are used for detecting ground movement, or displacement point low coherence [2]. Because of the large amount of data produced by satellites, it may be ineffective to check provided imagery manually and an automatic solution is required.

1.2 Goals

1. Create a program, which would allow to select a region of interest for which SAR imagery intensity analysis will be performed.

2. Implement process for analysing satellite the intensity of the images.

3. Provide visualisation of signal intensity analysis results.

1.3 Structure

Chapter 2 describes data, which is used for intensity analysis as well as how it is distributed and obtained. It compares different versions of collected data, which is available for processing, provides reasons for image parameters choices and highlights the advantages of these choices when compared with their counterparts.

Chapter 3 contains technical implementation details of the created program. It consists of description of the technologies used in the project, typical workflow from starting the analysis to receiving image analysis results with both user view and backend processes for collecting and analysing satellite data, as well as main steps in image processing.

Chapter 4 provides an overview and result comparison of two different areas of interest for validation. Additionally, potential reasons are provided to explain differences in the result sets and to confirm analysis result validity for the region.

2 Analysis input data

SAR imagery is created using satellites, which are equipped with synthetic-aperture radar. Data is collected while satellite is orbiting over the region, using motion of the radar antenna to create an image of the scanned area [3]. Those images later become available via different services associated with corresponding satellites.

2.1 Satellite platform selection

For the purpose of the current work SAR images are taken from Sentinel mission satellites managed by the ESA (*European Space Agency*). Sentinel mission consists of several sub-missions (Sentinel-1, Sentinel-2, etc.), each equipped with its different types of instruments to perform their specific tasks. SAR technology is used in Sentinel-1 mission satellites. Sentinel-1 mission is composed of a constellation of two satellites, Sentinel-1A and Sentinel-2B, orbiting Earth on the same orbital plane, providing consistent coverage of area of interest with 6-day period (Figure 1) [4] [5]. Sentinel-1 mission data is distributed free of charge and is available in online archives shortly after observation time.
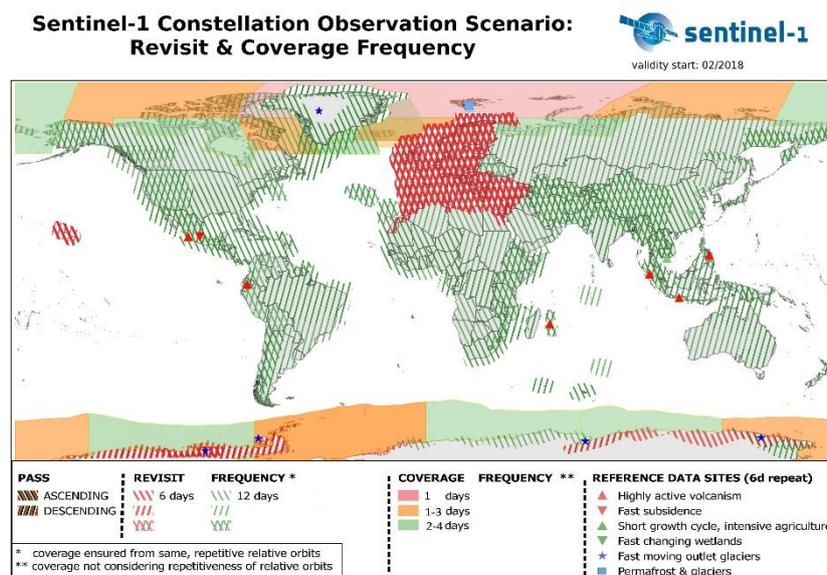

Figure 1. Sentinel-1 Constellation revisit frequency [4].

Sentinel-1 imagery can be searched and downloaded from a variety of different online resources specialising in mission data distribution. In order to avoid unnecessary work implementing multiple APIs (*Application Programming Interface*), only one data distributor is used: SciHub (*Copernicus Open Access Hub*). SciHub service provides free and open access to Sentinel mission satellite data, including both Sentinel-1 mission satellite imagery [6]. In addition to this, SciHub provides coverage for most of the world, which allows to perform signal intensity analysis of different regions of the Earth.

2.2 Image format

Sentinel-1 products are distributed in SAFE (*Standard Archive Format for Europe*) format, which contains product metadata as well as image data in binary format (Figure 2) [7].
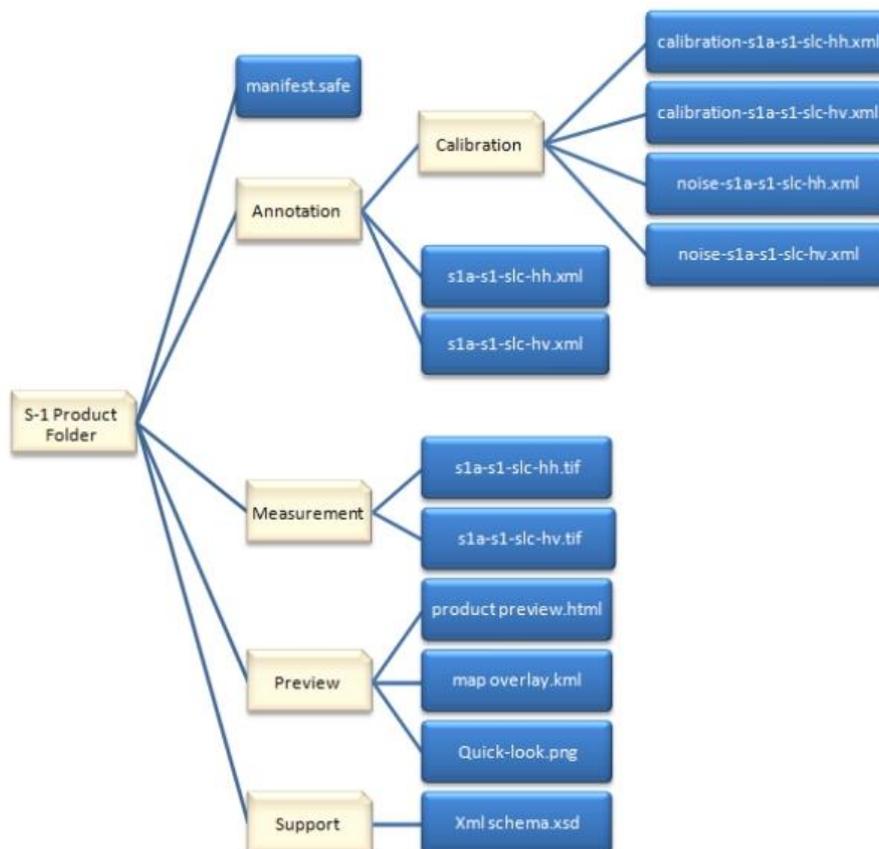

Figure 2. File structure for Level-1 SLC product in SAFE format [7].

Data products can be of multiple level: Level-0, Level-1 or Level-2. Level-0 product is raw data, which requires additional processing using SAR processor, resulting in a

Level-1 product. Level-2 products are derived from Level-1 products, which contain data derived from SAR data. Signal intensity data is contained in Level-1 products, which can be used for area of interest reflectivity analysis.

Level-1 products come in either SLC (*Single Look Complex*) or GRD (*Ground Range Detected*) product format. SLC products contain both phase information as well as signal intensity values [8]. GRD products are projected using an Earth ellipsoid model and multi-looked, resulting in a phase information loss, pixels with nearly square resolution and speckle reduction. Moreover, GRD products consist only of one band, while SLC product have multiple bands inside an archive [8]. Due to these differences, GRD generally have lesser file size than respective SLC product and are easier to process when working with signal intensity.

Preferred sensor operational mode is IW (*Interferometric Wide swath*), due to it being used for land and coastal areas, while EW (*Extra Wide swath*) is used primarily in polar regions (Figure 3) [4]. Products with IW sensor operational mode contain either single polarisation images VV (*vertical-vertical*) or dual polarisation images VV+VH (*vertical-horizontal*), but in any case image with VV polarisation is present.
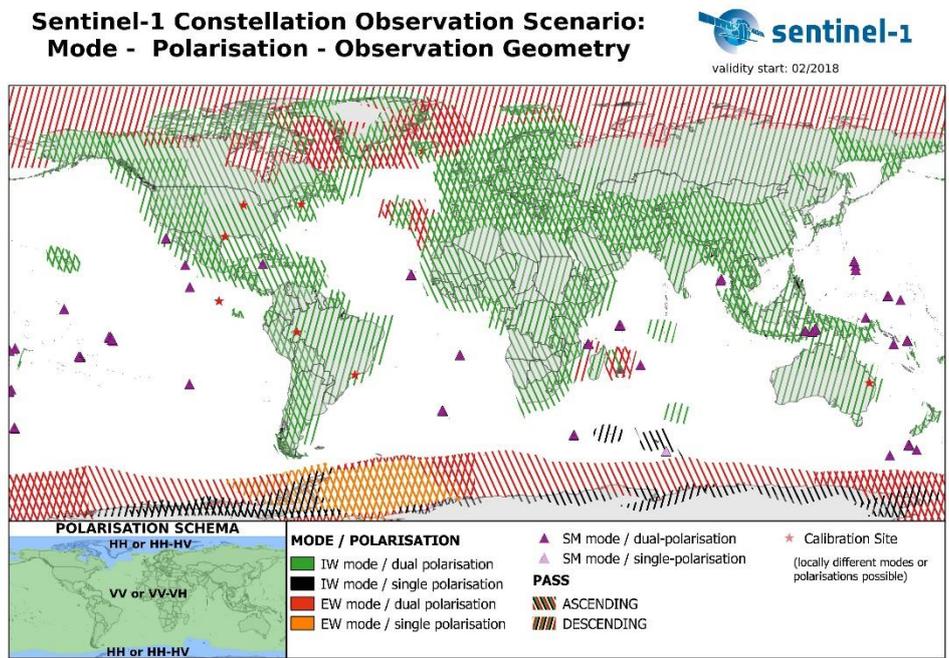

Figure 3. Sentinel-1 sensor operational mode and polarisation coverage [4].

For the purpose of this work products with the following parameters are used: Level-1 GRD products with IW sensor operational mode and VV polarisation. Moreover, only

GeoTIFF image files are required for performing intensity analysis. SciHub allows downloading parts of the product, effectively reducing required download size from approximate 1.5GB for a whole archived GRD product to 800-900MB of non-archived single GeoTIFF image (Figure 4), requiring less free space for data storage and speeding up download process.
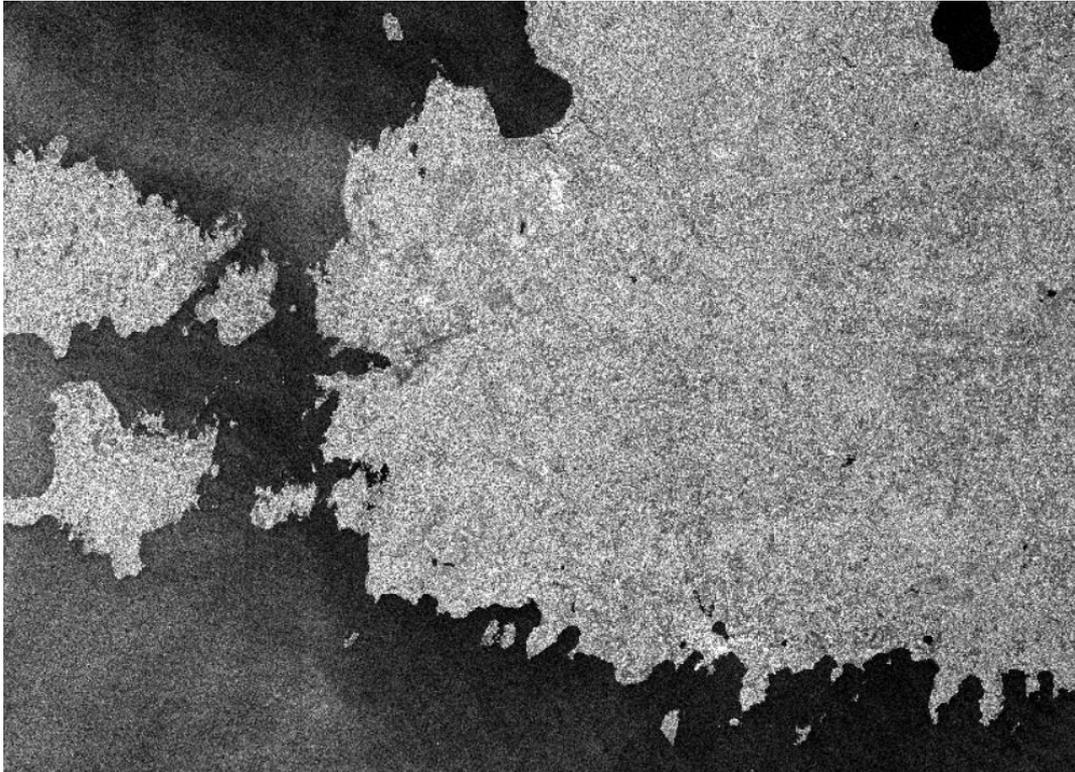

Figure 4. Level 1 GRD IW VV image, downloaded using SciHub. Ascending orbit.

3 Technical implementation

3.1 Technical information and used technologies

The implemented solution is a web application, which can be accessed via front end by a browser.

Backend part is implemented using Java 8 with Spring Boot Framework. Spring Boot provides MVC (*Model-View-Controller*) features, which are used for communication between the front end and the back end, as well as a convenient way to separate different program services by their layers and tasks, while still allowing to connect them together using dependency injection [9].

Interface for creating analysis and viewing analysis results is created using Thymeleaf [10]. Since application has only 2 primary views, Thymeleaf pages do not become hard to maintain and at the same time provide an easy way to connect with Spring Boot. Additionally, CSS (*Cascading Style Sheets*) and Bootstrap 4 framework [11] are used for creating custom page style. Extra interactivity is added using JavaScript with JQuery [12]. Map is created using Leaflet [13] with OSM (*OpenStreetMap*) tiles [14] and results are visualised using Chart.js [15].

PostgreSQL 10 [16] is used as a database because of its extension support, especially PostGIS [17], which is used when working with geometry and raster data types. In order to upload raster image data into the database, raster2pgsql utility is used [18]. Before importing an image into the database, it is processed using gdalwarp utility [19], which is a part of GDAL (*Geospatial Data Abstraction Library*) software library [20]. GDAL requires setting up PROJ_LIB environment variable in order to be able to transform geometries into a different projection, which is a necessary part of gdalwarp utility.

Application is deployed in Apache Tomcat 9.0 [21], which supports Java web applications. Application dependencies, build and deployment is handled by Maven [22].

3.2 Software requirements

Tool implementation must have following user interface functionality:

1. Creating analysis for any area of interest (U1).

2. Viewing previously created analyses (U2).

3. Viewing every available image intersecting with area of interest ordered by sensing date (U3).

4. Non-queued image can be added to download queue manually by selecting it through user interface (U4).

5. Viewing processed image results on a time scale to allow tracking signal intensity changes throughout the year (U5).

6. Y-axis for processed images can display multiple values. Options: mean, standard deviation, maximum, minimum (U6).

7. Images can be filtered by orbit direction. Options: ascending, descending, both (U7).

8. Available images are color-coded based on the state of an image (U8).

Tool implementation must have following image processing functionality:

1. Automatically search for all images intersecting with new area of interest (S1).

2. Images intersecting with new area of interest, are automatically grouped by time period and first image in each group is added to download and process queue (S2).

3. Automatic download for images placed in a download queue (S3).

4. Automatic signal intensity analysis of downloaded images (S4).

3.3 Main process workflow

The main process consists of two actors: user, who is interested in analysing satellite signal intensity in some area of interest, and a system, which collects, downloads and processes data covering user-selected region (Figure 5).

There are two main activities for user: creating new analysis (see Appendix 1) and viewing existing analysis results (see Appendix 2). When creating new analysis, user must specify geographical location, for which intensity analysis should be performed. After confirmation, system creates new analysis entry, searches for images, covering this location, and initiates download of queued Sentinel-1 products. When viewing existing analysis, user can add more images to download queue, if more data is required to assess signal reflectivity in the chosen area. Multiple revisits of the page may be required in case the user wants to see more data about some region, while there are more images still in the download queue.

System processes consist of image search done at the start of an analysis, parallel queued products download and image signal intensity analysis.
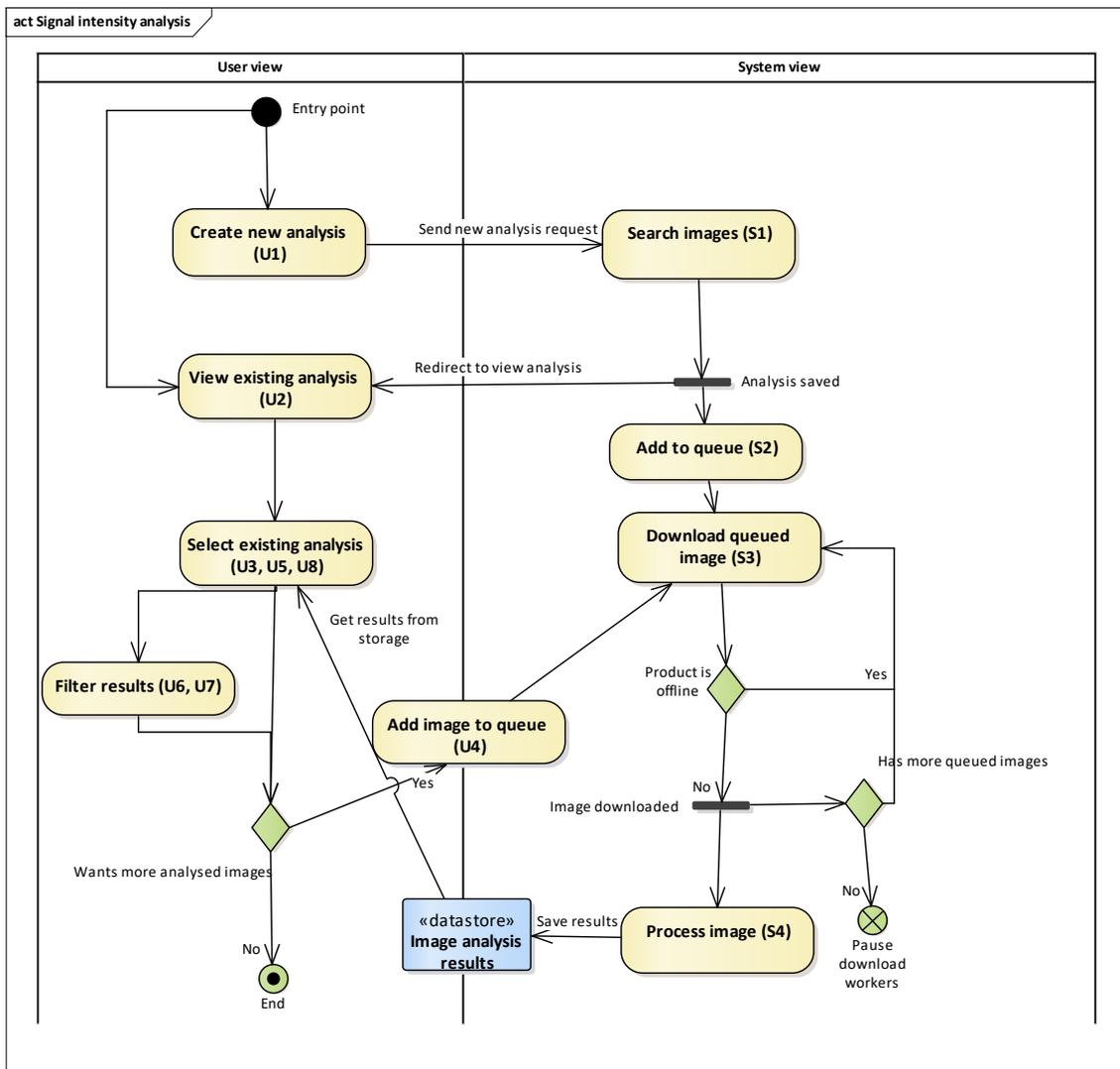
Figure 5. Signal intensity analysis main process workflow with requirements.

3.4 Image processing

3.4.1 Image search

The first step after creating initial analysis is searching for images, which intersect with the region of interest. In order to retrieve appropriate data, a search request is sent to the SciHub OData API with the following parameters:

1. Product type – GRD.

2. Platform name – Sentinel-1.

3. Intersection with analysis geometry.

4. Sensor operational mode – IW.

5. Polarisation – VV.

Because of the result pagination, multiple requests can be sent, if service has not retrieved all available image information from previous requests.

Image search results are merged with images already present in the database in order to apply actual image state, after which the images are grouped by year and month and groups are sorted by image sensing date. First image in each group is added to the download queue, if it is not already in the queue. Extra images can be added to download queue using the View Analysis page by clicking the download button on a non-downloaded image. This way the total amount of data required to be downloaded is drastically reduced and specific time periods can still be analysed by manually triggering image download.

3.4.2 Image download

Image download is performed by multiple separate threads, which extends base download class. Because of the restriction of only two simultaneous downloads from SciHub, a maximum of two SciHub download jobs are running at the same time. Each thread gets one of the images with status QUEUED or status RETRY, with priority being on the images in QUEUED state. If no images in any of these states are available, then download workers wait until more images will be queued for download.

QUEUED state images are images which were only added to download queue but has never been downloaded. RETRY state specifies images, where download worker tried to download product, but received HTTP Response Code 202, which means offline product in SciHub API. Offline products require extra download request: first request triggers moving product from offline storage to online storage, while the second request downloads product, if it has finished moving to online storage.

Once download is finished, image state changes to DOWNLOADED. In case of an error, image state is set to ERROR.

Images are downloaded into a file at a configurable location with filename format image_uuid.tiff

3.4.3 Image processing

Image processing is done using raster2pgsql and gdalwarp utilities and PostGIS raster and geometry functions. Both utilities are used with WGS84 (*World Geodetic System 1984*) for the main projection. Image import is done in a separate thread in order to remove long wait periods, while image is being imported into the database.

Gdalwarp utility is required to change downloaded image projection. Images from SciHub come with GCP (*Ground Control Point*) projection, which cannot be properly imported into database using raster2pgsql utility. Without image warping, imported result tiles contain invalid coordinates and cannot be matched with intersecting geometries. Following gdalwarp command is used for reprojection:

```
"gdalwarp -r bilinear -t_srs EPSG:4326 source_file_path output_file_path"
```

Raster2pgsql utility is used to import GeoTIFF images into the database table named raster_image. This table is overwritten each time new image is processed, which helps to remove old image data, for which analysis result has been generated before. Image is broken down into raster tiles with width 200 pixels and height 200 pixels. Using single raster without breaking image down into smaller chunks can result in an import error and out of memory exceptions, if image file size is big enough. Moreover, smaller tiles allow to find signal intensity values only from tiles intersecting analysis polygon. Intensity value is calculated using PostGIS st_summarystatsagg function grouped by analysis, for which image signal intensity is being processed, resulting in a mean intensity value of raster tiles with its standard deviation as well as minimum and maximum values of all tiles. Following raster2pgsql command is used for image import:

```
"raster2pgsql -s 4326 -d -C -M input_file_path -F -t 200x200 -I -e
public.raster_image | psql postgresql_database_url"
```

Calculated values are saved for each analysis and image pair, where analysis location intersects image location, after which new image data is imported.

3.5 Analysis result visualisation

In View Analysis page user can select analysis, which should be displayed. Object name and total number of images for object are displayed after analysis object selection. Image analysis results are also displayed in a form of chart containing two lines:

1. One line represents all available images for selected analysis object, except those, which were already downloaded. It shows image coverage of an object, and clicking on a point shows its state, sensing date and orbit direction. Image in state PENDING can also be added to download queue by clicking the download button. This line is always completely straight and in the middle of the chart, parallel to the X axis.

2. Second line shows image intensity analysis. Clicking on a point here opens more info about this image intensity. Y axis represents value, selected using select input above the chart.

Every image state is colour coded: grey = PENDING, green = DOWNLOADED, blue = QUEUED, yellow = DOWNLOADING, orange = RETRY, red = ERROR. Chart can be zoomed in to get a better look at the specific timeframe.

Chart can be filtered by orbit direction: ASCENDING, DESCENDING or both. This way the user can see if any of the orbits has worse signal reflection than the other.

4 Analysis result validation

For the purpose of result validation, two areas of interest were selected and signal intensity analysis. The first location chosen for validation is an urban area with a lot of highly reflective surfaces such as concrete, stone or glass and big number of straight angles between streets and buildings, acting as natural corner reflectors [23]. The second location is a region with a high level of vegetation, which can reflect satellite signal at an unexpected angle [2] [23]. Expected result is higher values of signal intensity for the images intersecting the first location, when compared to intensity values in the second location.

4.1 Case 1, Tallinn, Harjumaa, Estonia

A total of 767 images was found for this location (see Appendix 3). A total of 18 images were processed, 4 of which were from ascending orbit and 14 were from descending orbit.

Mean intensity values of both orbits tended to fall into 200-250 interval (Figure 6), with one outlier having mean intensity value of 181. Standard deviation values for a majority of the images were slightly higher than their respective mean values, suggesting that there are regions with worse signal reflectivity when compared to other parts of the same location. Minimum intensity values for a single image tended to remain in the 20-50 interval, while maximum intensity values went up to 65 000, hence strengthening the suggestion that reflected signal strength distribution is not homogenous.
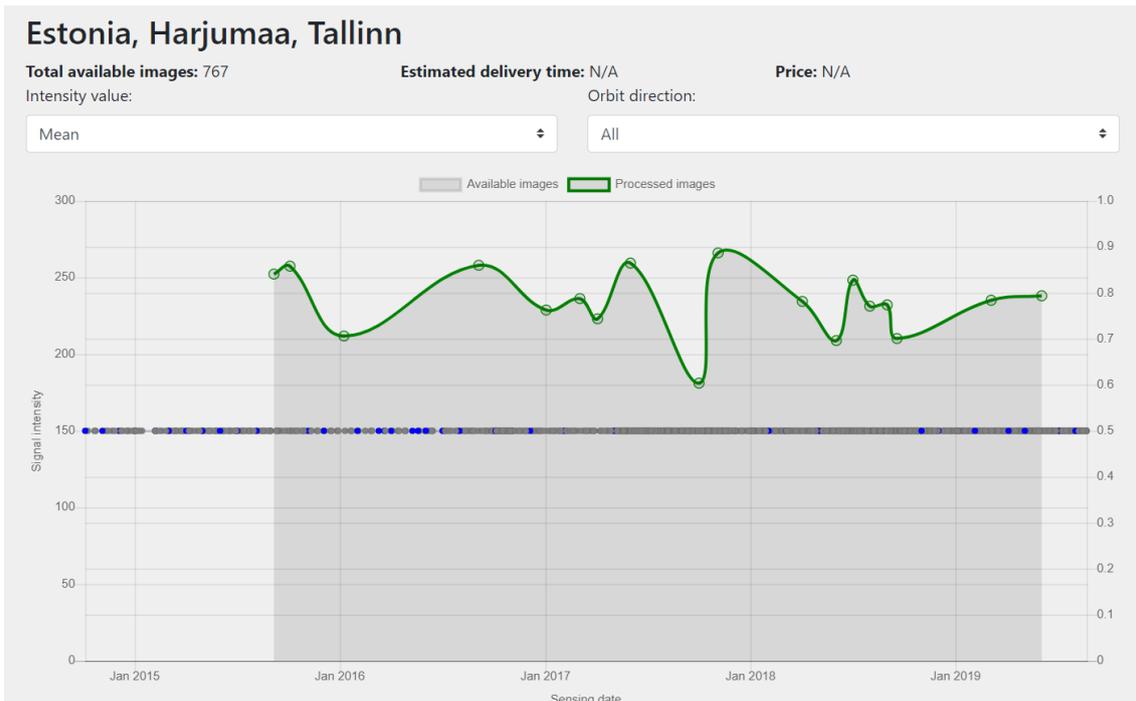
Figure 6. Signal intensity analysis for Tallinn, Estonia.

4.2 Case 2, Matsalu National Park, Estonia

A total of 804 images was found for this location (see Appendix 3). A total of 17 images were processed, 3 of which were from ascending orbit and 14 were from descending orbit.

Mean intensity values of both orbits were positioned in the 150-200 interval (Figure 7), with one outlier having a value of 244. Standard deviation values for a majority of processed images fell into 40-80 interval and were lower than their respective mean value, suggesting more homogenous intensity distribution in this area of interest, when compared with the first case. Minimum and maximum intensity values support that with falling into 10-50 and 1500-4500 intervals respectively.

Figure 7. Signal intensity analysis for Matsalu National Park, Estonia.

Achieved analysis results have met initial expectations. Tallinn analysis results had higher mean intensity values than the mean intensity values in Matsalu National Park area. Additionally, big gap between minimum and maximum intensity values as well as high standard deviation may suggest that some parts of Tallinn area may have worse reflectivity than others, resulting in lower coherence of displacement points calculated using InSAR methodology. A smaller selected area can be used to achieve more accurate results.

5 Summary

The goal of this work was to implement an automatic process, which would search and analyse signal intensity data presented in SAR images, presenting a way to assess if some area of interest can be analysed using InSAR methodology and provide the result with high amount of displacement points and high coherence of said points.

As a result, a web application was created, which takes an input from user, where a region of the Earth is specified, and searches corresponding Sentinel-1 products using SciHub API, downloads products to local storage and processes image data using PostGIS raster functions. Results of the signal intensity analysis is presented to the user on a line chart with time scale, where yearly changes in signal reflectivity can be seen, as well as additional images requested for analysis.

Result validation have confirmed that implemented process works correctly and completes all necessary steps to process image data and find signal intensity values for an area of interest. Provided test cases have met expectations when compared to each other.

References

[1] P. S. Agram, "PERSISTENT SCATTERER INTERFEROMETRY IN NATURAL TERRAIN," Stanford University, Stanford, 2010.

[2] S. Sinha, L. K. Sharma, C. Jeganathan, M. S. Nathawat, A. K. Das and S. Mohan, "EFFICACY OF INSAR COHERENCE IN TROPICAL FOREST REMOTE SENSING IN CONTEXT OF REDD," *International Journal of Advancement in Remote Sensing, GIS and Geography,* vol. 3, no. 1a, pp. 38-46, 2015.

[3] M. Kirscht and C. Rinke, "3D Reconstruction of Buildings and Vegetation from Synthetic Aperture Radar (SAR) Images," in *MVA*, Chiba, 1998.

[4] European Space Agency, "Sentinel Online: observation scenario," [Online]. Available: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/observation-scenario. [Accessed 22 August 2019].

[5] European Space Agency, "Sentinel Online: overview," [Online]. Available: https://sentinel.esa.int/web/sentinel/missions/sentinel-1/overview. [Accessed 12 August 2019].

[6] European Space Agency, "Copernicus Open Access Hub," [Online]. Available: https://scihub.copernicus.eu/userguide/. [Accessed 22 March 2019].

[7] European Space Agency, "Sentinel Online: SAR formats," [Online]. Available: https://earth.esa.int/web/sentinel/user-guides/sentinel-1-sar/data-formats/sar-formats. [Accessed 23 August 2019].

[8] European Space Agency, "Sentinel Online: Data Products," European Space Agency, [Online]. Available: https://sentinel.esa.int/web/sentinel/missions/sentinel-1/data-products. [Accessed 25 August 2019].

[9] P. Webb, D. Syer, J. Long, S. Nicoll, R. Winch, A. Wilkinson, M. Overdijk, C. Dupuis, S. Deleuze, M. Simons, V. Pavić, J. Bryant and M. Bhave, "Spring Boot Reference Guide," [Online]. Available: https://docs.spring.io/spring-boot/docs/current/reference/htmlsingle/. [Accessed 22 August 2019].

[10] "Thymeleaf," The Thymeleaf Team, [Online]. Available: https://www.thymeleaf.org/. [Accessed 25 August 2019].

[11] "Bootstrap," Bootstrap team, [Online]. Available: https://getbootstrap.com/. [Accessed 25 August 2019].

[12] "JQuery," The JQuery Foundation, [Online]. Available: https://jquery.com/. [Accessed 25 August 2019].

[13] V. Agafonkin, "Leaflet," Vladimir Agafonkin, [Online]. Available: https://leafletjs.com/. [Accessed 25 August 2019].

[14] OpenStreetMap, "OpenStreetMap," OpenStreetMap, [Online]. Available: https://www.openstreetmap.org/about. [Accessed 25 August 2019].

[15] "Chart.js," [Online]. Available: https://www.chartjs.org/. [Accessed 25 August 2019].

[16] The PostgreSQL Global Development Group, "PostgreSQL," The PostgreSQL Global Development Group, [Online]. Available: https://www.postgresql.org/. [Accessed 25 August 2019].

[17] OSGeo, "PostGIS," OSGeo, [Online]. Available: https://postgis.net/. [Accessed 25 August 2019].

[18] OSGeo, "Raster Data Managment, Queries and Applications," OSGeo, [Online]. Available: https://postgis.net/docs/using_raster_dataman.html. [Accessed 25 August 2019].

[19] OSGeo, "GDAL," OSGeo, [Online]. Available: https://gdal.org/programs/gdalwarp.html. [Accessed 25 August 2019].

[20] OSGeo, "GDAL," OSGeo, [Online]. Available: https://gdal.org/. [Accessed 25 August 2019].

[21] The Apache Software Foundation, "Apache Tomcat," The Apache Software Foundation, [Online]. Available: http://tomcat.apache.org/. [Accessed 25 August 2019].

[22] The Apache Software Foundation, "Apache Maven Project," The Apache Software Foundation, [Online]. Available: https://maven.apache.org/. [Accessed 25 August 2019].

[23] V. Aguado, A. Vink, B. James and E. Biescas, "An improved combination of Natural Reflectors and Corner Reflectors to monitor surface heave generated by SAGD operations using InSAR satellite technology," in *World Heavy Oil Congress*, Edmonton, Alberta, Canada, 2015.

Appendix 1 – Create analysis page
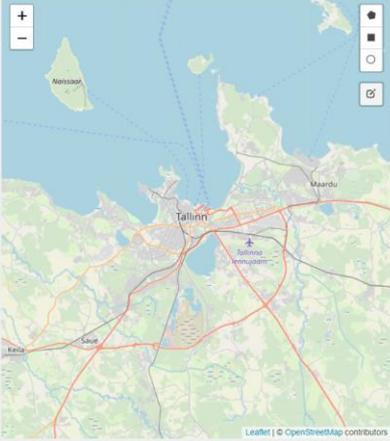
Appendix 2 – View analysis page
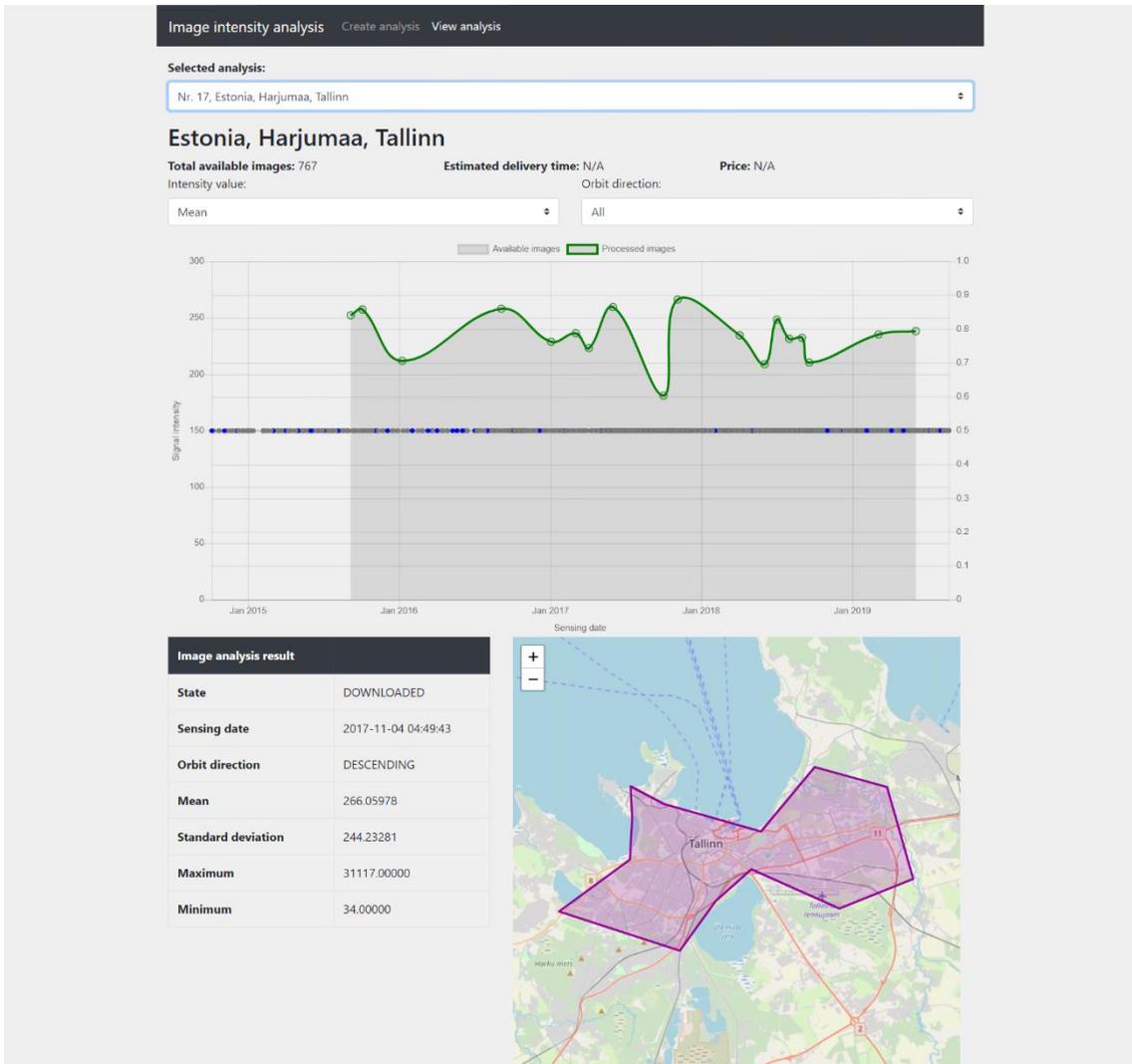
Appendix 3 – Performed analysis locations

```
Case 1. Estonia, Harjumaa, Tallinn. WGS84
{
      "type": "Polygon",
      "coordinates": [[
            [24.610199,59.405118], [24.721161,59.386661],
            [24.75412,59.41015], [24.787079,59.425242],
            [24.868378,59.406796], [24.936493,59.420771],
            [24.912323,59.46378], [24.845306,59.473268],
            [24.795868,59.44312], [24.706879,59.455964],
            [24.676117,59.464338], [24.677216,59.449264],
            [24.675018,59.429713], [24.610199,59.405118]
      ]]
}

Case 2. Estonia, Matsalu National Park. WGS84
{
      "type": "Polygon",
      "coordinates": [[
            [23.497559,58.805065], [23.463501,58.770906],
            [23.555786,58.772615], [23.616211,58.78799],
            [23.663452,58.781157], [23.718384,58.785143],
            [23.796387,58.746974], [23.660156,58.736713],
            [23.590942,58.726449], [23.526123,58.695637],
            [23.592041,58.696208], [23.739258,58.713899],
            [23.829346,58.718463], [23.855713,58.741844],
            [23.929321,58.731581], [23.972168,58.750394],
            [23.798584,58.810186], [23.742554,58.804496],
            [23.687622,58.794252], [23.620605,58.807341],
            [23.561279,58.794252], [23.497559,58.805065]
      ]]
}
```