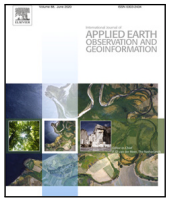
International Journal of Applied Earth Observation and Geoinformation 119 (2023) 103329

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Time-series analysis of Sentinel-1/2 data for flood detection using a discrete global grid system and seasonal decomposition

Florian Fichtner *, Nico Mandery, Marc Wieland, Sandro Groth, Sandro Martinis, Torsten Riedlinger

German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Münchner Strasse 20, Wessling, 82234, Germany

ARTICLE INFO

Dataset link: https://github.com/dlr-eoc/ukis-h3cellstore

Keywords:
Flood detection
H3
Time-series analysis
Seasonal decomposition
Anomaly detection
Sentinel-1
Sentinel-2
Copernicus EMS

ABSTRACT

Automated flood detection using earth observation data is a crucial task for efficient flood disaster management. Current solutions to identify flooded areas usually rely on calculating the difference between new observations and static, pre-calculated water extents derived by either single acquisitions or timely aggregated products. Such pre-calculated datasets, however, lack representation of real-world seasonality and short-term changes in trend.

In this paper we present a complete workflow to automatically detect hydrological extreme events and their spatial extent, which automatically adapts to local seasonality and trend. For that we rely on a novel combination of well-established algorithms and tools to detect anomalies in time-series of water extent across large study areas. The data is binned into a discrete global grid system H3, which greatly simplifies aggregation across spatial and temporal resolutions. For each grid cell of an H3 resolution we perform a time-series decomposition using Seasonal and Trend decomposition using Loess (STL) of the cell's proportion which is covered with surface water. All cells receive an anomaly score, calculated with extended isolation forest (EIF) on the residuals for each step in time. A burst of anomalies represents a hydrological extreme event like a flood or low water level.

The presented methodology is applied on Sentinel-1/2 data for two study areas, one near Sukkur, Pakistan and the other one in Mozambique. The detected anomalies correlate with reported floods and seasonal variations of the study areas. The performance of the process and the possibility to use different H3 resolutions make the proposed methodology suitable for large scale monitoring.

1. Introduction

Monitoring flood and hydrological drought conditions at large scale using satellite data has become a key part of efficient disaster management. In the past years, a number of water and flood monitoring systems have been set up, e.g. Pekel et al. (2016) or Salamon et al. (2021) that provide satellite-derived surface water extents over time. The Global Surface Water product (Pekel et al., 2016) globally maps the location and temporal distribution of water surfaces derived from Landsat data. The newer Global Flood Awareness System (GloFAS) (Salamon et al., 2021) relies on Copernicus Sentinel-1 Synthetic Aperture Radar (SAR) satellite data and allows monitoring of floods in near real-time (NRT). Periodically updated water extents (e.g. included in Brown et al., 2022) become increasingly available. Such information alone can be valuable for disaster management. An automated spatial and temporal identification of potentially harmful hydrological conditions would provide more context, but poses many challenges. The main question that arises in this context is whether the surface water extent outlined in a satellite scene is anomalously large or small and hence can be considered as a significant flood or a hydrological drought. This is particularly relevant for triggering event alerts during continuous monitoring.

To date, flood mapping is mostly done by comparing the observed water extent to a product representing the *normal* hydrological situation. This could simply be a pre-event satellite scene taken during the same season or just before the anomaly event, like in Hostache et al. (2012). Such approach is typically used in rapid mapping activations like for example (Copernicus Emergency Management Service, 2022). Applying change-detection approaches based on comparing two observations, flood mapping results may be inaccurate due to insufficient choice of the reference image, that could either include another hydrological anomaly or was taken during a different hydrological season.

* Corresponding author.
E-mail address: florian.fichtner@dlr.de (F. Fichtner).

https://doi.org/10.1016/j.jag.2023.103329
Received 17 January 2023; Received in revised form 14 April 2023; Accepted 23 April 2023
Available online 2 May 2023

Schlaffer et al. (2015) reduce such noise by using multi-temporal data to assemble the pre-flood reference.

To overcome the limitations of such approaches, Schneibel et al. (2022) suggests a scene-based anomaly detection. In this approach the number of water pixels are counted over time and when the count is outside the interquartile range the scene is classified as anomaly. While being simple and efficient, this approach is limited to the detection of anomalies at scene-level and does not allow a more refined identification of the exact locations where an anomaly has happened.

In more complex approaches, many observations over time are aggregated to create a reference water product (e.g. Pekel et al., 2016). Flood or drought events are classified as deviations from the pre-calculated reference water masks. Like shown in Martinis et al. (2022), the use of a reference water product that represents permanent water bodies such as the Shuttle Radar Topography Mission (SRTM) Water Body Data (SWBD) (NASA JPL, 2013) could lead to significant overestimation of inundated areas since they do not take seasonal water dynamics into account. Martinis et al. (2022) creates reference water masks based on the frequency of observations in the previous two years. Seasonal masks are created monthly by using the same principle, but with frequencies of the same month. Thresholds are applied to create a binary water mask. For existing reference water products, the choice of globally applicable water frequency thresholds and data fusion criteria are a complex and computationally expensive task and has significant impact on the water extent prediction. Thresholds used are commonly fixed independent on the spatial scale of analysis and neglect trends or limit seasonality to monthly or quarterly products.

In hydrology, it is common to do time-series analysis on river gauge observations, e.g. Bormann et al. (2011). Those approaches are valid for the location of the gauging station, but lack information of the spatial extent of a flood like it is possible with hydrological models and by integrating Earth Observation data. There are attempts to combine this information, using river gauge measurements as ancillary data (Huang et al., 2014), however access to such data is not always possible.

To fill observation gaps and to represent seasonal variations (Bai et al., 2022) suggest the use of time-series.

Like shown in Li and Stefanakis (2020) Discrete Global Grid System (DGGS) offer a wide range of advantages over traditional GIS (Geographical Information System). One of them is the ability to perform multiscale analysis. This is achieved by defining a new spatial reference system which partitions the surface of the earth into defined grid cells. Using a DGGS allows to work with heterogeneous data in a unified way (Purss et al., 2019) and could overcome some computational issues by spatially aggregating pixel information. That means they offer potential to observe successive phenomena at the same geographic location defined by a cell or a group of cells in various scale. A DGGS offers performance, scalability and operational flexibility.

To provide fast and refined information on anomaly events to support emergency or disaster management authorities, satellite mission planners, and science communities we need to answer the following questions:

1. How can a hydrological extreme event and its spatial extent be automatically detected?
2. How can such a model automatically adapt to local seasonality?
3. How can such a model automatically adjust to local trends (e.g. river changes course or lake dries out)?

In this paper, we aim to combine extracted surface water from Sentinel-1 and Sentinel-2 data in a DGGS with time-series analyses (TSA) and decomposition to answer these questions. Our goal is to showcase that using a DGGS based time-series analysis for flood detection is more dynamic than the commonly used solutions described above. Using the observations in each cell of the DGGS within the study area for time-series decomposition we aim to extract seasonal dynamics and anomalies of surface water extents independent of background knowledge about the study areas.

We apply the proposed methodology to different study areas, having high seasonal dynamics and severe flood events. Compared to previous studies in this direction, we want to be able to not only say that an anomaly has happened, but also where exactly it occurred. While doing so we rely on seasonal decomposition of the observations into season, trend and residual. The whole process is based on established methodologies that are combined in a novel way. Finally, we evaluate our methodology by correlating results with reported anomaly events.

2. Study areas and data

In this study, we use image data from the Copernicus Sentinel-1 and Sentinel-2 satellite missions, which both carry sensors with complementing characteristics that are suitable for large-scale systematic surface water monitoring. The Sentinel-1 C-Band Synthetic Aperture Radar (SAR) sensor has cloud-penetrating capabilities and performs image acquisition day and night. While SAR images are more prone to misclassification in water segmentation due to landcover classes having similar backscattering properties (e.g. sand, asphalt), the multi-spectral sensor on-board Sentinel-2 is less susceptible to false classification. Sentinel-2 images are, however, affected by cloud-cover and their availability is limited to daylight conditions. For Sentinel-1 we use a dual polarization of VV,VH and for Sentinel-2 the bands Blue, Green, Red, NIR, SWIR1, SWIR2.

To showcase our approach, we selected two different study areas:

(a) Pakistan: Indus flood plain near Sukkur, see Fig. 1(a). The area is dominated by hot desert climate. There is almost no rainfall and the surface water extent and inundation of the Indus river is caused by melting snow in Himalaya in spring and summer and monsoon rains in summer. This study site was selected because of the severity and seasonality of floods. 1035 Sentinel-1 and 2104 Sentinel-2 scenes from 2018 to 2022 were used for the analysis.

(b) Mozambique, see Fig. 1(b). The area is dominated by tropical savanna climate with more arid patches further inland. Floods are usually caused by cyclones between October and April. This study site was selected because of the severity of floods across a large and diverse area. 4921 Sentinel-1 scenes from 2019 to 2022 were considered in this analysis.

3. Methodology

Each satellite scene from within an area of interest (AOI) and within a monitoring period is semantically segmented into water, non-water and nodata pixels. For the segmentation we are using a Multi-Spectral Flood Processor based on convolutional neural networks (MS-FP-cnn) (Wieland and Martinis, 2019) for Sentinel-2. This generic processor for water segmentation of optical satellite data has been optimized to deal with the variability of spectral signatures across different types of water bodies. Clouds and cloud shadows are masked with the UKIS Cloud Shadow Mask (Wieland et al., 2019). For the segmentation of Sentinel-1 imagery we use an adaption of the same CNN (Helleis et al., 2022) within the same workflow. Both CNNs were trained using the data of Wieland et al. (in review). The resulting raster data are stored in a DGGS for further analyses. This workflow is shown in Fig. 2, which provides a general overview of the methodology introduced in this work. In short, the presented workflow starts with the segmented imagery, then inserts the results into the database before the time-series analysis and anomaly detection is done on each H3 cell.

As DGGS we use the H3 grid system (Sahr, 2011; Brodsky and Contributors, 2018) a hierarchical grid providing 16 spatial resolutions and a 64-bit unsigned integer index, identifying each cell. The grid cells have a hexagonal shape.

Fig. 3 exemplifies the spatial hierarchy and interconnections between grid-cells at different resolution. The main reasons for selecting

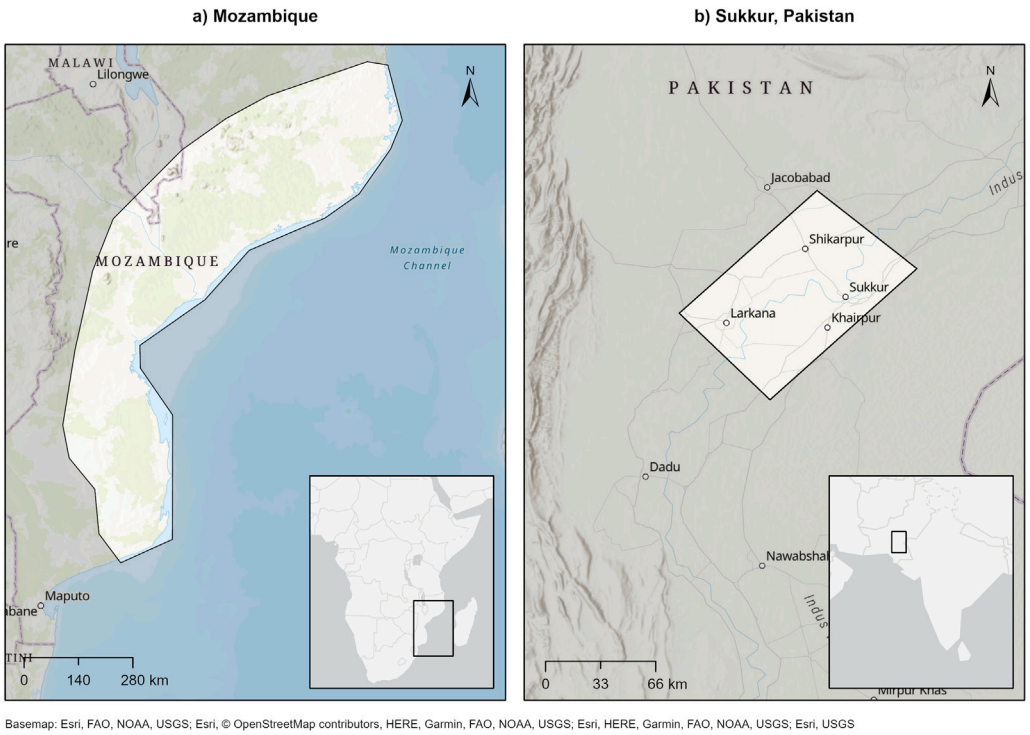


Fig. 1. Overview map of study sites. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

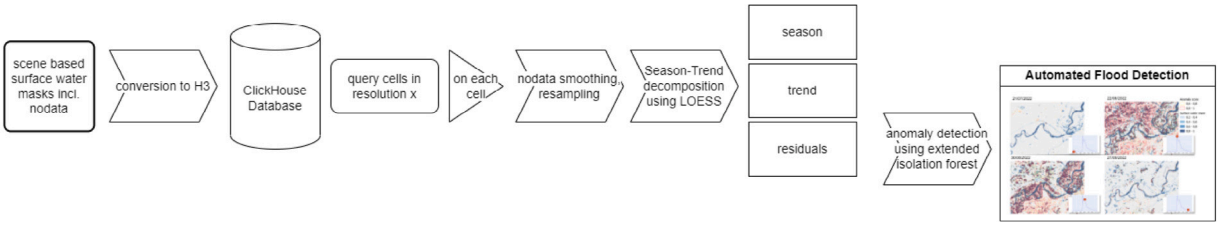


Fig. 2. Simplified flowchart of methodology steps.

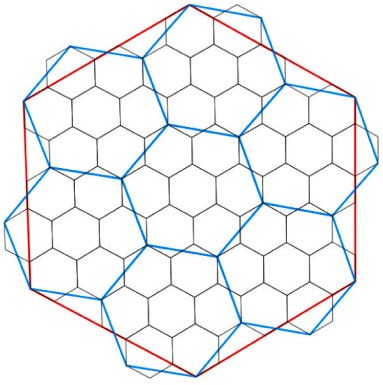


Fig. 3. H3 hierarchy, exemplary showing 3 H3 resolutions with black as blue's children and red as its parent. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

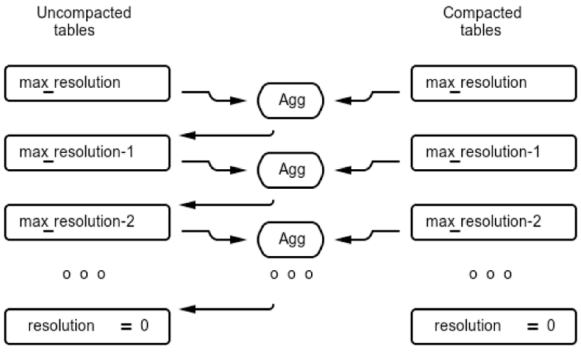


Fig. 4. The aggregation process creates uncompacted H3 resolutions.

the H3 grid system were (a) the good documentation, (b) the range of existing programming language integrations, (c) the pleasing visualization possibilities of the hierarchical hexagons structure and (d) the fact that it is currently widely adopted.

The output raster masks from the semantic segmentation are converted into the H3 grid system in several steps using the Python bindings of h3ron-ndarray (Mandery and Contributors, 2022). When both are available, raster masks created with Sentinel-1 or Sentinel-2 are created equally throughout the following process. First, the raster masks are cut into rectangle shaped tiles, represented by array slices. Tiles exclusively containing nodata are ignored. All remaining tiles are distributed across a thread pool to derive the H3 index of each pixel, based on the cell containing the centroid of the pixel. We choose H3 cell resolution 12 as target resolution with an average edge length of

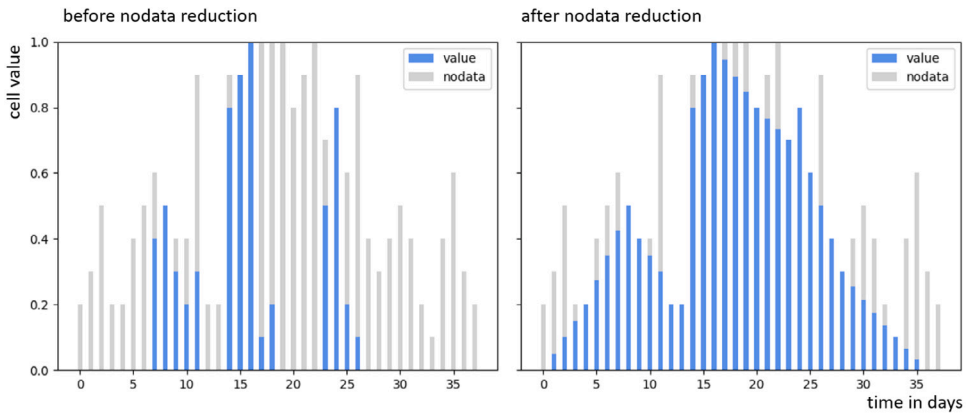


Fig. 5. Schematic illustration of nodata smoothing. Nodata and value filling of cell before nodata reduction on the left. Nodata and value filling of cell after nodata reduction, at the end of the iterative optimization on the right.

approximately 9.4 m because it is relatively close to the raster's resolution without oversampling it. Values for cells in lower H3 resolutions are derived from automated aggregation based on their mean relative to the cell area.

These cells and their associated values are stored in a ClickHouse (ClickHouse, 2022) database using the ReplacingMergeTree storage engine and the compaction process provided by the ukis_h3cellstore library (Mandery and Contributors, 2023). The ReplacingMergeTree engine allows to insert large amount of data into the table. Instead of rewriting the data in storage during insert, merging and removing duplicates is happening in the background. Duplicates with the same $h3index, timestamp$ sorting key value are removed. As ClickHouse is a column-based storage system, the values of a single column are stored together (Kleppmann, 2017). This property, combined with the order key leads to relatively homogeneous chunks of data getting stored in the files of the individual columns. H3 cells spatially located closely together show only minor numeric differences in the value of their indexes. We take advantage of that by applying Z-Standard compression (Meta Platforms, Inc., 2022) to data stored in the database system. These features optimize the storage for reading by reducing seeks on disk as well as the amount of data to be read to satisfy a database query. The aggregation to derive values in lower resolutions is based on the cell's area and the cell's parent cell, as well as the cell's siblings' values and areas. To reduce the number of cells stored in our database the hierarchical character of the H3 grid is used to compact data belonging to the same scene. This compaction and aggregation process is showcased in Fig. 4. The uncompacted resolution N is created by applying an aggregation function to the combined contents of the compacted and uncompacted tables of resolution $N + 1$. This process is repeated consecutively until resolution 0 is reached. Compaction reduces the number of cells or rows necessary in the database by forming the lowest possible resolution cell for children with the same attributes. This represents a form of loss-less compression as de-compaction using the child cells re-creates to original data again. To take advantage of the partitioning functionality built into ClickHouse, low resolution parent cells together with time intervals are used for spatio-temporal partitioning the stored data on disk.

At this point the data is stored in the database and can be referenced using the cell's H3 index. There are three attributes written to the database for each cell.

- is_water: The area percentage of the cell covered with water. A value of 0.0 marks a cell without detected surface water, the value 1.0 is stored for a cell completely covered with water.
- is_nodata: The area percentage of the cell affected by nodata values (e.g. cloud cover). Nodata values are required to compute the proportion of water per grid cell during aggregation of grid cells at different resolutions.
- is_border: Boolean flag to mark scene border cells. This flag is needed to exclude border cells from further analyses because of their skewing effects.

Having no water detected by semantic segmentation is the most frequent case. Therefore, this information is not explicitly stored in the database to further reduce the number of rows.

When aggregating to lower resolutions it is assumed that the water values of child cells are complete. At scene borders, during conversion to lower H3 resolutions, the water observations might be incomplete because child cells outside the scenes do not have data available in the set. This can cause an underestimation of water coverage as the aggregation advances to lower resolutions. This effect is exaggerated in time-series because scenes' footprints are often stacked at a similar location. Cells are flagged as border cells when they are directly within the exterior ring of the scene's footprint. When analyzing, we query all cells in the AOI which had water at one point, excluding the ones flagged as border cells. During analysis, the information stored in such cells is not considered, but misleading border effects in visualizations and analyses are mitigated.

After the data is stored in the database it can be efficiently accessed using ukis_h3cellstore (Mandery and Contributors, 2023) to further clean and prepare the data on-the-fly. To be able to use time-series analysis with a daily interval, multiple satellite observations for the same cell at the same day are merged using their mean value. To fill gaps of cases without observation on a day, the data is resampled and interpolated to daily data with a forward fill.

Fig. 5 visualizes the proportion of a cell which is covered by is_water and is_nodata before and after applying the following nodata smoothing algorithm. Nodata, mainly present in Sentinel-2, will be smoothed out using an optimization algorithm under the assumption that a cell's $no\_water = 1.0 - is\_water - is\_nodata$. The aim of the smoothing is to reduce the effect of is_nodata on the performance of the analysis and therefore replace it with the most likely is_water value. For this the minimization criterion assumes that a cell's is_water value is more probable to be stable than fluctuating. The proportion of is_nodata of each time-step acts as bound for how much has to be added to is_water to minimize the effect of nodata. In Fig. 5, this means that only the gray nodata bar can be filled. Using these assumptions, we move a sliding window over the vector of time steps and calculate the absolute difference of is_water to the timestep before. This difference is iteratively decreased by adding to is_water. By providing a limit of iterations at the beginning and end of the timeseries, we avoid over-reducing for cells with an extensive amount of nodata. This reduces the influence of nodata and smooths the overall curve representing the water coverage over time.

At this point there is a daily value of relative surface water coverage for each cell which is called *optimized water*. With these *optimized*

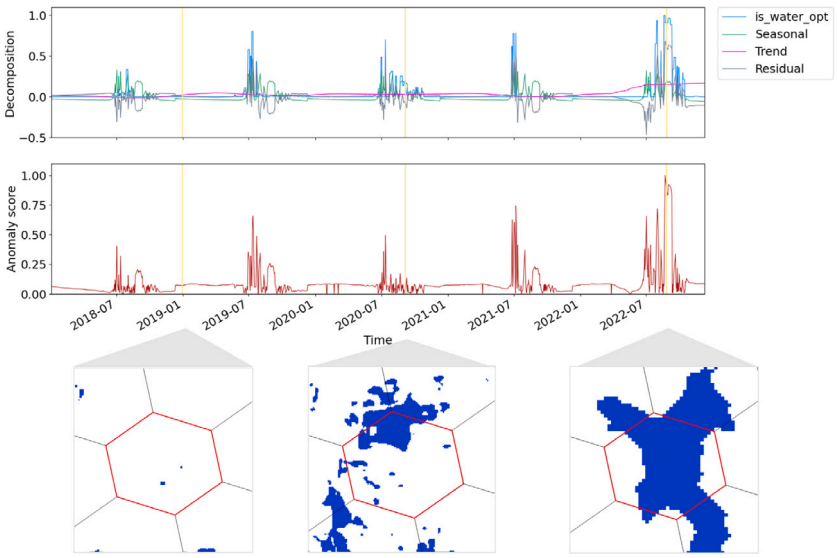


Fig. 6. Timeseries decomposition and anomaly score of a cell (88425400edfffff) with exemplary water masks for three timestamps.

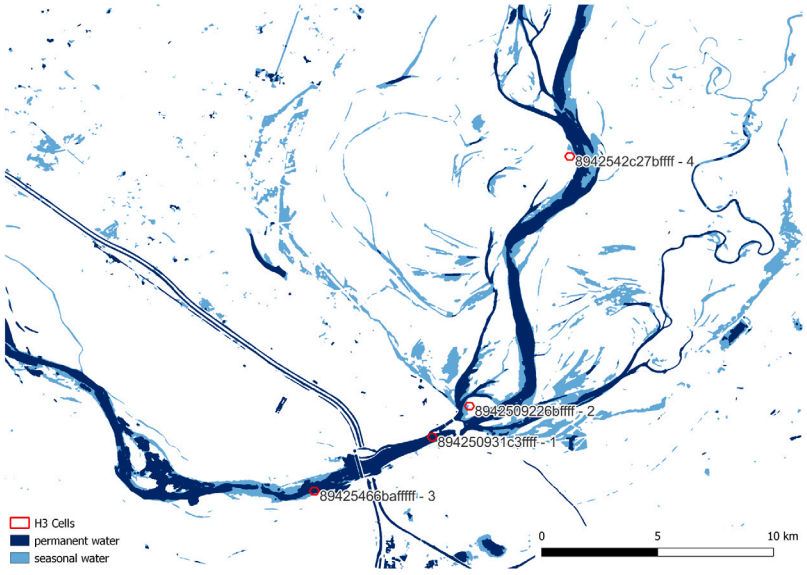


Fig. 7. Location of selected H3 cells in resolution 9 and their placement on Reference Water Masks (Martinis et al., 2022) for September, derived from data of 2020 and 2021 around Sukkur, Pakistan.

water values a univariate time-series is created for each H3 cell on the resolution of choice. The time-series is decomposed into residual, season and trend via a Seasonal and Trend decomposition using Loess (locally estimated scatterplot smoothing) (STL) (Cleveland and Cleveland, 1990) using a period cycle of 365 days. STL was selected because of its robustness and versatility. With such decomposition, seasonal changes and long-term trends are automatically anticipated. The change of a river's course after a flooding event, for example, will therefore not be classified as anomaly for the whole remaining time-series. An anomaly score is calculated on the remaining residuals using an extended isolation forest (EIF) (Hariri et al., 2021). The EIF was selected because of its ability to work on complex cases, not limited to outside of range outlier detection like interquartile range and because it provides flexibility to also work with multivariate timeseries. This algorithm uses the assumption that anomalies are few and different from non-anomalous or normal observations. The extended isolation forest is built out of decision trees cutting the data into sub-samples with a random slope on a random intercept for the branch cut. Data of the whole timeseries is used during this training phase. The depth a data point reaches in these trees is the basis for the anomaly score. The settings used for the EIF are 20 trees, a sample size of the training data of 20 and the extension level of 1. Before normalization the EIF anomaly score is clamped to the range $[0.44, 0.6]$ and normalized to the range $[0, 1]$. This is done to have a more universal scale instead of relying on implementation details of the EIF. The range is used for the following figures in Section 4.

To assign anomaly scores to the input values, the results of the anomaly detection are joined to the non-resampled timeseries using a join by key distance with forward strategy. This is necessary because we truncate to the day during resampling and the information about the hour of the day is lost.

The timeseries analysis and anomaly detection described above are applied to each H3 cell over an AOI with the H3 resolution of choice. Due to the parallel character of this workload, this process is spread across a pool of threads for a runtime reduction correlated to the number of available CPUs. The result is a complete picture of water coverage in each cell. Fig. 6 showcases the principle on a single cell. The curves on top represent the time-series decomposed in its parts of

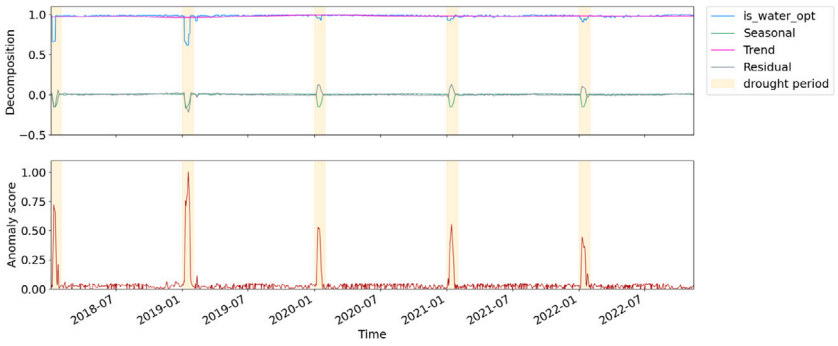


Fig. 8. Timeseries decomposition and anomaly detection of cell 894250931c3ffff - 1 at a river location, with seasonal less water coverage in January.

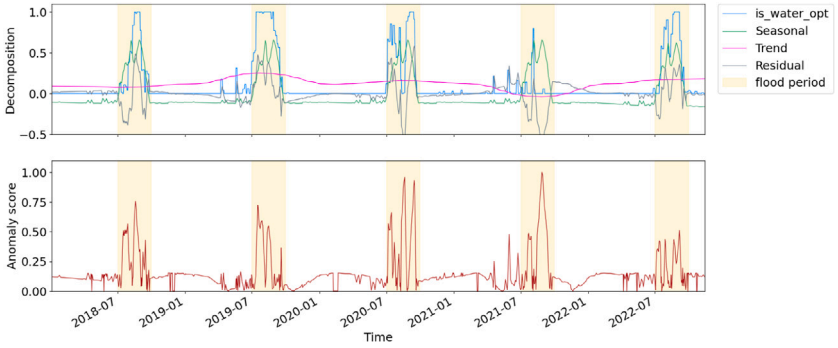


Fig. 9. Timeseries decomposition and anomaly score of cell 8942509226bffff - 2 which is prone to seasonal flooding in between July and September.

seasonal, trend and residual. In the frame below, the red curve is the anomaly score of the residual at each timestep. The three frames at the bottom show the water mask derived from Sentinel-1 at different steps in time.

To evaluate the results, the methodology is applied to areas prone to hydrological anomalies. Time and location of anomalies can be verified by comparing the derived results with data from other sources.

4. Results

For the results presented in this work we are using the AOIs around Sukkur, Pakistan and Mozambique (see Fig. 1). First, the results and behavior on single H3 cells is shown, then the results across whole study areas is showcased. Finally, the performance of the implementation is presented.

4.1. Results of TSA on single cells

To showcase results and the behavior of our proposed methodology on single cells four different kind of locations were exemplary selected, see Fig. 7.

Figs. 8 and 9 show two examples of timeseries with seasonal patterns. The anomaly score differs in each season and is higher when the season behaves different from others. To distinguish anomalies and to flag floods or droughts, a threshold has to be set on the anomaly score. For simplicity, the threshold is set to 0.8. Positive anomalies have residual values above 0.0 and negative anomalies below 0.0.

Fig. 10 shows that the smoothing and resampling of (periodic) nodata works as expected. The amount of nodata gaps are created by scene border effects of Sentinel-1. Even though the *is_water* observation is interrupted by regular *nodata* peaks at scene borders, the optimized water-value *is_water_opt* looks smooth and is interpolated as expected. The decomposition and further analysis of the residuals are based on the *optimized water*-value. When combining radar and optical sensors a similar smoothing effect would occur in scenarios of prolongated cloud-cover where the nodata proportion of optical observation is filled with known *is_water* of the next radar observation.

When comparing flood extents with cells showing an anomaly it is important to select the appropriate time range for our time-series analysis. Like seen in the comparison of Fig. 11, the time after the anomaly occurs can influence the anomaly score of a cell. On the left the analysis stops earlier, showing that water coverage change is initially marked as anomaly: On the right, the change becomes the new normal. In this example, the suddenly rising water coverage in summer 2020 is initially seen as anomaly when the TSA is performed at the turn of the year. This can be observed on the left tile of the figure. Right, on the other hand, the water coverage becomes the new normal, hence also the rising water in summer 2020 is to be expected according to the long-term trend and therefore has a low anomaly score.

4.2. Results of TSA across an AOI

To show how the introduced approach works in comparison to solutions using static reference water masks, we compare results with the one of Martinis et al. (2022). Both methodologies rely on the same water segmentation data.

Fig. 12 shows this comparison on the basis of the flooding around Sukkur of 11th of September 2020. The map shows the anomaly score based on our introduced method and the observed water difference to permanent and seasonal reference water masks. The figure is zoomed and does not show the whole AOI to make it easier to distinguish between different layers and H3 cells. Comparing the results, one can observe strong similarities with locations of cells with high anomaly score. Cells tend to be flagged as anomalies, where the observed water extent is different from the combination of seasonal mask and permanent reference water mask. The differences in some patches come from the optimization of *is_water* in the timeseries approach and the fact that some floods are not anomalies on the long term like also observed in Fig. 11. Trend is not considered when using static reference water

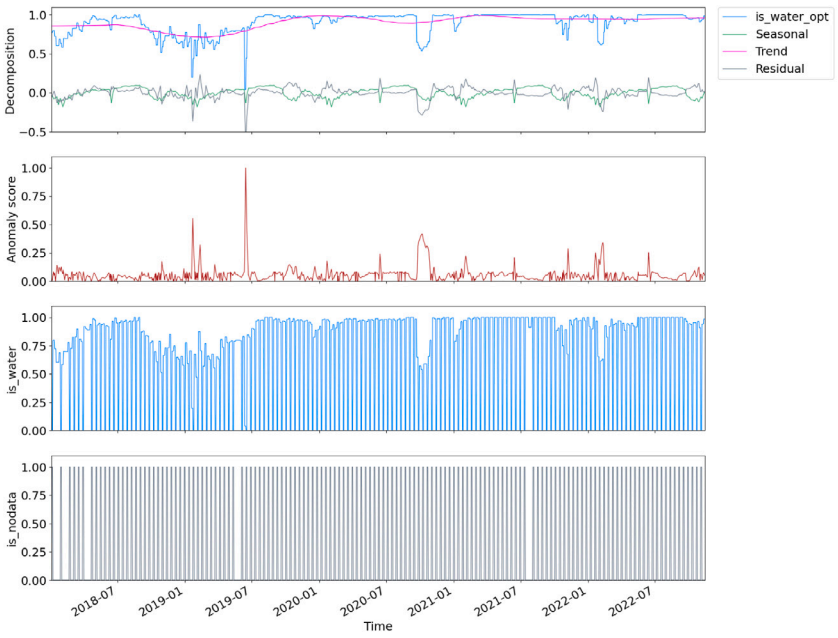


Fig. 10. Timeseries decomposition and anomaly detection of cell 89425466baffffff - 3 showing *is_water* and *nodata* and optimized *is_water_opt*.

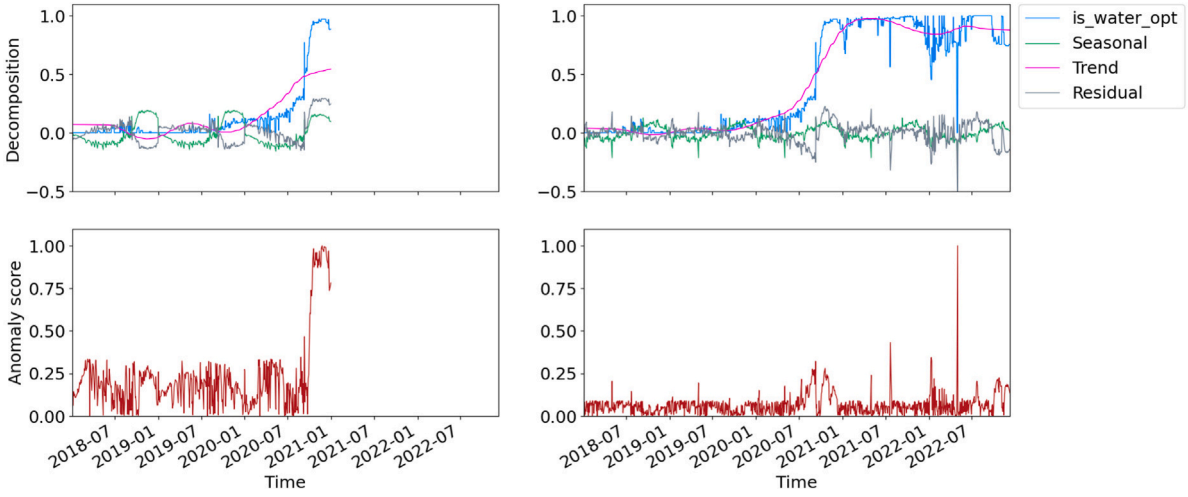


Fig. 11. Timeseries decomposition and anomaly detection of cell 8942542c27bffff - 4 at a place where the course of river changes. On the left the analysis only includes data until December 2020.

masks. Seasons are also not simply based on the water extent of the same month in the two years before the observation, but rather on the season component of STL. Additionally, the observed water layer was created from Sentinel-2 scenes on the 10th of September, because there were no scenes within the AOI in our dataset for the 11th of September. All other layers are based on a mix of Sentinel-1 and Sentinel-2. Our methodology can still create results and flag anomalies here due to the resampling, optimization and nodata handling. This shows, that for a date with no observation or with extensive cloud-cover there would still be meaningful results.

As another example, the comparison with the reference water mask product introduced in Fig. 12 are repeated in Fig. 13 for the flood in summer 2022 in Pakistan around Sukkur. It is visible that the high number of anomalies do not necessarily refer to the biggest extent of surface water. This can be observed in the southern part of the map, where also small relative changes of surface water are flagged as anomaly. The observed water layer is again only based on Sentinel-2 scenes because there were no Sentinel-1 scenes available for that date.

Fig. 14 showcases the situation and anomaly detection around Sukkur in four steps in time, before, during and after the flood in summer 2022 (GDACS, 2022b). In the first frame, before the flood, anomalies are detected in the north-western part of the map. These anomalies are negative, hinting to a dry period. The next two frames show the two peaks in the anomaly detection, where across the whole AOI the most cells are classified as anomaly. The last frame shows the situation after most of the water is gone.

Fig. 15 shows the results of the time-series analysis and decomposition across an AOI covering the eastern part of Mozambique. The sum of the timeseries having a positive anomaly across time hints to potential flooding events when the appertaining cells cluster. To showcase the results, 9 days have been selected and GDACS-alerts have been filtered to find the event belonging to the anomaly clusters. Due to the thickness of the outline of cells having an anomaly in their timeseries and the size of the AOI, the floods look exaggerated, especially when anomalies are not formed around a cluster. Day 1, 3, 4, 7 and 9 show scenarios where time and location of GDACS-alert and the approach showcased here coincide. The time range of the event on the third selected day was only

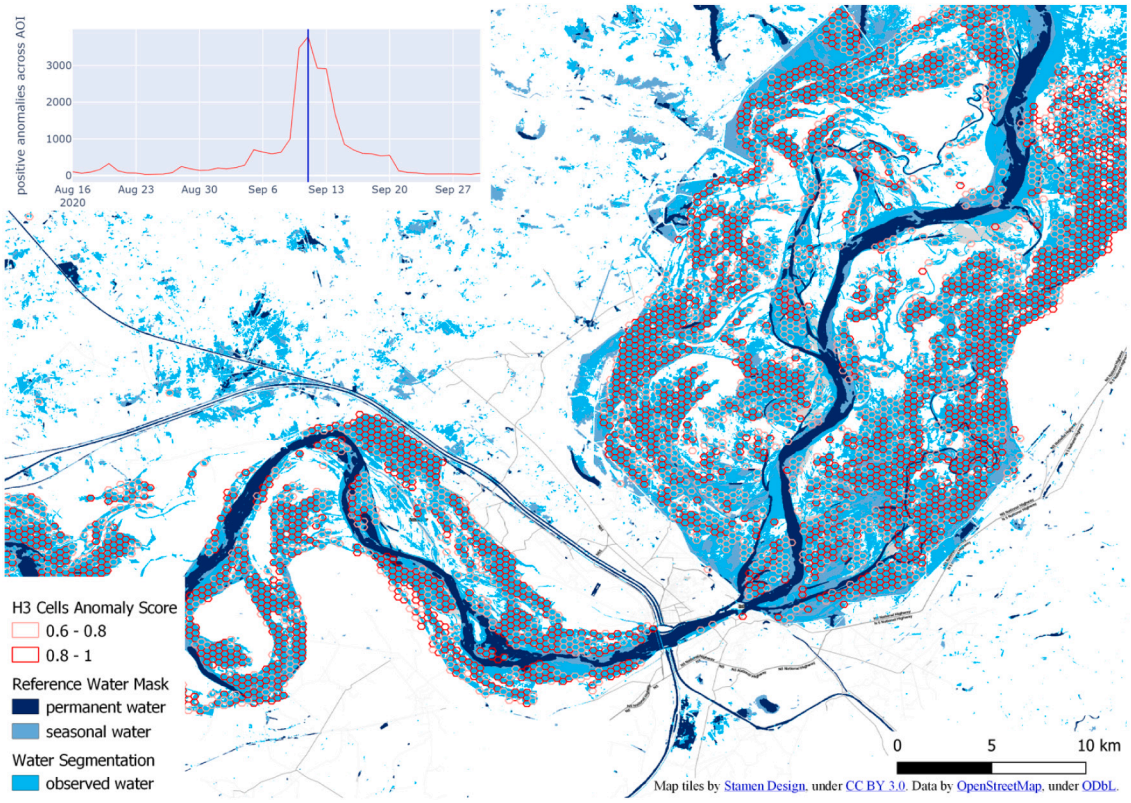


Fig. 12. Flood extent (observed water) drawn from reference water masks in relation to anomaly H3 cells in resolution 9 for 11th of September 2020, zoomed into the region around Sukkur. The diagram in the top left corner shows the number of cells flagged as anomaly across the whole AOI.

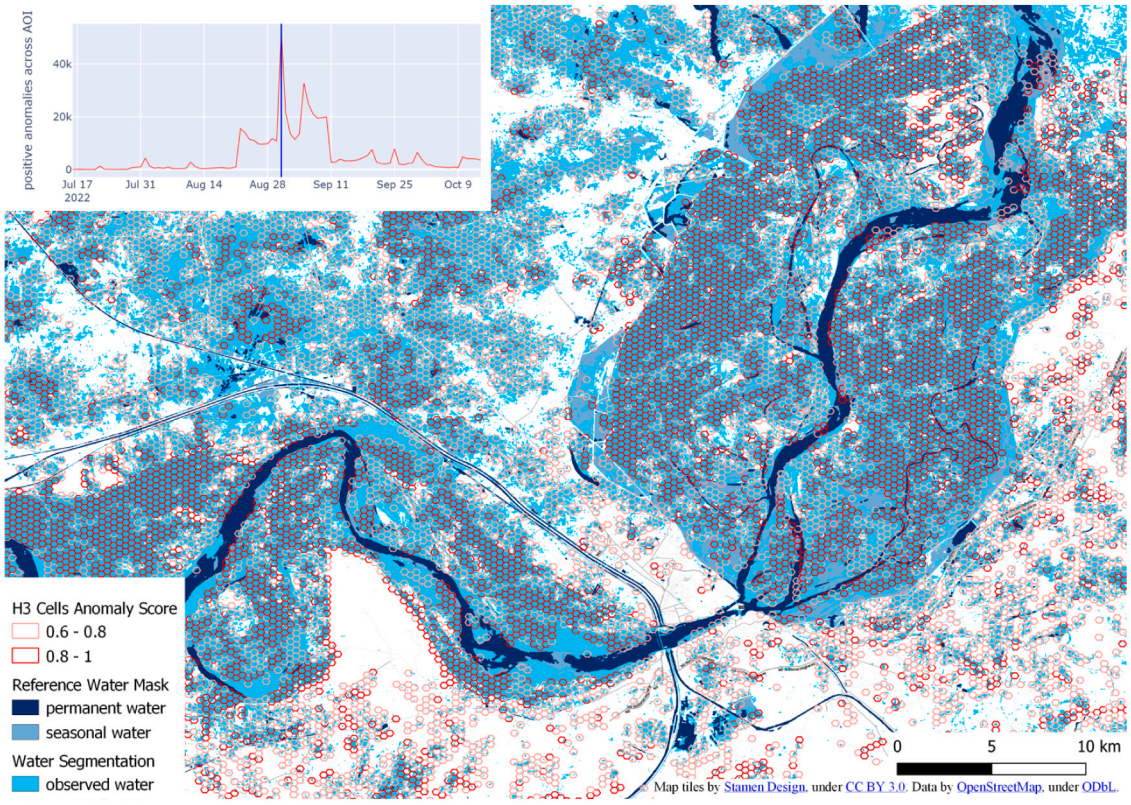


Fig. 13. Flood extent drawn from reference water and segmented satellite imagery of the same dates in comparison to anomaly H3 cells based on Sentinel-1 and Sentinel-2 in resolution 9 for 31st of August 2022. The diagram in the top left corner shows the number of cells flagged as anomaly across the whole AOI.

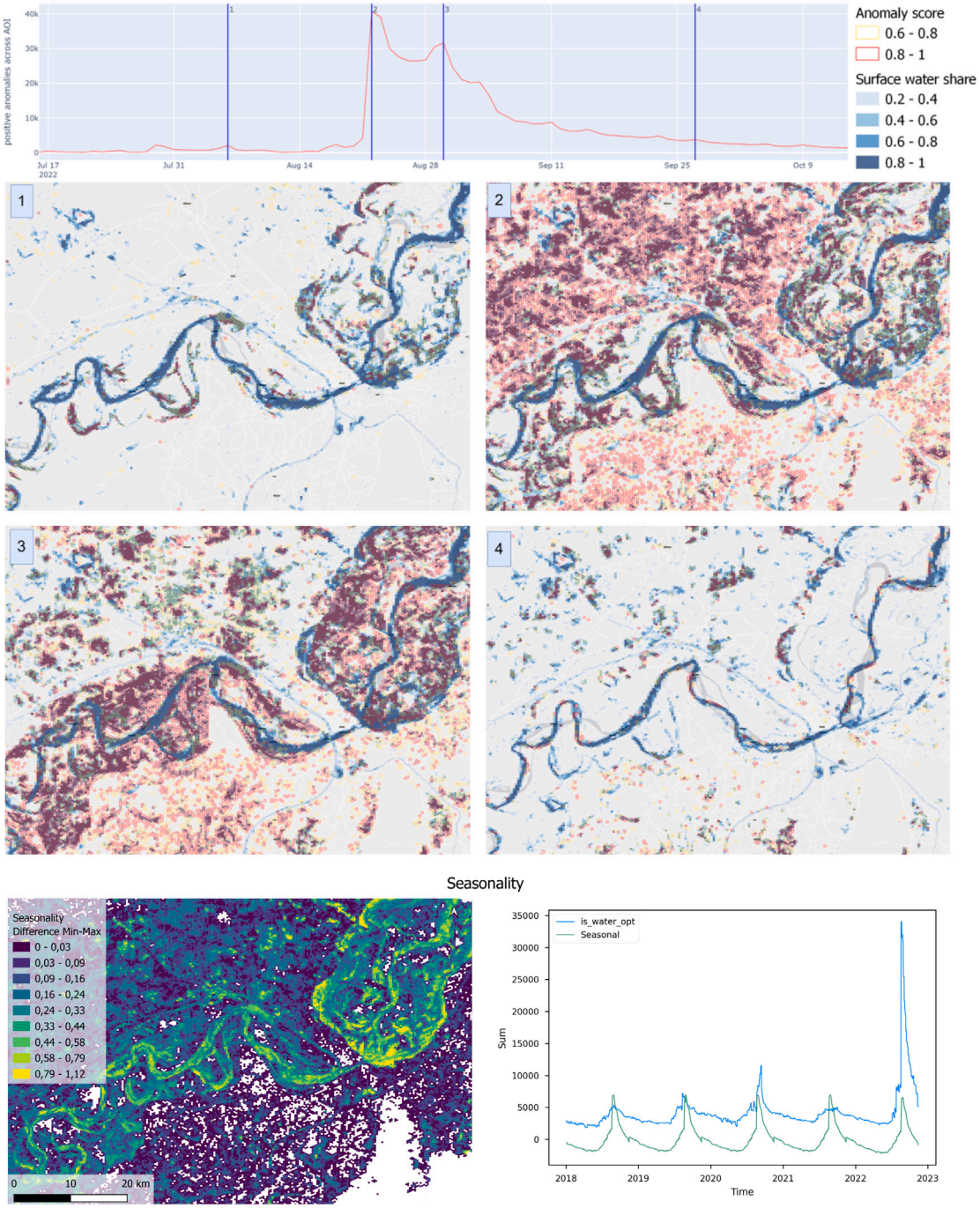


Fig. 14. Time-series analysis of area around Sukkur before and after the flood of summer 2022 based on water segmentation on Sentinel-1 and H3 resolution of 8. The linegraph shows the anomalies over time, the maps show the anomaly score and corrected water share of cells. The intensity of seasonality is shown as the difference between the minimum and maximum of the seasonal time series component. The bottom line plot shows the sum of observed water values across time. Sources Basemap: Esri, HERE, Garmin, FAO, NOAA, USGS, ©OpenStreetMap contributors, and the GIS User Community.

one day according to GDACS. However, according to FloodList (2020) also the Sofala Region was heavily affected, especially around the date of the peak in anomalies. The detected anomalies cluster around the Sofala region, further south of the GDACS polygon, showing that the flood was well detected. The peaks at Day 2 and 6 (and also the other investigated peaks without flood reference in official sources or in the news) paradigmatically show that the anomalies are caused by natural phenomena which are not considered to be a harmful flood. There are for example flooded clusters around mangrove areas south of Beira and, especially visible on day 6, flooded ponds further inland which are usually dry. They are shown in the small Sentinel-2 extract. Therefore, they do not coincide with any GDACS-alerts. Day 5 was classified as flash flood (GDACS, 2020) and also the underlying segmented imagery did not catch the event. The alert of day 8 was outside the AOI, the number of anomalies hints to a rainy period.

4.3. Performance

An inspection of the database contents after the ingestion of the segmented imagery shows the effectiveness of the compaction process as detailed in Table 1. The number of cells written to the ClickHouse database gets reduced by more than 98%. As the aggregated resolution

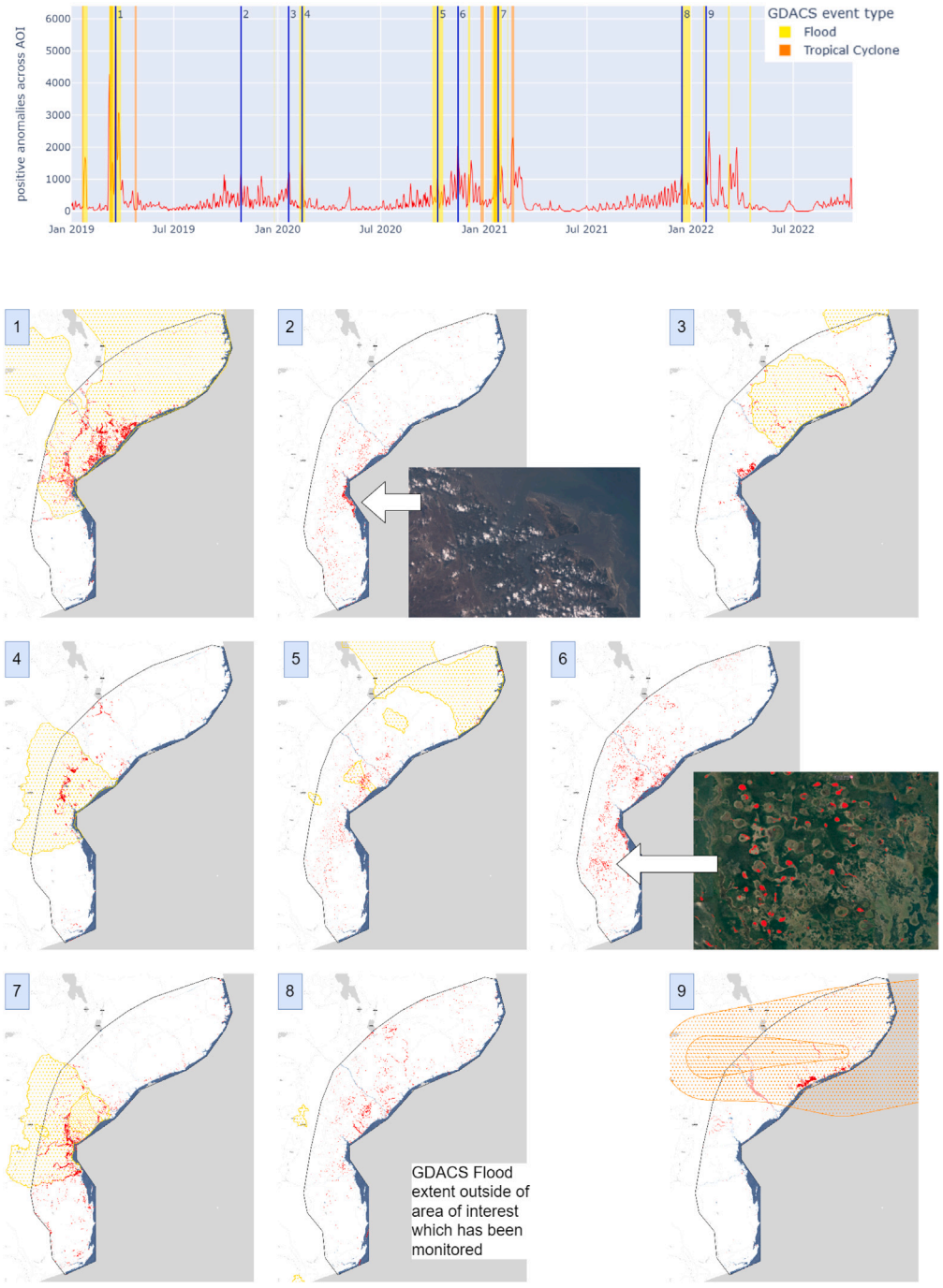


Fig. 15. Time-series analysis of area of interest used in Schneibel et al. (2022) based on water segmentation on Sentinel-1 and H3 resolution of 8. Graph shows the anomalies over time, the (GDACS, 2022a) alerts for floods (yellow) and cyclones (orange) in country Mozambique for the same time range along with 9 maps showing exemplary analysis results, Sources Basemap: Esri, HERE, Garmin, FAO, NOAA, USGS, ©OpenStreetMap contributors, and the GIS User Community. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

levels are able to reuse the compacted cells of the highest resolution level, the space-saving effect of the compaction process gets even increased.

Querying the database, decomposition and analysis of 58 months long daily resampled timeseries on resolution 10 H3 cells within the approximately 7550 km^2 large AOI around Sukkur, Pakistan takes around 6 minutes on a laptop with 6 Intel(R) Core(TM) i7-10850H CPU @ 2.70 GHz and 32 GB RAM and results in approximately 26 GB of ZSTD-compressed parquet files. TSA is only performed when there was a water observation at least once in the cell. When using a lower resolution of 9, the analysis is around 7 times faster.

5. Discussion

Based on the results presented in the previous chapter, the proposed methodology provides a suitable solution to automatically detect hydrological extreme event such as floods and droughts using earth observation data. This was shown on two different study areas in Mozambique and in Pakistan. Generally, the approach is applicable to any segmentation of surface water under the precondition that the segmentation quality is consistent over time and nodata information is available. It is also efficient enough for flood-monitoring on AOIs, where monitoring can be performed in lower resolution and in case a

Table 1
Effect of cell compaction on the Sentinel data ingested into H3 resolution 12. These database cell counts exclude aggregated resolution levels.

Sensor	Number of scenes	Average number of cells per scene with compaction	Average number of cells without compaction	Reduction percentage
Sentinel-1	1.030	667.285	138.523.157	99.52%
Sentinel-2	2.192	47.586	4.210.711	98.87%

certain anomaly threshold is reached, the analysis can be repeated in higher resolution. With the properties of the DGGS, a change in resolution is simple, but a higher resolution requires more processing and storage resources. Doing monitoring in lower resolution is beneficial for monitoring because the TSA has to be constantly repeated.

The key benefit is the ability to dynamically adapt to the scenarios within the AOI and its independency from resolution and sensors. The seasonal curve can look and behave different in each cell and is solely based on pattern of previous observations. Considering seasonality also improves differentiation between unusual flood events and seasonal inundation such as practiced in rice agriculture, for example. Sudden cuts of seasons or steps caused by pre-processed seasonal water mask datasets, e.g. at the turn of a month are eliminated. It also works seamlessly when results of sensors like Sentinel-1 and Sentinel-2 are mixed, as long as the segmentation quality and resolution is reasonably comparable. This can be used to fill gaps in satellite overpasses or periods of long cloud covers. Gaps in observations are automatically filled using resampling of the information available. Currently, even though it is queryable in the database, no sensor is preferred over other ones. A peak in the detection of anomalies over an AOI does not represent the date of the biggest water surface extent, but rather the date of the biggest change. Especially in cases of strong trend, like in Fig. 11, this behavior can be misleading. To find anomalies, it is not necessary to use external or pre-calculated water masks with fixed thresholds. Everything can be calculated on-the-fly and rely on well-established algorithms like STL or EIF. Because we rely heavily on well-established components, we assume that the proposed methodology is transferable to other study areas.

It should be noted however, that a validation of the results remains challenging due to the lack of reference data. Using the polygons provided by GDACS help to get an indication of the general usefulness.

Like shown, not every cluster of anomalies hinting to a flood can be considered a disaster. In many cases humans and infrastructure might not have been negatively affected, like in case of mangroves. The observed water extent can still be a statistical anomaly.

The size of residuals per timeseries and cell cannot be compared to the size of residuals in surrounding cells. That means each timeseries stands for itself. For example, if the water coverage is linear, a small residual will receive a high anomaly score like shown in Fig. 8 even though the real change might be neglectable. This could happen when a cell remains completely dry throughout the timeseries and at one observation a single pixel was (mis-)classified as water. Flood monitoring will therefore profit from longer timeseries and the severity of a flood is better drawn from the number of clusterings of anomalies than on a single cell.

When a timeseries is continuously recalculated as soon as new observations become available, like for example during a monitoring process, the anomaly score of a cell for the same timestep can change. Such behavior was already shown in Fig. 11 and can be explained by the adaption of trend and season with new observations.

Doing the TSA on a similar H3 resolution than the original rasters is expensive in computation, but possible in theory. The results are expected to be noisy, because the values at that resolution are booleans and the algorithms used works better on real numbers. This effect can be avoided to some extent by taking neighboring cells into account.

Another interesting aspect is the combination of different sources and the influence on results. Figs. 14 and 15 show a different anomaly curve for the same AOI Fig. 1. The first was created from a combination of Sentinel-1 and Sentinel-2, while the second one only uses Sentinel-1. There are different detection rates with both sensors and the segmentation models have different properties. This allows us to profit from the strength of each sensor, for example that Sentinel-1 also provides regular results in cloudy periods. Using a DGGS simplifies combining and querying different sensors. That means an analysis can be done using a combination of sensors or only using results of one of them.

Generally, the properties of H3 allow for many further extensions. Using a DGGS provides the possibility to add contextual data, like population and infrastructure datasets to the analyses, using a simple join by H3 index.

Fig. 15 shows that a smaller AOI than the one used in this study might be beneficial to clearly identify floods based on the anomaly score as flooding are common throughout the whole period. This is at least the case when the total sum of anomalies is used as an indicator.

For the future, a flood alert might depend on cluster-detection of anomalies and auxiliary information, like the population count of cells. Even with the currently arbitrary threshold of anomaly score above 0.8 to flag anomalies, the results look promising. Such thresholds could vary depending on the AOI in the future. On top of that, if anomalies point to a scenario where no flooding or drought can be identified in the real world it can help to improve segmentation. They might hint to saliences of the segmentation model and can help to identify patterns which could be added to the training dataset.

As pointed out earlier, using a DGGS creates many new opportunities. However, before benefiting from its properties it adds a layer of complexity. There is custom tooling necessary which has to be maintained. Conversion and storage add to the costs as well. Nevertheless, using H3 enables us to do things which would otherwise be impossible. Such things are a low-cost change in resolution or the normalization of different sensor results to cell values.

6. Conclusions and outlook

We have shown that combining several well-established methodologies and algorithms and applying them to a new use cases improves automated flood and drought detection. By introducing H3 as DGGS we simplify binning of surface water extents across time. Using a DGGS enables us to do fast resolution traversal necessary for timeseries analysis on each H3 cell within the AOI. With decomposition of timeseries we are able to detect anomalies on residuals, taking season and trend into consideration. This eliminates the need of pre-calculated and static reference water masks. We showed the general functionality of the introduced approach on real world scenarios in Pakistan and Mozambique.

In the future, we plan to further exploit the benefits of H3 to automatically detect cluster of cells showing an anomaly. This could be combined with repeating the analysis in different resolutions to create a real monitoring service as well as transferring the methodology to different study sites (e.g. agricultural areas). Anomaly clusters are hydrological anomalies which can then be combined with any desired auxiliary data, like population or infrastructure. Using H3's properties, we can use graph-algorithms on the changing scenarios to provide even more information to emergency responders faster. Additional data could also play a role in a potentially multivariate anomaly detection not limiting our approach to surface water extents. We also plan to do analyses with focus on hydrological droughts. Furthermore, we aim to experiment using weighting factors for different sensor types.

CRediT authorship contribution statement

Florian Fichtner: Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Nico Mandery:** Conceptualization, Formal analysis, Methodology, Software, Visualization, Writing – review & editing. **Marc Wieland:** Conceptualization, Methodology, Writing – review & editing. **Sandro Groth:** Conceptualization, Data curation, Visualization, Writing – review & editing. **Sandro Martinis:** Funding acquisition, Writing – review & editing. **Torsten Riedlinger:** Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

ukis-h3cellstore is available on https://github.com/dlr-eoc/ukis-h3cellstore

Acknowledgments

This work was conducted partially by the DLR-funded project VisPlore and through DLR internal funding.

References

Bai, B., Tan, Y., Zhou, K., Donchyts, G., Haag, A., Weerts, A.H., 2022. Time-series surface water gap filling based on spatiotemporal neighbourhood similarity. Int. J. Appl. Earth Obs. Geoinf. 112, 102882. http://dx.doi.org/10.1016/j.jag.2022.102882, URL: https://www.sciencedirect.com/science/article/pii/S156984322200084X.

Bormann, H., Pinter, N., Elfert, S., 2011. Hydrological signatures of flood trends on German rivers: Flood frequencies, flood heights and specific stages. J. Hydrol. 404, 50–66. http://dx.doi.org/10.1016/j.jhydrol.2011.04.019, URL: https://www.sciencedirect.com/science/article/pii/S0022169411002605.

Brodsky, I., Contributors, 2018. H3: Hexagonal hierarchical geospatial indexing system. https://h3geo.org/.

Brown, C.F., Brumby, S.P., Guzder-Williams, B., Birch, T., Hyde, S.B., Mazzariello, J., Czerwinski, W., Pasquarella, V.J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., Tait, A.M., 2022. Dynamic world, near real-time global 10 m land use land cover mapping. Sci. Data 9, 251. http://dx.doi.org/10.1016/j.rse.2022.113077.

Cleveland, R.B., Cleveland, W.S., 1990. STL: A seasonal-trend decomposition procedure based on loess. J. Off. Stat. 3, 3–33, URL: https://www.proquest.com/scholarly-journals/stl-seasonal-trend-decomposition-procedure-based/docview/1266805989/se-2.

ClickHouse, 2022. Fast open-source OLAP DBMS. https://clickhouse.com/.

Copernicus Emergency Management Service, 2022. List of EMS rapid mapping activations. https://emergency.copernicus.eu/mapping/list-of-activations-rapid.

FloodList, 2020. Mozambique – floods leave 28 dead, 58,000 affected, says UN. https://floodlist.com/africa/mozambique-floods-january-2020-update.

GDACS, 2020. Overall Green alert flood for mozambique FL 1100649. https://www.gdacs.org/report.aspx?eventid=1100649&episodeid=1&eventtype=FL.

GDACS, 2022a. Global disaster alert and coordination system. https://www.gdacs.org/Alerts/default.aspx.

GDACS, 2022b. Overall red alert flood for Pakistan FL 1101522. https://www.gdacs.org/report.aspx?eventid=1101522&episodeid=1&eventtype=FL.

Hariri, S., Kind, M.C., Brunner, R.J., 2021. Extended isolation forest. IEEE Trans. Knowl. Data Eng. 33, 1479–1489. http://dx.doi.org/10.1109/TKDE.2019.2947676.

Helleis, M., Wieland, M., Krullikowski, C., Martinis, S., Plank, S., 2022. Sentinel-1-based water and flood mapping: Benchmarking convolutional neural networks against an operational rule-based processing chain. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 2023–2036. http://dx.doi.org/10.1109/JSTARS.2022.3152127.

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? Int. J. Appl. Earth Obs. Geoinf. 19, 205–213. http://dx.doi.org/10.1016/j.jag.2012.05.003, URL: https://www.sciencedirect.com/science/article/pii/S0303243412000992.

Huang, C., Chen, Y., Wu, J., 2014. Mapping spatio-temporal flood inundation dynamics at large river basin scale using time-series flow data and MODIS imagery. Int. J. Appl. Earth Obs. Geoinf. 26, 350–362. http://dx.doi.org/10.1016/j.jag.2013.09.002, URL: https://www.sciencedirect.com/science/article/pii/S0303243413001013.

Kleppmann, M., 2017. Designing Data-Intensive Applications. O'Reilly.

Li, M., Stefanakis, E., 2020. Geospatial operations of discrete global grid systems—A comparison with traditional GIS. J. Geovisualization Spat. Anal. 4, 26. http://dx.doi.org/10.1007/s41651-020-00066-3.

Mandery, N., Contributors, 2022. H3ron. https://github.com/nmandery/h3ron.

Mandery, N., Contributors, 2023. Ukis_h3cellstore. http://dx.doi.org/10.5281/zenodo.7534213, https://github.com/dlr-eoc/ukis-h3cellstore.

Martinis, S., Groth, S., Wieland, M., Knopp, L., Rättich, M., 2022. Towards a global seasonal and permanent reference water product from sentinel-1/2 data for improved flood mapping. Remote Sens. Environ. 278, 113077. http://dx.doi.org/10.1016/j.rse.2022.113077.

Meta Platforms, Inc., 2022. Zstandard compression algorithm. https://facebook.github.io/zstd/.

NASA JPL, 2013. NASA shuttle radar topography mission water body data shapefiles & raster files [dataset]. NASA EOSDIS land processes DAAC. http://dx.doi.org/10.5067/MEaSUREs/SRTM/SRTMSWBD.003.

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, 418–422. http://dx.doi.org/10.1016/j.rse.2022.113077.

Purss, M.B.J., Peterson, P.R., Strobl, P., Dow, C., Sabeur, Z.A., Gibb, R.G., Ben, J., 2019. Datacubes: A discrete global grid systems perspective. Cartogr. Int. J. Geogr. Inf. Geovisualization 54, 63–71. http://dx.doi.org/10.3138/cart.54.1.2018-0017.

Sahr, K., 2011. Hexagonal discrete global grid systems for geospatial computing. Arch. Photogramm. Cartogr. Remote Sens. 22, URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84888220662&partnerID=40&md5=57ec30f27dc88c88514d93395cfbf5c8.

Salamon, P., Mctlormick, N., Reimer, C., Clarke, T., Bauer-Marschallinger, B., Wagner, W., Martinis, S., Chow, C., Böhnke, C., Matgen, P., Chini, M., Hostache, R., Molini, L., Fiori, E., Walli, A., 2021. The new, systematic global flood monitoring product of the Copernicus emergency management service. In: 2021 IEEE International Geoscience and Remote Sensing Symposium. IGARSS, pp. 1053–1056. http://dx.doi.org/10.1109/IGARSS47720.2021.9554214.

Schlaffer, S., Matgen, P., Hollaus, M., Wagner, W., 2015. Flood detection from multi-temporal SAR data using harmonic analysis and change detection. Int. J. Appl. Earth Obs. Geoinf. 38, 15–24. http://dx.doi.org/10.1016/j.jag.2014.12.001, URL: https://www.sciencedirect.com/science/article/pii/S0303243414002645.

Schneibel, A., Merkle, N., Wieland, M., Lechner, K., Azimi, S., Henkel, F., Henry, C., Kiefl, R., Yuan, X., Gähler, M., 2022. User driven flood response: And monitoring information - key findings of the Data4Human project. In: IEEE Global Humanitarian Technology Conference, GHTC Technology for the Benefit of Humanity.

Wieland, M., Fichtner, F., Martinis, S., Krullikowski, C., Plank, S., Motagh, M., in review. S1S2-Water: A global dataset for semantic segmentation of water bodies from Sentinel-1 and Sentinel-2 data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens..

Wieland, M., Li, Y., Martinis, S., 2019. Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network. Remote Sens. Environ. 230, 111203. http://dx.doi.org/10.1016/j.rse.2019.05.022, URL: https://www.sciencedirect.com/science/article/pii/S0034425719302159.

Wieland, M., Martinis, S., 2019. A modular processing chain for automated flood monitoring from multi-spectral satellite data. Remote Sens. 11, http://dx.doi.org/10.3390/rs11192330, URL: https://www.mdpi.com/2072-4292/11/19/2330.