

## DATA PROCESSING ARCHITECTURES FOR MONITORING FLOODS USING SENTINEL-1

W. Wagner<sup>1,2\*</sup>, V. Freeman<sup>3</sup>, S. Cao<sup>1</sup>, P. Matgen<sup>4</sup>, M. Chini<sup>4</sup>, P. Salamon<sup>5</sup>, N. McCormick<sup>5</sup>, S. Martinis<sup>6</sup>,  
B. Bauer-Marschallinger<sup>1</sup>, C. Navacchi<sup>1</sup>, M. Schramm<sup>1</sup>, C. Reimer<sup>2</sup>, C. Briese<sup>2</sup>

<sup>1</sup> Department of Geodesy and Geoinformation, TU Wien, 1040 Vienna, Austria - wolfgang.wagner@geo.tuwien.ac.at

<sup>2</sup> EODC Earth Observation Data Centre, 1040 Vienna, Austria

<sup>3</sup> Spire Global Luxembourg, 33 rue Sainte Zithe, 2763 Luxembourg

<sup>4</sup> Luxembourg Institute of Science and Technology, 4362 Esch-sur-Alzette, Luxembourg

<sup>5</sup> European Commission, Joint Research Centre (JRC), Ispra, Italy

<sup>6</sup> Deutsches Zentrum für Luft- und Raumfahrt, 82234 Wessling, Germany

### Commission III, ICWG III/IVa

**KEY WORDS:** SAR, Sentinel-1, Floods, Water Bodies, Data Cubes, Big Data, Model Calibration, Change Detection

#### ABSTRACT:

Synthetic Aperture Radar (SAR) images acquired by Earth observation satellites often constitute the only source of information for monitoring the progression of flood events over larger regions. Particularly attractive are the SAR data acquired by the Copernicus Sentinel-1 satellites because they are free and open, and combine a short revisit time with a good spatial and radiometric resolution. In this contribution, we discuss how a Sentinel-1 data processing system should be designed to optimally benefit from the dense Sentinel-1 time series and advanced algorithms such as change detection or machine learning methods. This was one of the questions addressed by an expert group tasked by the Joint Research Centre of the European Commission to investigate the feasibility of an automated, global, satellite-based flood monitoring product for the Copernicus Emergency Management Service. Drawing from the expert group report, we distinguish three broad categories of data processing architectures, namely single-image, dual-image, and data cube processing architectures. While the latter architecture is the most demanding in terms of large storage and compute capacities, it is also the most promising to derive high-quality Sentinel-1 flood maps comprised not just of the flood mask but also of data fields describing the retrieval uncertainty and masks showing where Sentinel-1 cannot detect floods due to physical reasons. Therefore, we recommend to use data cube processing architectures and showcase the use of the Austrian Data Cube for monitoring a small-scale flood event that occurred in Austria in November 2019.

### 1. INTRODUCTION

Sentinel-1 is a constellation of polar-orbiting radar satellites operated by the European Space Agency (ESA) as part of the European Union's Copernicus programme. At present two Sentinel-1 satellites (1A and 1B) are in orbit; preparations for the launch of Sentinel-1C are on-going. Each of these satellites provides day-and-night all-weather Synthetic Aperture Radar (SAR) imaging capabilities at C-band (5.404 GHz). Unlike previous SAR missions, Sentinel-1 was designed to achieve systematic coverage in a limited number of acquisition modes that meet most user requirements (Potin et al., 2012). Over land, Sentinel-1 has been acquiring data almost exclusively in Interferometric Wide (IW) swath mode (250 km wide swath, VV and VH polarisations) in pre-defined coverage patterns. A key advantage of this systemic approach is that processing lines can be streamlined and made fully automatic. Another important advantage is that coverage is maximised, becoming de facto mainly limited by the maximum duty cycle of the SAR instrument which, in the case of IW mode, is about 25 min per 100 min orbit. As this precludes a uniform worldwide coverage, acquisitions are prioritised according to region. As shown in Figure 1, Sentinel-1A and 1B cover Europe exceptionally well, providing SAR measurements every 2-3 days on average. Over the other continents tectonic zones and agricultural areas are also often well covered, with measurements taken every 5 to 10

days. Least-prioritised regions, which are typically situated in arid and cold environments, are covered every 10 to 15 days.

The systematic coverage undisturbed by clouds and lightning conditions makes Sentinel-1 highly attractive for numerous applications. The high temporal sampling rate is not only beneficial for the mapping of urban areas (Lisini et al., 2018), forests (Dostálová et al., 2018), rice (Bazzi et al., 2019) and other land cover types, but even more so for the monitoring of dynamic land surface variables such as soil moisture (Bauer-Marschallinger et al., 2019), vegetation (Vreugdenhil et al., 2018), snow (Lievens et al., 2019) and dynamic water bodies (Huang et al., 2018). In this contribution we discuss how Sentinel-1 can be used for systematic and fully-automatic monitoring of flood events. Our prime interest is not in a specific algorithm for turning an incoming Sentinel-1 image into a flood map, but in the question of how to set up the data processing system in such a way as to optimally benefit from the unique properties of the Sentinel-1 data and novel algorithmic approaches such as change detection and machine learning models.

This paper is a synthesis of selected sections from an expert group report that investigated the feasibility of introducing an automated, global, satellite-based flood monitoring product based on Sentinel-1, in order to complement and enhance the capabilities of the Copernicus Emergency Management Service (CEMS) for mapping and monitoring floods (Matgen et al.,

\* Corresponding author

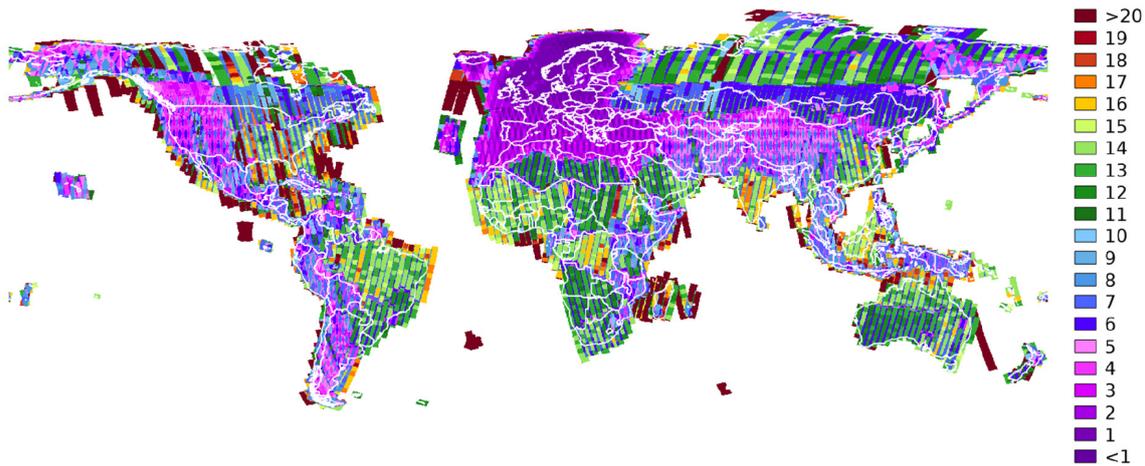


Figure 1. Average revisit time achieved by Sentinel-1A and 1B over land in Interferometric Wide Swath mode, based on 2017 data acquisitions.

2019). Based upon a short overview of scientific algorithms for floodwater mapping in Section 2, we introduce three different system architectures in Section 3 and discuss their requirements in Section 4. The importance of choosing the right system architecture is illustrated in Section 5 by analysing a recent flood event that occurred in Austria in November 2019.

## 2. SCIENTIFIC METHODS

### 2.1 Flood Mapping Algorithms

Water bodies can be detected in SAR images because the mirror-like reflection of microwave pulses by the water surfaces leads to backscatter intensities that are much lower than for most other land cover types. This physical mechanism renders the mapping of open, calm water in principle rather straightforward. Therefore, even thresholding algorithms applied to single images of flood events often produce quite satisfying results (Manjusree et al., 2012). However, applying such algorithms over larger regions in a fully automatic fashion is prone to errors as there are many confounding effects that lead to signal ambiguities (e.g. wind-induced waves, low signal contrast to sand, asphalt or grassland, etc.) or even prevent the detection of water at all (e.g. dense vegetation, urban regions, radar shadows, etc.).

The impact of confounding effects can be minimised by using change detection approaches that are less sensitive to the generation of false positives. Changes can thus be directly attributed to sudden changes occurring on the ground. Using two or more SAR images rather than a single-image transforms the flood mapping issue to a classification problem between change and no change. Following the computation of the difference image, different histogram thresholding approaches (often in combination with image tiling and region growing methods) can be applied to generate a binary classification. These methods assume that one type of change (i.e. a decrease of backscatter due to the specular reflection on water bodies) dominates all others. For example, O'Grady et al. (2011) showed that misclassifications of non-water pixels due to low backscatter over dry regions can be effectively reduced with the use of dual-image processing approaches. The selection of adequate reference images is a prerequisite for an efficient change detection. Different auto-

matic selection algorithms have been presented in the literature to select a suitable reference image (Hostache et al., 2012).

The availability of many backscatter observations distributed over time makes it possible to go one step further and to apply time series analyses in an attempt to understand and model the seasonality of the backscatter coefficient (Schlaffer et al., 2015). The rationale is that pixels can be classified as 'flooded' when they show a specific deviation from a seasonal trend inferred from a statistical model. Time series analyses applied on SAR data archives are also well suited to improve the characterization of permanent water bodies (Santoro et al., 2015).

### 2.2 Error Characterisation

The value of Sentinel-1 flood data products can be much increased when layers on the retrieval uncertainty are added, describing the retrieval uncertainty and areas where the sensors and algorithms fail to detect flood affected areas for physical reasons (forests, urban areas, radar shadow regions, etc.). This is crucial for the usability of the flood data products, because users can generally deal with known unknowns, but not with unknown unknowns. For example, a quantification of the exclusion areas and classification uncertainty is required when assimilating such datasets into numerical models for improved flood forecasting and monitoring. The identification of exclusion areas can be done on the basis of the Sentinel-1 data themselves or through the use of topographic and land cover data. For example, Huang et al. (2017) used Digital Elevation Models (DEMs) to compute two popular terrain indices that can assist the masking of Sentinel-1 data in mountainous catchments. The uncertainties in SAR-based flood mapping is often the by-product of the classifier that was used for generating the flood maps. Other methodologies produce uncertain flood maps based on fuzzy set theory (Pulvirenti et al., 2011) and Bayesian statistics (D'Addabbo et al., 2016).

## 3. SYSTEM ARCHITECTURES

Considering the algorithmic approaches discussed above, we propose to distinguish three broad categories of data processing architectures, namely *single-image*, *dual-image*, and *data cube* processing architectures. They are illustrated by Figures 2 to 4 and discussed in the following sub-sections.

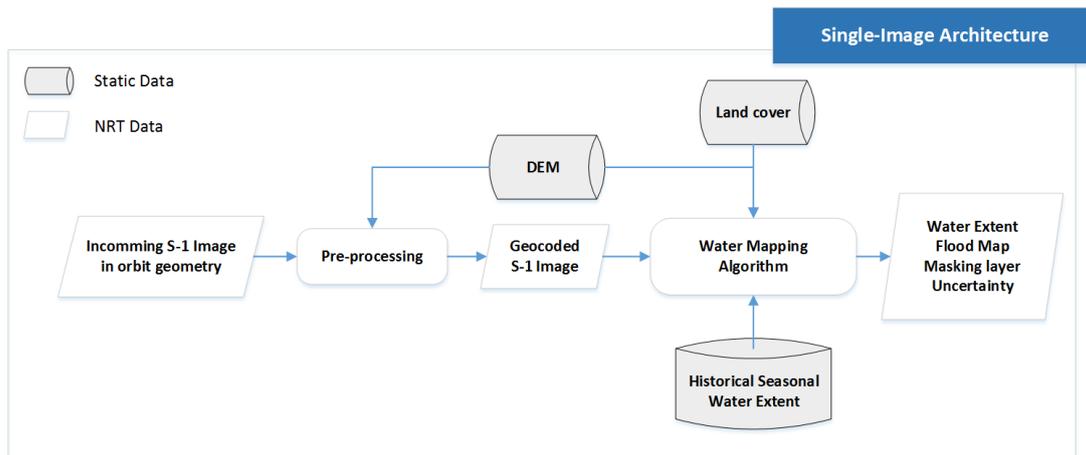


Figure 2. Single-image processing architecture for global flood monitoring using Sentinel-1 SAR image data.

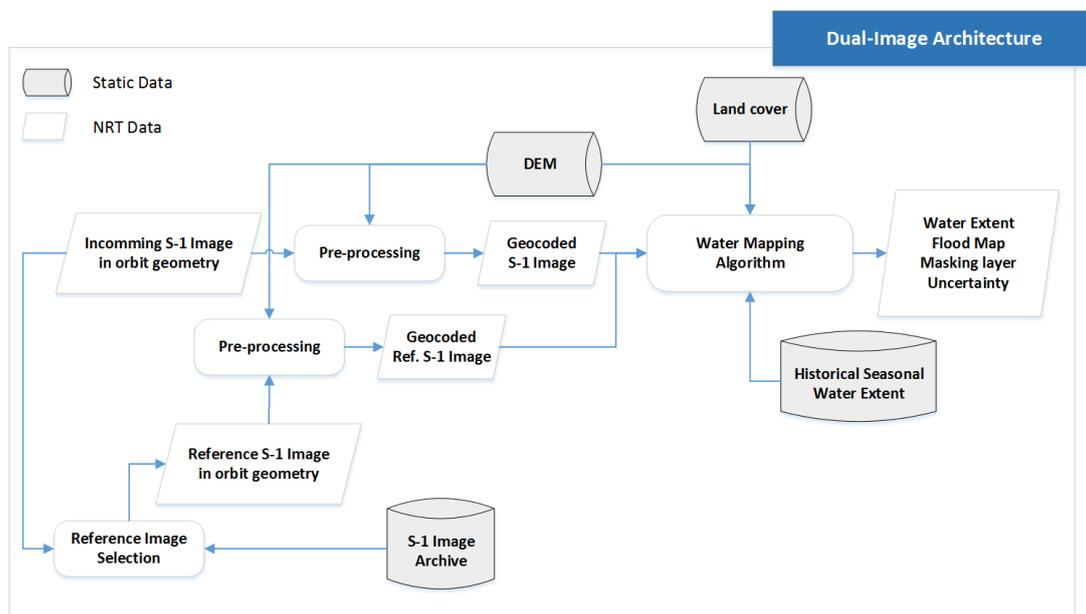


Figure 3. Dual-image processing architecture for global flood monitoring using Sentinel-1 SAR image data, based on change detection approaches using individual historic Sentinel-1 images as a reference.

### 3.1 Single-Image Processing Architecture

In the first, most basic single-image processing architecture (Figure 2), the water mapping algorithm is applied to a single Sentinel-1 image, using some ancillary data, such as a digital elevation model, and maps of land cover and historical water extent. In this case, the data flow is simple, whereby incoming Sentinel-1 SAR Level-1 data are converted, step by step, to geocoded imagery and the flood mapping product. The algorithms are designed to work with single SAR images, typically relying on processing techniques that, for example, combine thresholding with region-growing and noise reduction. From an engineering point of view, the single-image processing architecture is easily implemented. Drawbacks are: (a) It is difficult to deal with spatial heterogeneity of the land surface, due to the limited information content of single SAR images; (b) Training and calibration of algorithms is not naturally built into the processing architecture, and so is usually done only for a limited number of flood events.

### 3.2 Dual-Image Processing Architecture

Change detection approaches, which are based on a comparison of flooded and non-flooded Sentinel-1 SAR images, enable the detection of flooded areas more reliably than single-image approaches. In the simplest case, change detection algorithms can be implemented using a dual-image processing architecture (Figure 3) that allows the comparison of the incoming SAR image to an historic SAR image extracted from the Sentinel-1 data archive. This historic image may e.g. be the most recent pre-flood image or any image characteristic of non-flooding conditions. This architecture is somewhat more difficult and costly to implement than the single-image processing architecture, for example due to the need to maintain an Sentinel-1 data archive, and provide fast access to it. However, this architecture has the important advantage that change detection algorithms are better able to handle the spatial heterogeneity of the land surface than single-image methods. This simply reflects the fact that two custom-selected SAR images hold more information than a single image. Additionally, change detection algorithms

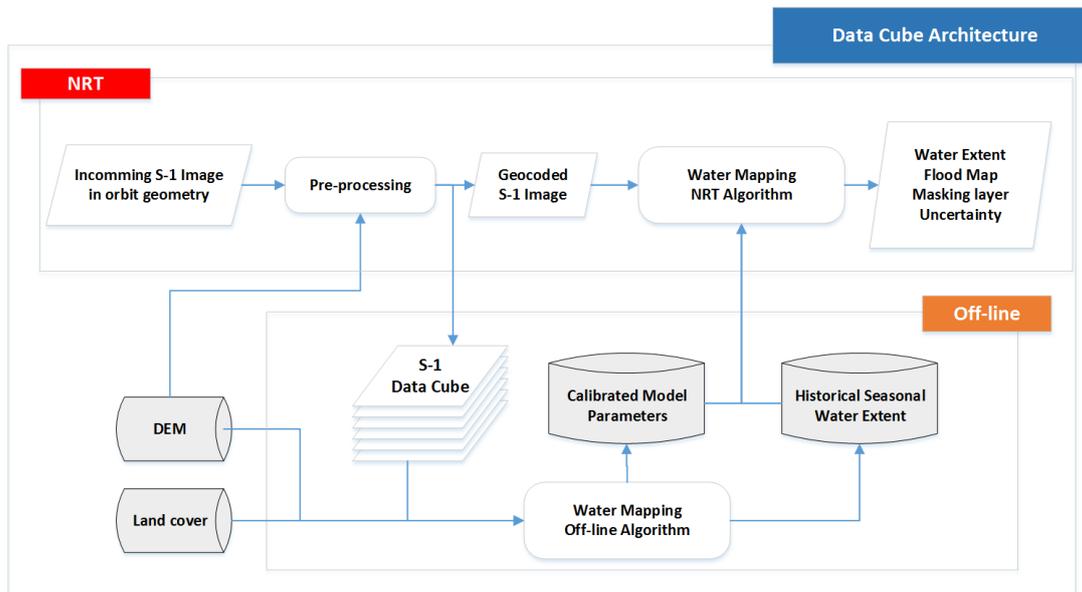


Figure 4. Data cube processing architecture for global flood monitoring using Sentinel-1 SAR image data, consisting of an NRT data flow and off-line components.

are more easily applied to different geographic regions. Nonetheless, a dual image processing architecture does not fully solve the training and calibration problem, since region-specific thresholds and / or model parametrisations are still required.

### 3.3 Data Cube Processing Architecture

The most sophisticated data processing architecture is based on the “data cube” concept, whereby incoming SAR images are geocoded, gridded and added as *analysis ready data* (ARD) to an existing spatio-temporal SAR data cube (Figure 4). By using a data cube processing architecture, where the temporal and spatial dimensions are treated alike, each incoming Sentinel-1 image can be compared with the entire backscatter history, in a straightforward manner. The entire backscatter time series for each pixel can then be analysed, in order to derive pixel-specific backscatter statistics, which can then be used for example to derive pixel-specific thresholds and model parametrisations. Using this architecture, the full Sentinel-1 SAR information content is used (by analysing the entire backscatter history) and standardised throughout the complete time series (by also gridding the temporal dimension into a constant frequency). Therefore, model training and calibration may be carried out systematically for each pixel. Advantages are: (a) Algorithms are better able to handle land surface heterogeneity; (b) Uncertainties can be better specified; (c) Regions where open water cannot be detected for physical reasons (e.g. dense vegetation, urban areas, deserts), can be determined *a priori*. Additionally, historic water extent maps are produced, essentially as a by-product of the model calibration, which may serve as a reference for distinguishing between floods and the normal seasonal water extent.

## 4. SYSTEM REQUIREMENTS

The system resources required for a worldwide Sentinel-1 flood monitoring product depend primarily on the volume of Sentinel-1 SAR data, the computation time per image, and the user requirements regarding product timeliness and availability. While the exact specifications of the flood data product

are important, probably the most important drivers of costs are the data pre-processing efforts, the spatial sampling of the flood product, and the need to provide historic data that are consistent (e.g. regarding format, software version) with the near real-time (NRT) data stream. The selected data formats, data latencies (time delays), and system availability also play a role. In the following, we discuss some of the requirements arising from a fully automatic, worldwide Sentinel-1 data service that aims to deliver flood data to users within 8-12 hours after sensing (Figure 4).

### 4.1 Sentinel-1 Data Access

A fundamental requirement for the NRT generation of products, is a fast, uninterrupted access to input data. In the present context, the lowest data latencies (i.e. 1-2 hours) would be achieved by receiving the Sentinel-1 data at local ground receiving stations. However, costs for running dedicated reception services at one or more ground receiving stations are probably large, and coverage would be limited to a few regions worldwide. Acceptable data latencies (i.e. 8-12 hours) should also be attainable by downloading the data from the Copernicus Service Data Hub and Collaborative Nodes. Already now Sentinel-1 data are provided via these hubs with different data latencies. Besides the standard non-time-critical Sentinel-1 data products there is also a “Fast24h” data stream that provides ground range detected Sentinel-1 Level 1 data on a world-wide basis within about 6 hours on average. Over Europe and some other selected regions there is even a “NRT-3h” data stream that disseminates Sentinel-1 Level 1 to users in under 3 hours. In this scenario for data access, costs arise due to high bandwidth internet access, to allow downloading several Terabytes per day, and dedicated efforts to ensure a fast, uninterrupted data stream for downloading. These costs might be reduced by accessing Sentinel-1 data at one of the Copernicus Data and Information Access Services<sup>1</sup> or at equivalent Earth observation cloud platforms such as EODC Earth Observation Data Centre<sup>2</sup>.

<sup>1</sup> <https://www.copernicus.eu/en/access-data>

<sup>2</sup> <https://www.eodc.eu>

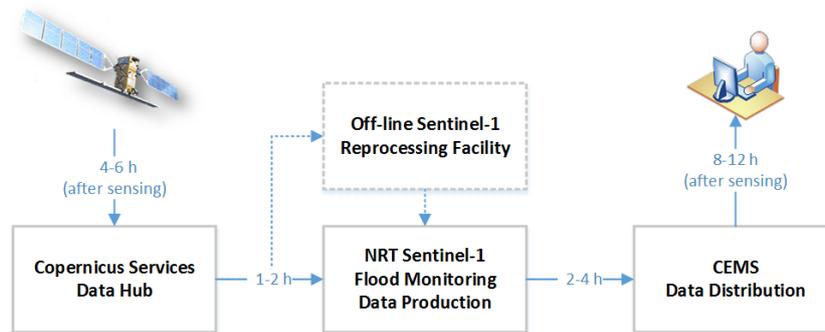


Figure 5. Data flow and envisaged data latencies (time delays) from the Copernicus Service Data Hub to the CEMS data distribution facilities. Note that the indicated timeliness of 8-12 hours represents the total time from Sentinel-1 data acquisition to flood monitoring product distribution to users.

## 4.2 Storage Capacity

Depending on the product specifications and system set-up, the storage required for data processing and archiving can be substantial. The Sentinel-1 Ground Segment operations generate about 250 terabytes (TB) of Interferometric Wide Swath (IW) mode - Level 1 - Ground Range Detected (GRD) data per year (i.e. about 0.7 TB per day). For NRT processing of incoming Sentinel-1 data, at its most basic level it is sufficient to have storage capacity for processing only one day of Sentinel-1 data (e.g. a few TB to store Levels 1, 2, and intermediate data products). In practice however, such minimalistic requirements are not realistic, as users wish to have access to historical data, which implies keeping at least one copy of the derived flood maps and ancillary internal data - leading to a requirement of a few hundreds of TB per year. In the event that users want the NRT- and historical data to be consistent (in terms of data format or software versions used along the entire processing chain), storage needs to be large enough to hold also the Level 1 data, in order to allow for regular re-analysis efforts (which ensures that algorithmic updates can also be applied to historic data). In such a scenario, the required storage space to store all data (Levels 1, 2 and intermediate data) is about two to four times that of the Level 1 data volume. Hence, the required storage capacity would be about 0.5 to 1 petabytes (PB) per year.

## 4.3 NRT Data Processing

The NRT data processing system must have the capability to handle all daily acquisitions of Sentinel-1 data. A typical flood mapping processing chain includes pre-processing of SAR scenes (calibration, noise removal, terrain correction, geo-referencing), and water mapping. Computational performance solely for the Sentinel-1 pre-processing (i.e. to generate geometrically and radiometrically corrected images on a 10x10 metres rastergrid), based on evaluations done with ESA's Sentinel-1 Toolbox, is about 5 megabits (Mbit) per second - equivalent to 1.6 seconds per megabyte (MB) - using a high performance computing node (i.e. two processors of Intel Xeon 2.6 GHz each having 8 cores). For example, for pre-processing 0.7 TB of daily Sentinel-1 data, based on a performance factor of 5 Mbit per second, 311 node-hours are required (e.g. 20 nodes running for 16 hours daily). The computation effort required for flood mapping (which varies depending on the selected algorithm) must be added to this estimate, but is typically (much) less effort that is needed for the pre-processing. Therefore, a conservative estimate of the computing resources needed to run

the NRT service is 30-40 computing nodes. (A certain overhead is needed to handle fluctuations in the incoming Sentinel-1 data stream).

## 4.4 Offline Data Processing

In order to either allow a regular re-processing of the historic flood maps, or to support a data cube processing architecture (Section 3.3), an off-line high performance computing (HPC) environment is needed in parallel to the NRT system. Data processing in such an environment is not as time-critical as in NRT, in that the requirements for hardware availability are not stringent. However, sufficient HPC resources are required to perform analysis over hundreds of TB of historical Sentinel-1 data. Hence, the HPC facilities must be large enough to be capable of providing in the order of millions of core hours per month in order to complete re-processing efforts in reasonable time periods (i.e. a few weeks to months per re-analysis cycle). For example, when using compute nodes with 2 processors, each with 8 cores, then 87 compute nodes are needed to provide 1 million core hours per month. In practice, re-processing jobs should be run on 200-500 compute nodes, so the HPC system must consist of several hundred to a few thousand compute nodes to provide sufficient resources when needed. Such HPC resources are e.g. available at the EODC which connects its worldwide Sentinel-1 data archive with the Vienna Scientific Cluster<sup>3</sup> via a high-bandwidth network fabric.

## 4.5 System Bandwidth

An important characteristic of any Sentinel-1 data processing environment - whether online or off-line - is the bandwidth between the processing components and the storage. High bandwidth is needed, as data transfer rate must be high in order to parallelise the computations of the Sentinel-1 data. In view of the Sentinel-1 data volume acquired per day, the network bandwidth capacity between the computation units and storage must be several tens of gigabits (Gbits) per second, in order to generate flood maps seamlessly, with no input / output restriction. Similarly, a high bandwidth is crucial when re-processing historical Sentinel-1 data.

## 4.6 Metadatabase

Metadata contain information for understanding, interpreting, and managing the data, which is needed for correct and effective data processing. For any system capable of the automated

<sup>3</sup> <http://vsc.ac.at/>

processing of global Sentinel-1 data, the availability of an active metadatabase, suitable for steering processing efforts, is a prerequisite. Throughout the data processing chain, different kinds of information must be gathered to set up configurations for accessing and processing various data sources. A fast and reliable querying of both input and output data products, should be available, based on region of interest, acquisition time, and satellite data specifications (acquisition mode, polarization, orbit, etc.). The query results are used to make decisions, select the right model parameters, and read relevant information from auxiliary data sources (e.g. land cover, historical flood maps, advisory flags, and masking layers). The metadatabase should be automatically updated as soon as a data product is generated, in order to keep track of processed and non-processed data-files in near real-time.

#### 4.7 System Redundancy and Product Availability

Timely and reliable delivery of data products is an essential aspect of the NRT system. The system should be reliable enough to provide a non-stop (24 hours per day, seven days a week) service operation, with a high product availability. Therefore, a monitoring system must be in place, to automatically detect failures and to recover data processing instances, using redundant system components, including the following:

- *Access node redundancy*: In the event of failing to access input data or auxiliary data, Sentinel-1 data should be accessed from alternative data hubs or cloud platforms.
- *NRT hardware redundancy*: In the event of any failure of the storage or computing nodes needed for the NRT processor, a redundant NRT processor should take over
- *Software redundancy*: The processing chain should be implemented in redundant environments ready for running identical code. Furthermore, a protocol is required to steer the switching between the individual processing chains.

## 5. DISCUSSION

From the above discussion of system requirements it is clear that the data cube processing architecture is the most demanding approach in terms of storage, compute, bandwidth, and system maintenance capacities. However, as already noted in Section 3.3, it brings important benefits that makes this architecture attractive for operational users and scientists alike.

#### 5.1 Scientific Benefits of Data Cube Architecture

From a scientific perspective, the key advantage of the data cube architecture is that it provides fast access to Sentinel-1 time series. Therefore, it becomes possible to implement advanced algorithms such as change detection (Schlaffer et al., 2015) or machine learning (Kreiser et al., 2018) methods that require model training. Since the data cube provides access to the backscatter time series of each land surface pixel, it is even possible to parameterise the models for each pixel separately. This simplifies the task of coming up with algorithms that are transferable across geographic domains. Moreover, this implies that retrieval uncertainties and exclusion areas may also be derived on a per pixel basis.

Note that these scientific tasks are further simplified when the Sentinel-1 data are pre-processed to higher-value data formats

that allow compare the backscatter data across space and time. At the very least, Sentinel-1 stored in the cube have been geocoded and local incidence angles are provided. But to better deal with changes in the imaging geometry related to the orbit number and pass direction, it is preferable to work with radiometrically corrected, terrain-flattened backscatter (Small, 2011). This was for example recognised by the Committee on Earth Observation Satellites (CEOS) in their definition of the ARD format for radar backscatter<sup>4</sup>.

#### 5.2 User Benefits of Data Cube Architecture

Users of the Sentinel-1 flood data products will directly benefit from the improved accuracy, robustness and characterisation of data quality that one can expect from per-pixel trained water body mapping algorithms. But equally important for many users is that they do not just received the latest flood images, but have also access to the complete historic data record. This allows them to validate and calibrate their flood forecasting models, and assess the severity of flood events. Last but not least, historic water extent maps can be produced, essentially as a by-product of the calibration of the water mapping algorithm. These historic maps can then serve as a reference for distinguishing between inundation areas caused by the flooding and the normal seasonal water extent. This reference or baseline information is best obtained directly from historical Sentinel-1 time series, in order to ensure high consistency with the NRT data product. The use of other global datasets, such as those derived from optical satellite imagery - e.g. the Global Surface Water Explorer<sup>5</sup>, developed by Pekel et al. (2016) - would be problematic for this purpose, given that surface water areas seen by optical sensors are not identical to those seen by Sentinel-1 SAR. Therefore, subtracting an optical reference water map from an Sentinel-1 water extent map would produce systematic errors in the derived flood inundation area. This problem also applies to datasets derived from other SAR sensors, operating at different wavelengths, polarisations or spatial resolutions from Sentinel-1.

#### 5.3 Use Case

Let us illustrate the benefits of the data cube architecture by discussing the experiences made with using either single Sentinel-1 images or the complete Sentinel-1 data cube for documenting a recent smaller-scale flood event in Austria. This event happened in November 2019 and was caused by severe weather which mainly affected northern Italy and south-eastern France but also brought record snow- and rainfall to the southern parts of Austria. While Austria was spared a major disaster, many creeks, rivers and lakes overflowed. Together with the accompanying mudflows and landslides, the flooding caused significant damage to public infrastructure (roads, railway, electricity, etc.) and private property (housing, agricultural land and forests). Just for the federal state of Carinthia alone, the economic damages are estimated to be in the three-digit million € range<sup>6</sup>.

For this event, the CEMS Rapid Mapping service was activated (EMSR414<sup>7</sup>). The Rapid Mapping service operates in a

<sup>4</sup> <http://ceos.org/ard/index.html#slide3>

<sup>5</sup> <https://global-surface-water.appspot.com/>

<sup>6</sup> <https://www.diepresse.com/5729310/nach-unwetter-versicherungen-rechnen-mit-schaden-in-millionenhohe>

<sup>7</sup> <https://emergency.copernicus.eu/mapping/list-of-components/EMSR414>

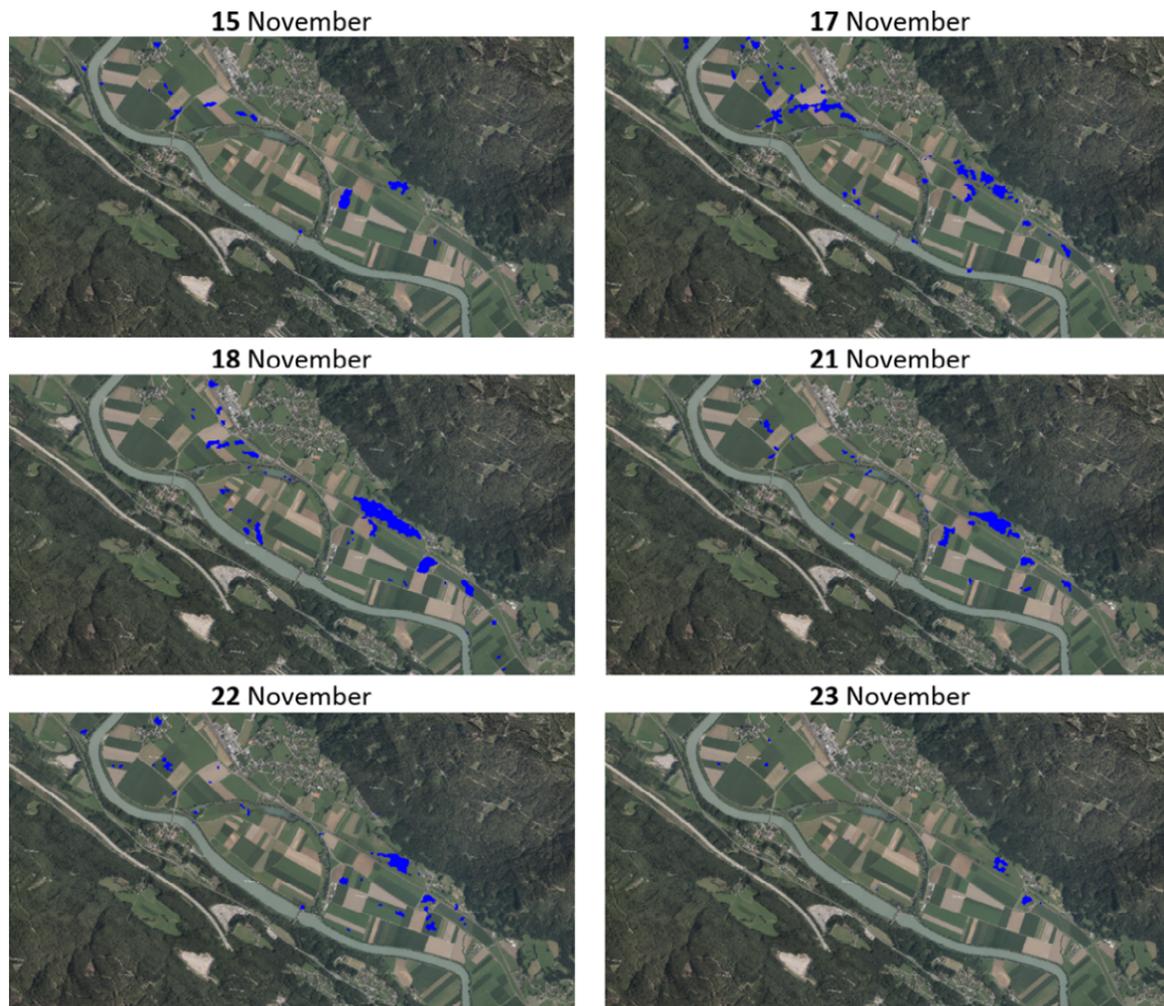


Figure 6. Progression of the November 2019 flooding along the river Drau near Weißenstein in Carinthia for the period 15 to 23 November 2019. The flood extent maps were derived from a sequence of six Sentinel-1 images by comparing each SAR image (in VV polarisation) to a Sentinel-1 composite image depicting normal (i.e. non-flooding) conditions during this part of the year. The images show an overlay of the Sentinel-1 derived water extent maps on top of a Sentinel-2 image. The SAR differences images were post-processed (filtered, classified) and cleaned for outliers.

way that, once activated, the CEMS team collects available optical and radar imagery acquired before, during and after the event, and carries out an expert interpretation of these images to delineate inundated areas. For the November 2019 flooding the CEMS team used mostly pre- and post-event Sentinel-1 and Sentinel-2 imagery to document the size (magnitude and extent) of the event, but failed to detect inundation area in the Sentinel-1 images (“No impact detected”). Only in one post-event image acquired by the COSMO-SkyMed satellite over a region in Carinthia some inundation areas could be detected.

To better understand why the EMRS14 activation did not result in Sentinel-1 flood maps, we analysed analysis-ready Sentinel-1 data made available via the Austrian Data Cube (ACube). The ACube has been developed within a research project funded by the Austrian Space Application Programme, and is envisioned to become an open government data service. It hosts Sentinel-1 and Sentinel-2 data pre-processed with state-of-the-art algorithms according to standards (cartographic projection, terrain model, definition of output variables, etc.) as agreed upon by the Austrian public user community. Our own analysis of the 2019 flood event confirmed that it is indeed diffi-

cult to depict inundation areas for this localised and short-lived flood event for which the extent of affected areas is often not much larger than the size of individual Sentinel-1 pixels (10-20 m). Moreover, the flood took place in a mountainous region, which meant that topographic effects and the presence of wet snow further complicated the analysis. Nonetheless, by benefiting from the complete Sentinel-1 data cube and calculating anomalies, inundated areas started to emerge. After some post-processing and clearing for erroneously classified areas, we obtained for the Carinthia region covered by the EMSR414 activation a realistic, and temporally dense (every 1-3 days) sequence of inundation maps (Figure 6).

## 6. SUMMARY

Floods are the most frequent and costliest natural disasters worldwide. State-of-the-art scientific methods for automatically detecting and identifying flood events, based on a global, continuous supply of all-weather, day-and-night SAR images, such as those provided by Europe’s Copernicus Sentinel-1 satellites, are now mature and in principle ready for operational implementation. Nonetheless, transferring the science al-

gorithms into an operational setting is challenging and requires careful attention to the way of how the processing system is designed. Whilst many published algorithms could be implemented in processing architectures designed to work with single or two images (one flood and one non-flood image), we recommend to use data cube processing architectures. This allows to train advanced change detection or machine-learning methods on a per pixel basis, which can be expected to improve the accuracy, transferability and data quality characterisation of the flood data products. Additionally, users benefit from a historic data record for training their flood forecasting models and for assessing how the current water extent compares to regular seasonal fluctuations and past flood events.

## ACKNOWLEDGEMENTS

This study was partly funded by the Copernicus Emergency Management Service. We also acknowledge funding by the Austrian Research Promotion Agency (FFG) for the Austrian Data Cube project, and by ESA for the Gamma2Cloud (Feasibility of Using Sentinel-1 Terrain-flattened Gamma Naught Backscatter Across EO platforms) project.

## REFERENCES

- Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., Wagner, W., 2019. Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles. *IEEE Transactions on Geoscience and Remote Sensing*, 57(1), 520-539.
- Bazzi, H., Baghdadi, N., El Hajj, M., Zribi, M., Minh, D. H. T., Ndikumana, E., Courault, D., Belhoucette, H., 2019. Mapping Paddy Rice Using Sentinel-1 SAR Time Series in Camargue, France. *Remote Sensing*, 11(7).
- D'Addabbo, A., Refice, A., Pasquariello, G., Lovergine, F. P., Capolongo, D., Manfreda, S., 2016. A Bayesian network for flood detection combining SAR imagery and ancillary data. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6), 3612-3625.
- Dostálová, A., Wagner, W., Milenković, M., Hollaus, M., 2018. Annual seasonality in Sentinel-1 signal for forest mapping and forest type classification. *International Journal of Remote Sensing*, 39(21), 7738-7760.
- Hostache, R., Matgen, P., Wagner, W., 2012. Change detection approaches for flood extent mapping: How to select the most adequate reference image from online archives? *International journal of applied earth observation and geoinformation*, 19, 205-213.
- Huang, C., Nguyen, B. D., Zhang, S., Cao, S., Wagner, W., 2017. A comparison of terrain indices toward their ability in assisting surface water mapping from Sentinel-1 data. *ISPRS International Journal of Geo-Information*, 6(5), 140.
- Huang, W., DeVries, B., Huang, C., Lang, M. W., Jones, J. W., Creed, I. F., Carroll, M. L., 2018. Automated Extraction of Surface Water Extent from Sentinel-1 Data. *Remote Sensing*, 10(5).
- Kreiser, Z., Killough, B., Rizvi, S. R., 2018. Water across synthetic aperture radar data (wasard): Sar water body classification for the open data cube. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 437-440.
- Lievens, H., Demuzere, M., Marshall, H.-P., Reichle, R. H., Brucker, L., Brangers, I., de Rosnay, P., Dumont, M., Giroto, M., Immerzeel, W. W. et al., 2019. Snow depth variability in the Northern Hemisphere mountains observed from space. *Nature Communications*, 10(1), 1-12.
- Lisini, G., Salentinig, A., Du, P., Gamba, P., 2018. SAR-Based Urban Extents Extraction: From ENVISAT to Sentinel-1. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(8), 2683-2691.
- Manjusree, P., Prasanna Kumar, L., Bhatt, C. M., Rao, G. S., Bhanumurthy, V., 2012. Optimization of threshold ranges for rapid flood inundation mapping by evaluating backscatter profiles of high incidence angle SAR images. *International Journal of Disaster Risk Science*, 3(2), 113-122.
- Matgen, P., Martinis, S., Wagner, W., Freeman, V., Zeil, P., McCormick, N., 2019. Feasibility assessment of an automated, global, satellite-based flood-monitoring product for the Copernicus Emergency Management Service. *European Commission, Ispra, JRC119812, EUR 30073 EN*, 1-47.
- O'Grady, D., Leblanc, M., Gillieson, D., 2011. Use of ENVISAT ASAR Global Monitoring Mode to complement optical data in the mapping of rapid broad-scale flooding in Pakistan. *Hydrology and Earth System Sciences*, 15, 3475-3494.
- Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A. S., 2016. High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418-422.
- Potin, P., Bargellini, P., Laur, H., Rosich, B., Schmuck, S., 2012. Sentinel-1 mission operations concept. *2012 IEEE International Geoscience and Remote Sensing Symposium*, 1745-1748.
- Pulvirenti, L., Pierdicca, N., Chini, M., Guerriero, L., 2011. An algorithm for operational flood mapping from Synthetic Aperture Radar (SAR) data using fuzzy logic. *Natural Hazards and Earth System Sciences*, 11(2), 529-540.
- Santoro, M., Wegmüller, U., Lamarche, C., Bontemps, S., Defourny, P., Arino, O., 2015. Strengths and weaknesses of multi-year Envisat ASAR backscatter measurements to map permanent open water bodies at global scale. *Remote Sensing of Environment*, 171, 185-201.
- Schlaffer, S., Matgen, P., Hollaus, M., Wagner, W., 2015. Flood detection from multi-temporal SAR data using harmonic analysis and change detection. *International Journal of Applied Earth Observation and Geoinformation*, 38, 15-24.
- Small, D., 2011. Flattening gamma: Radiometric terrain correction for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 49(8), 3081-3093.
- Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., Strauss, P., 2018. Sensitivity of Sentinel-1 Backscatter to Vegetation Dynamics: An Austrian Case Study. *Remote Sensing*, 10(9).