

UNIGIS

Master Thesis

submitted within the UNIGIS MSc programme
Interfaculty Department of Geoinformatics - Z_GIS
University of Salzburg

Surface water mapping to support humanitarian action A South Sudan case study

by

Francesco Collivignarelli
106720

A thesis submitted in partial fulfilment of the requirements of
the degree of
Master of Science (Geographical Information Science & Systems) – MSc (GISc)

Advisor:

Dr. Lorenz Wendt

Salzburg, 9.2.2022

Science pledge

By my signature below, I certify that my thesis is entirely the result of my own work. I have cited all sources I have used in my thesis, and I have always indicated their origin.

Luino, 9.2.2022

Francesco Collivignarelli

A handwritten signature in black ink, reading "F. Collivignarelli". The signature is written in a cursive style with a large initial 'F'.

Acknowledgments

I would like to acknowledge the Paris Lodron University of Salzburg for the comprehensive Geographic Information Science and Systems Master of Science program. The expertise I have gained in these studies is certainly invaluable. My sincere thanks to Dr. Lorenz Wendt and Dr. Andreas Braun, who guided me through the present research. Your commitment to supervision was extremely helpful. Finally, a special mention to the MSF GI staff, Edith Rogenhofer and Léo Tremblay and to Léo Martine from UNHCR. The activities in which you engage every day are commendable and your insights have allowed me to make a small contribution in support of it.

Table of Contents

| | |
|---|-----------|
| Abstract | 1 |
| Introduction | 3 |
| <i>Context</i> | 3 |
| <i>EO for humanitarian action</i> | 4 |
| <i>Problem statement</i> | 6 |
| Aims and objectives | 9 |
| <i>Aims</i> | 9 |
| <i>Objectives</i> | 9 |
| Material and methods | 10 |
| <i>Material</i> | 10 |
| Study area..... | 10 |
| Sentinel-1 | 12 |
| Sentinel-2..... | 15 |
| WorldPop | 17 |
| OpenStreetMap..... | 21 |
| ESRI land cover | 23 |
| Google Earth Engine..... | 24 |
| <i>Methods</i> | 26 |
| SAR data preparation..... | 27 |
| Selection..... | 27 |
| Radiometric slope correction | 28 |
| Speckle filtering | 30 |
| Surface water extraction | 32 |
| Ordinary adaptive thresholding..... | 32 |
| Adaptive-constrained thresholding | 34 |
| Bimodality test | 37 |
| Cross-validation | 40 |
| Population data intersection..... | 45 |
| VGI intersection..... | 46 |
| Results | 49 |
| <i>Surface water and population</i> | 49 |
| <i>Water occurrence</i> | 57 |
| <i>Applications</i> | 60 |
| Single date | 60 |

| | |
|--|-------------|
| Multitemporal..... | 61 |
| Discussion..... | 62 |
| Conclusions and outlook | 63 |
| References..... | 64 |
| Annex 1 | I |
| Annex 2 | XVI |
| <i>Single date application</i> | <i>XVI</i> |
| <i>Multitemporal application.....</i> | <i>XX</i> |
| <i>Utility and processing functions.....</i> | <i>XXII</i> |

Abstract

The use of Geo Information (GI) and Earth Observation (EO) in the context of crisis and emergency mapping has experienced a rapid growth in recent years (Lang et al., 2017). In parallel, the uptake of geospatial technologies within the humanitarian sector happened at an unprecedented rate (Lang et al., 2018) as consequence of the urgent information need in challenging situations involving a considerable number of people.

Particularly, the use of remote sensing contributed to achieve promising results in supporting refugee camps mapping and monitoring (Wendt et al., 2017) as well as in the exploitation of their environmental resources optimizing water supply (Wendt et al., 2015).

To increase the effectiveness of humanitarian assistance, several solutions have been proposed and realized, demonstrating that information services can be made operational for Non-Governmental Organizations (NGOs) such as Médecins Sans Frontières (MSF) (Lang et al., 2019).

In the surface water mapping domain, some visualization platforms are currently available. Global Surface Water Explorer provides global coverage, although the contained data is updated until 2020 and the methodology (Pekel et al., 2016) relies exclusively on optical imagery. Therefore, both of these characteristics hamper the timely relevant disaster response in cloudy condition (Braun and Hochschild, 2017), especially in tropical areas and during the rainy seasons where and when flooding is more likely to occur. Commercial solutions, such as ICEYE (ICEYE, 2021) are not affordable as the imagery is provided at high cost and acquisitions need to be planned in advance. In addition, the data exploited is more suitable to target small and urban areas due to the limited swath and the short-wavelength sensing band (Ignatenko et al., 2020).

The present study aims to build a robust surface water mapping solution, to support disaster response and risk analysis (Plate, 2002) intersecting data from different sources such as satellite imagery, Volunteered Geographic Information (VGI) and demographic maps with no to low human interaction. The resulting information is eventually integrated into the context of humanitarian action (Lang et al., 2019) and delivered to organizations operating in problematic areas, primarily South Sudan.

Mapping of surface water with Synthetic Aperture Radar (SAR) is well established, and allows to operate irrespective from weather and hour of the day. However, global approaches (Markert et al., 2020, Cao et al., 2019, Uddin et al., 2019) struggle with the conditions found in South Sudan, particularly in terms of land cover types and their shares. Usual approaches aim to find an optimal threshold for backscatter to distinguish water from land, starting from the ordinary intensity thresholding theory (Otsu, 1979).

In this research an alternative adaptive-constrained thresholding method is explored, making exclusive use of free of charge data from the Copernicus program (Geudtner et al., 2014), and auxiliary layers. The method is tested using acquisitions from the years 2016 to 2021 over Pibor, South Sudan, and validated against optical data showing an overall accuracy of more than 88%.

The resulting water maps are intersected with demography layers (Tatem, 2017), to estimate the amount of population living in flooded areas in current and past dates. The results consist therefore of a multi-temporal set of surface water maps, one each

suitable satellite acquisition, to which corresponds a set of involved population maps. In addition, the possible intersection with VGI data (Ramm et al., 2011) such as roads and places is demonstrated. The surface water multi-temporal dataset is eventually condensed in a water occurrence map, which can be used for flood risk assessment.

The data processing and information generation pipeline is designed making extensive use of JavaScript applications for Google Earth Engine (Gorelick et al., 2017), in order to enable the application of the methodology with limited effort on large scales and replicable on different areas beyond those targeted here.

Introduction

This chapter introduces the research area of the present study and is divided into three parts. In the first part, Context, the larger frame of information need in the humanitarian sector is discussed. In the second, EO for humanitarian action, the role of Earth Observation (EO) in the theme is introduced. The final part, Problem statement, includes a literature review and explains the challenges which this research aims to tackle, with main focus on the subject of surface water mapping in South Sudan.

Context

Large-scale population displacements within complicated emergency circumstances caused by armed conflicts and protracted crises have heightened the need for more efficient humanitarian aid in recent years, particularly in the Middle East and Eastern Africa (Lang et al., 2019).

Humanitarian organizations need up-to-date, trustworthy information regarding the situation on-site and the environmental context when reaching out displaced people (Lang et al., 2017). The requirement for accurate data concerns all stages of humanitarian assistance, from mission planning in an emergency to long-term care and maintenance, including the availability of natural resources such as ground and surface water and estimating the immediate environment's nutritional capacity.

In crisis situations, acquiring information on-site through gathering evidence on the ground is frequently limited, with time and accessibility being the most limiting considerations (Lang et al., 2015). Satellite image mapping and other geospatial technologies are used to provide situational awareness in complex humanitarian disasters (Voigt et al., 2014). In the humanitarian sector, optimizing information flows is a major asset, and technological innovation is critical, as it is in many other domains, with the value of any new technology being severely tested according to whether it eventually helps to save lives or not. In the humanitarian field, GI and EO technologies are popular since two decades ago (Bjorgo, 2001). However, converting this data into useful geographic information products for humanitarian aid workers is still difficult (Füreder et al., 2015).

In this regard, humanitarian relief organizations play a significant user role; they are the final stakeholders, and they work closely with those with technology providers. In recent years, several programs and initiatives have been developed in which academic institutions and humanitarian actors have exchanged knowledge. GI and EO has emerged as one of the most important enabling technologies in humanitarian efforts. Many different types of optical and radar sensors, as well as remotely piloted aircraft systems have begun to give a wide range of spatial, temporal, and spectral resolutions.

Humanitarian aid is required to guarantee that humanitarian action is targeted to the appropriate people at the right place at the right time in all mandates and missions. If applied correctly, new tools and techniques may be beneficial. Non-Governmental Organizations (NGOs) such as Médecins Sans Frontières (MSF) rely on up-to-date information to plan, direct, and, on occasion, adjust their operations and logistics on the ground.

EO for humanitarian action

Earth observation (EO) is a broad term that refers to all conventional observational instruments as well as prediction and forecasting models that integrate and use these measurements. The Group on Earth Observation (GEO) (Group on Earth Observation, 2005) which is made up of governments and international organizations, aims to bring together a variety of observation systems into the Global Earth Observation System of Systems (GEOSS) for monitoring the whole planet to address global key challenges, such as natural disasters.

One of the major technological contributors to GEOSS is EO from space using satellite remote sensing, which includes the space infrastructure (satellite sensors) and data provisioning and archiving subsystems, as well as the image analysis and information extraction subsystem, which includes information delivery and interoperable usage (Li, 2010).

In comparison to traditional field mapping, humanitarian response benefits from the broad advantages of remotely sensed data. The characteristics of remote sensing data are summarized with its potential humanitarian intervention in Table 1.

EO-based technology has evolved into a decision-supporting tool for humanitarian professionals when combined with improved information extraction techniques (Spröhnle et al., 2014). Remotely sensed data and associated information products complement time and effort-intensive field-based surveys and thereby increase the pool of spatially aware technologies for humanitarian assistance support due to their observational power, pervasive usage, and availability.

Table 1. Assets in remote sensing data use and potential within humanitarian action. Source: (Lang et al., 2015).

| Asset | Potential for humanitarian action |
|---|---|
| From a distance | No direct access required |
| | Objective source of information |
| | Cost-effective compared to field data collection |
| Area-wide coverage | Same conditions over large areas |
| | Time-efficient data collection |
| Level of detail | Control of level of detail |
| | Overview vs. detailed information |
| Global availability | Comparable data |
| | Fast and easy access to data |
| Time series monitoring (including retrospectively) | Objective source of documentation on how a situation looked before a critical situation emerged |
| | Validate other sources of information (e.g., from field data) |
| | Ex post evidence on past situations, events, changes |

Remote sensing takes data over broad areas in similar conditions from a fixed orbital view. As a result, it has an objective component, which aims to disclose the situation as it is, undistorted, unmanipulated, and potentially available to everybody at no cost. Regular capture of data allows monitoring of constant or changing ground conditions

using standardized methodologies, in addition to on-demand assessments of the situation on the ground, mainly using commercial data.

Remote sensing data archiving methods offer retroactive perspectives, allowing specific events to be analyzed historically rather than *ex post*, such as the growth or dismantlement of a camp. New technologies, such as the Sentinel fleet, which was installed by the European Copernicus programme (ESA, 1998), provide high-frequency global overviews.

While several studies use effectively optical EO satellite data and equivalent passive remote sensing techniques (Wendt et al., 2015, Wendt et al., 2017), active remote sensing, particularly radar, has a great potential in the field of humanitarian assistance because of its weather and daylight independence (Braun et al., 2016).

Quality tests are required for remote sensing products. In order to assess the information content's accuracy and dependability, it is frequently compared to ground reference data (Congalton and Green, 2019). Ground records are typically limited and difficult to obtain in humanitarian crisis circumstances, and in-field assessments are rarely conducted in parallel or within a similar time frame (Checchi et al., 2013). To compensate for the lack of ground reference data, independent evaluation such as image interpretation or cross-validations might be performed using optical data.

Concerning operational solutions, the Austrian Space Programme (ASP) funded project EO4HumEn (2013–2016) first established an EO-based information service to support NGOs, particularly MSF, humanitarian activities. It included services about population monitoring, groundwater drilling, and environmental impact assessment of disasters.

EO4HumEn+ follow-up project (2016–2018) aimed to broaden the service offering additional tools. While data on displaced people remained the most important part, the products on camps concentration have been expanded to include metropolitan contexts.

As outcome of these projects, currently, the department of Geoinformatics (Z_GIS) of the Paris Lodron University of Salzburg, still carries out dedicated services for MSF in the framework of GEOHUM (University of Salzburg - Z_GIS, 2021).

In conclusion, in each phase of humanitarian action, EO-based information is crucial. Satellite data provide quick information on the condition on the ground during an emergency and beyond. This can help saving time and money, for instance on staff and logistics. In addition, remote sensing provides a good source of information for cross validation, as different types of sources (Optical, SAR) are available and use different acquisition techniques concerning spectral bands and configurations.

Problem statement

As introduced above, NGOs and especially MSF need time-relevant and reliable information about several aspects concerning the ground status, in order to plan their operations as best.

Flooding has recently become a particularly hot topic especially for South Sudan, with thousands of people displaced and huge problems in organizing the necessary aid (MSF, 2019, MSF, 2020, MSF, 2021, MSF, 2022).



Figure 1. Floods in Pibor, South Sudan, 2019. Camp of displaced people is visible on the right. Source: MSF.

In addition, flooding is a general challenge for logistics. Some areas in eastern South Sudan, due to their geographic setting, have limited supply routes, with only two roads operational becoming inaccessible during the rainy season. As result, it is often necessary to fly towards these areas for nine to ten months of the year, approximately from May to February. This is reportedly expensive and difficult to arrange.



Figure 2. A World Food Program (WFP) helicopter on the flooded airstrip of Pibor, South Sudan, 2019. Source: REUTERS/Andreea Campeanu.

In the Pibor area, flooding can be so severe that the airstrip becomes inaccessible, and flights have to be diverted to Gumuruk, from where Pibor town has to be reached by car or at worst by boat, increasing the travel time required as well as the transportation costs, and decreasing safety.



Figure 3. People use canoes to cross water in Pibor, South Sudan, 2019. Source: REUTERS/Andreea Campeanu.

In this type of situation, it seems clear that EO technologies can help the generation of information in support of emergency planning and management. Particularly, surface water mapping at acceptable resolution can contribute in this sense, both for current and past dates, as archive data can highlight the most flood prone locations.

Using remote sensing, with particular reference to high-resolution and free of charge data, various approaches have been proposed exploiting optical (Acharya et al., 2016, Du et al., 2016, Feyisa et al., 2014, Main-Knorn et al., 2017, Verpoorter et al., 2012), SAR (Bioresita et al., 2018, Gulácsi and Kovács, 2020, Liang and Liu, 2020, Markert et al., 2020) or both techniques (Hong et al., 2015).

In the framework of SERVIR-Mekong (Asian Disaster Preparedness Center - NASA JPL, 2022b) the methodologies set for flood monitoring are possibly applicable as relying on SAR data and the relevant tools developed satisfy the platform requirements aimed by this study (Markert et al., 2020). However, these methodologies rely on specific conditions found in the Mekong delta region, specifically on large water bodies or well-delineated features from which edges can be extracted with a satisfactory reliability.

Other methodologies using SAR combine change detection and thresholding (Long et al., 2014, United Nations, 2020) involving the selection of a reference (pre-flood) image, which is time consuming and not straightforward in the case of a fast response is needed. In addition, when change detection is involved, permanent water and slow, persistent flooding is likely to be lost in the analysis, resulting in an inaccurate detection of the total surface water extent at a given time.

Artificial intelligence techniques, such as neural networks, or unsupervised classification methods (Gobeyn et al., 2017, Huang et al., 2018, Li et al., 2019) do not offer satisfactory control or involve modelling phases prior the actual detection. In addition, the computation is often time consuming and models are specifically tuned for a certain and limited area of interest.

For what concerns historical aggregated data, GSWE (EC - JRC, 2020) features global coverage. However, an analysis over South Sudan performed in the section Water occurrence of the present study, highlights its underestimation. The reason of the underestimation has to be found in the methodology itself (Pekel et al., 2016) which exploits optical data only and at the resolution of 30 m, likely to miss small features such as streams or small rivers.

SERVIR-Mekong offers an historical flood analysis platform, part of the decision support tools (Asian Disaster Preparedness Center - NASA JPL, 2022a), which is largely based on GSWE data and can't be used in South Sudan as the coverage is limited to the Mekong river basin.

Iceye (ICEYE, 2021) provides comprehensive flood hazard solutions, both concerning monitoring of single events and historical analysis, using SAR X-band data. The commercial nature of the platform, combined with the high-frequency band used, hampers its applicability in the humanitarian context over large areas due to the cost and technical reasons, as X-band has a limited strength over specific land cover types, such as flooded vegetation.

In summary, no of the mentioned methodologies and solutions seem to fulfill the requirements needed for effective humanitarian action. All of them, in fact, rely on demand acquisitions that reduce timeliness and increase costs or on optical data which hamper the operability irrespective from the weather present at the time of need. When free of charge, systematically acquired data is used, the resolution is often not satisfactory. In addition, all the existing studies are tuned for specific areas of interest, where no of them feature the same environmental conditions present in South Sudan.

For this reasons, the present study aims to fill the existing gap, by researching a methodology which beside satisfying the requirements above, can contribute with a ready-to-use tools which can be utilized by final stakeholders, primarily MSF. Final goals are further detailed in the next chapter, Aims and objectives.

Aims and objectives

Aims

The aim of this research is setting up a methodology to produce surface water maps which can be used for humanitarian purposes, particularly in South Sudan.

The methodology should be applicable over large areas, using free data, with no or limited human intervention, in a short time frame and allow the possibility to be operational when needed irrespective from the local conditions (weather, hour, status of roads and infrastructures).

In addition, surface water maps from past dates must be produced in order to fulfill a twofold purpose: first, to support the delineation of areas featuring the major flood risk. Second, to help to assess the robustness of the methodology that can be applied under different environmental and seasonal conditions, expected to cyclically vary over time.

The results, in terms of maps and statistics, are used for fast humanitarian action, with focus on the estimation of the population living in flooded areas.

Therefore, the approach adopted is realized by means of the objectives described below.

Objectives

To achieve the research aims, covering all the requirements and overcoming the challenges of the outlined scenario the objectives described hereafter are identified.

- Produce surface water maps from current and past dates. This main objective enables all the others and must be considered the most important milestone of the present research. Given the constraints set by the aims, satellite remote sensed data are the most suitable to be used.
- Enhance the surface water maps by means of intersection with other data sources. The surface water maps are spatial layers, delineating the areas covered by open water at a given time. These layers can be intersected with other data sources such as population, infrastructure, or places. In this research major focus is given to population.
- Generate information in support of flood risk maps, to be used to improve existing solutions. This objective makes use of historical surface water maps, produced by the application of the methodology on archive data.
- Develop tools which can be used to extract the desired information with no or limited human intervention. For this objective, JavaScript applications are developed respecting high quality coding standards to be used by the stakeholder's staff and improved by other researchers in future.
- Test the approach over hot spots agreed with the stakeholders. The full methodology is applied over the Pibor administrative area level 3, Greater Upper Nile region.

Material and methods

In this chapter, material and methods are discussed in detail. The first section, material, contains a description of the study area, of the datasets and of the tools used to implement the methodology presented in the second section, methods. In this last part, a deep level of description is used for each individual step included in the methodology.

Material

To implement a robust surface water mapping solution, this study exploits remote sensing (SAR), demographic data, land covers and VGI. All these sources are merged and harmonized by using a popular geospatial processing and visualization tool. In the following section, the study area is first introduced. Subsequently, Sentinel-1 (SAR), Sentinel-2 (optical), World Pop (demographic data), OpenStreetMap, ESRI world cover (land cover) and Google Earth Engine are described.

Study area

The study area of this research coincides with the level 3 administrative unit of Pibor, South Sudan. Its location within Africa and South Sudan is shown by the map layout in Figure 4.

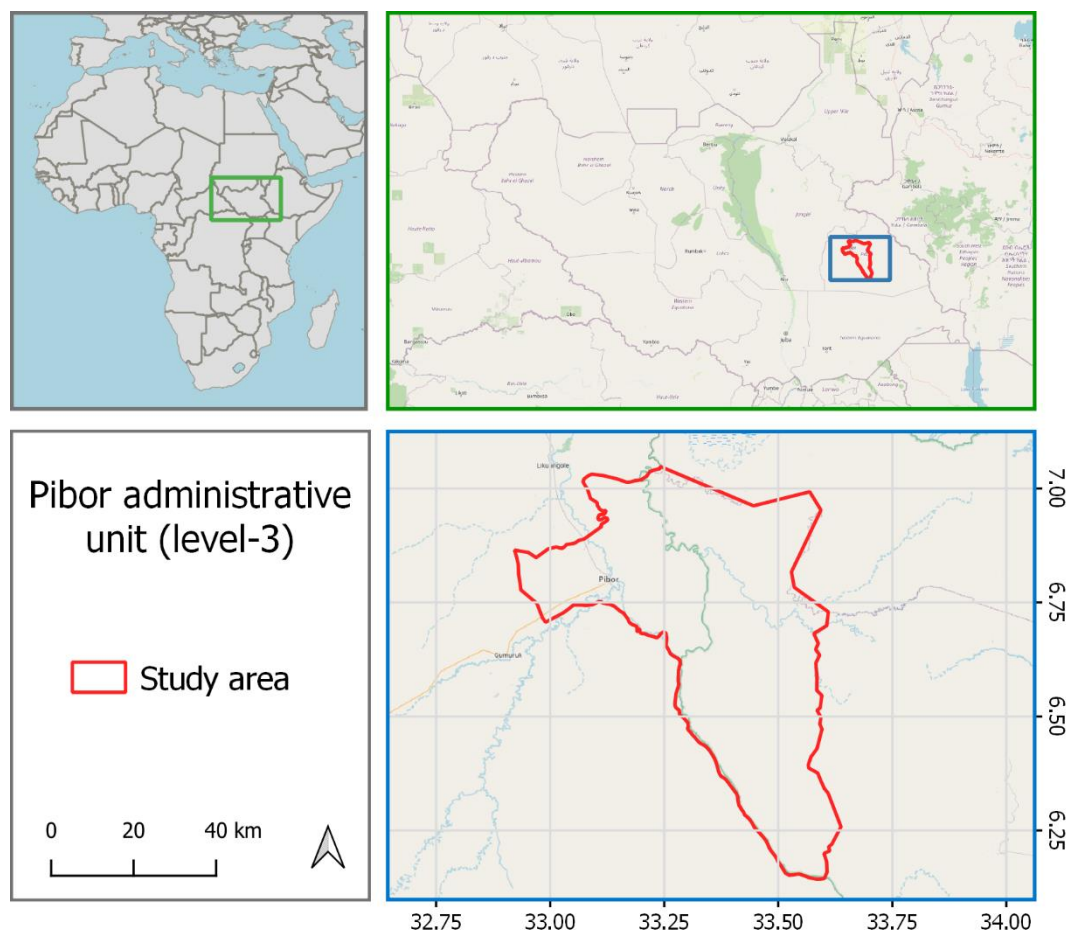


Figure 4. Map layout showing the location of the study area. Background: © OpenStreetMap contributors.

The study area, having a total extent of 3612 km², has its main center in Pibor Post (or Pibor, or Pibor town) that is a colonial-era outpost built in 1912 by the British and originally named Fort Bruce (Thorburn, 1922). The town serves as the county seat of

Pibor County, one of the constituent counties of Jonglei State prior to the creation of new states in 2015. The total population of the study area is uncertain, and official figures do not exist at the granularity of administrative level 3. Recent assessments for the relevant administrative level 2 is 911416 (HUMDATA, 2022), while the best estimate for the study area is 20781 people (University of Southampton, 2013).

The Pibor River, formed by the confluence of several smaller streams, begins its journey in the area. The river flows northward, receiving the Akobo river near Akobo. Eventually, after receiving the Gilo river and the Bela river, it joins the Baro river to form the Sobat river. This hydrologic system is cause of frequent flooding.

The topography of the area is almost flat, with limited variation in altitude. Elevation with respect to the sea level is slightly above 400 meters, with a range between 410 and 482 meters (CGIAR, 2018).

The climate of Pibor is classified as *Aw* (tropical savanna climate or tropical wet and dry climate) according to the last update of the Köppen-Geiger classification (Peel et al., 2007).

This climate is characterized by two seasons, wet and dry. The first involves intense rainfall that lasts for many days, with frequent thunderstorms and rare days without phenomena, similarly to the tropical climate. The second is very dry, with almost no rain.

The rains are distributed in a characteristic way in correspondence of the passage of the sun at the zenith and therefore are called zenithal rains: the periods of maximum intensity of rainfall are joined at the equator. These effects are clearly visible in the graph in Figure 5, which shows the average mm of rain per month over Pibor for the years from 2015 to 2020. The data is estimated using Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) (Funk et al., 2015).

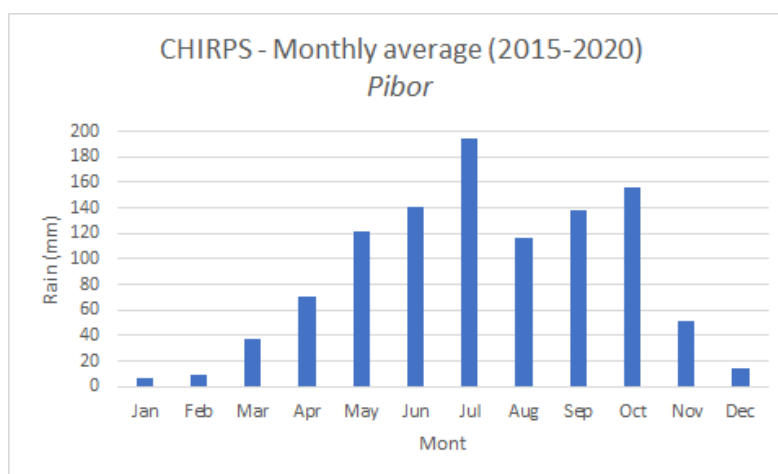


Figure 5. Pibor. Monthly precipitation average (CHIRPS) for the years 2015-2020.

In the *Aw* zone, temperatures are always quite high (generally higher than 21-22 °C). However, the temperature ranges are quite great: during the day high temperatures (up to 40 °C) are reached while during the night they can fall more (until 17 °C). This climate favors the growth of a dense and lush grass.

According to the ESRI land cover classification, visible below in Figure 18 (ESRI, 2020), the study area features the land cover classes summarized in Table 2.

Table 2. Pibor. Land cover type and relevant extents.

| Land Cover | Extent [Km ²] |
|--------------------|---------------------------|
| Water | 5.0 |
| Trees | 2499.0 |
| Grass | 246.5 |
| Flooded vegetation | 60.8 |
| Scrub/shrub | 797.8 |
| Built Area | 2.6 |
| Bare ground | 0.4 |

These land cover types and extent are, however, to be considered only an indication as the complexity of the environment and climate can change their shares with seasonal trend. This is particularly valid for what concerns water, object of the present study. Its extent, in fact, can vary significantly over the year due to the fact that flood prone zones are particularly widespread in the area.

Sentinel-1

The Sentinel-1 mission (European Space Agency (ESA), 2014), is the European radar observatory for the Copernicus joint initiative of the European Commission (EC) and ESA. Copernicus is a European program for the implementation of information services dealing with environment and security. It is based on observation data received from Earth Observation satellites and ground-based information.

The Sentinel-1 mission operates with a constellation of two satellites, carrying a C-band sensor acquiring in four exclusive imaging modes with different resolution (down to 5 m) and coverage (up to 400 km). It provides dual polarization capability, very short revisit times and rapid product delivery. For each observation, precise measurements of spacecraft position and attitude are available.

Synthetic Aperture Radar (SAR) has the advantage of operating at wavelengths not impeded by cloud cover or a lack of illumination and can acquire data over a site during day or nighttime under all weather conditions. Sentinel-1, with its C-SAR instrument, can offer reliable, repeated wide area monitoring.

The mission is composed of a constellation of two satellites, Sentinel-1A and Sentinel-1B, sharing the same orbital plane.

Sentinel-1 is designed to work in a pre-programmed, conflict-free operation mode, imaging all global landmasses, coastal zones and shipping routes at high resolution and covering the global ocean with frames of 20x20km. This ensures the reliability of service required by operational services and a consistent long term data archive built for applications based on long time series.

Sentinel-1 carries a single C-band synthetic aperture radar instrument operating at a center frequency of 5.405 GHz. It includes a right-looking active phased array antenna providing fast scanning in elevation and azimuth, data storage capacity of 1410 Gb and 520 Mbit/s X-band downlink capacity.

The C-SAR instrument supports operation in dual polarization (HH+HV, VV+VH) implemented through one transmit chain (switchable to H or V) and two parallel receive chains for H and V polarization. Dual polarization data is useful for land cover classification and sea-ice applications.

Sentinel-1 operates in four acquisition modes as depicted in Figure 6:

- Strip map (SM)
- Interferometric Wide swath (IW)
- Extra-Wide swath (EW)
- Wave mode (WV).

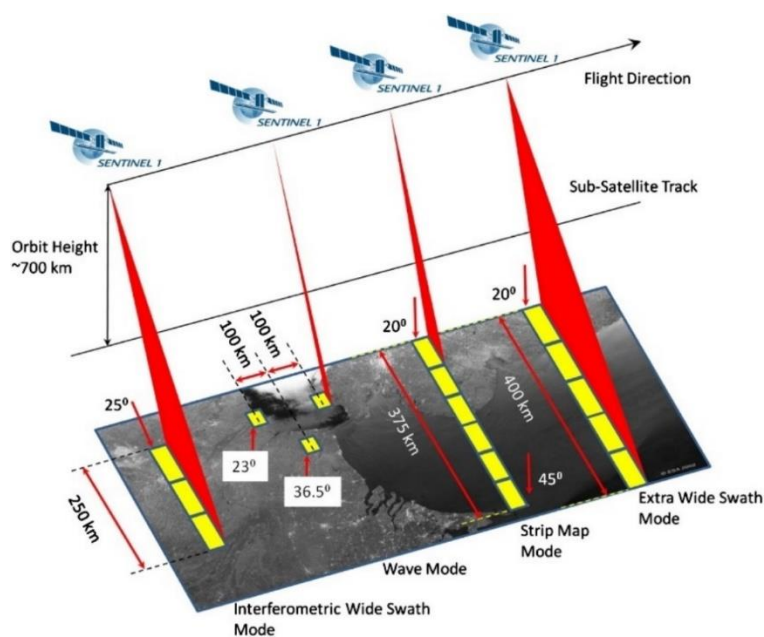


Figure 6. Sentinel-1 acquisition modes. Source: ESA.

IW mode allows combining a large swath width (250 km) with a moderate geometric resolution (5 m by 20 m). The IW mode images three sub-swaths using Terrain Observation with Progressive Scans SAR (TOPSAR). With the TOPSAR technique, in addition to steering the beam in range as in Scanning Synthetic Aperture Radar (SCANSAR), the beam is also electronically steered in azimuth to reach a higher quality image. Interferometry is ensured by sufficient overlap of the single bursts. The TOPSAR technique ensures homogeneous image quality throughout the swath.

The IW mode at 10m resolution is the default acquisition mode over land and the one used in the present study. This configuration of the background mission guarantees global coverage and a repetition frequency of at least 12 days for the same viewing geometry, which becomes of 4 days in the case more than one is available. A graphical representation of the observation scenario in terms of revisiting time and mode is visible in Figure 7 and Figure 8.

Sentinel-1 Constellation Observation Scenario: Revisit & Coverage Frequency

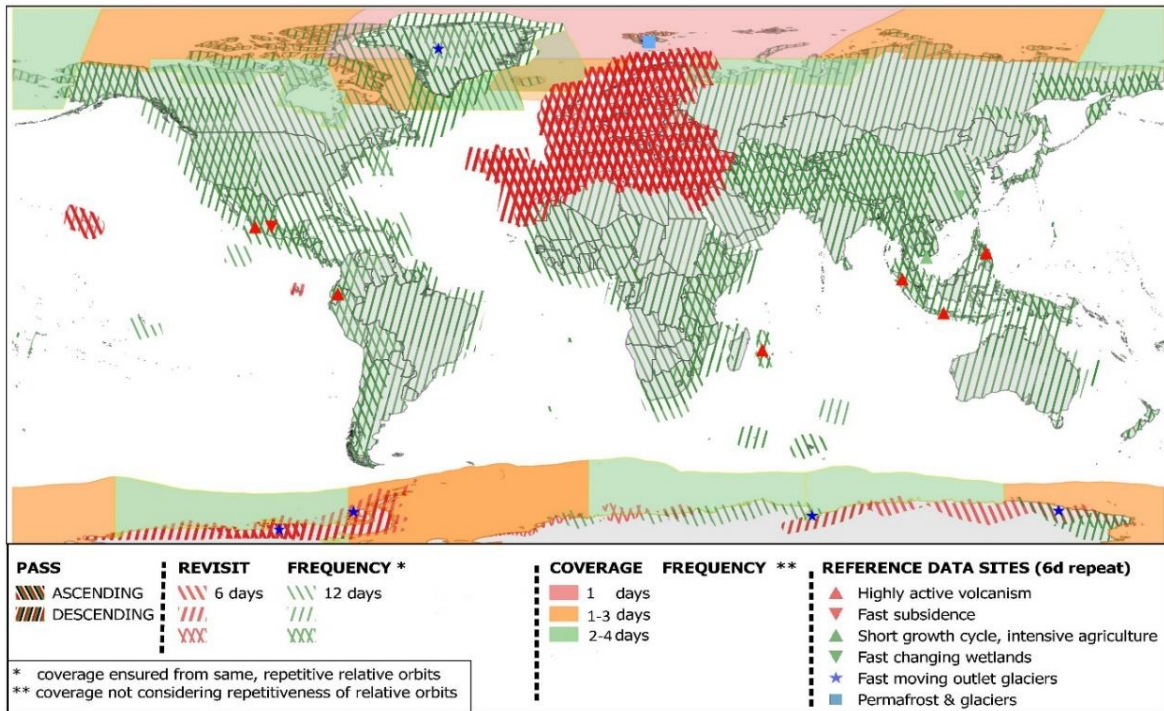


Figure 7. Sentinel-1 observation scenario: revisit & coverage frequency, 2-4 days repetition time are ensured on land, except for the arctics. Source: ESA.

Sentinel-1 Constellation Observation Scenario: Mode - Polarisation - Observation Geometry

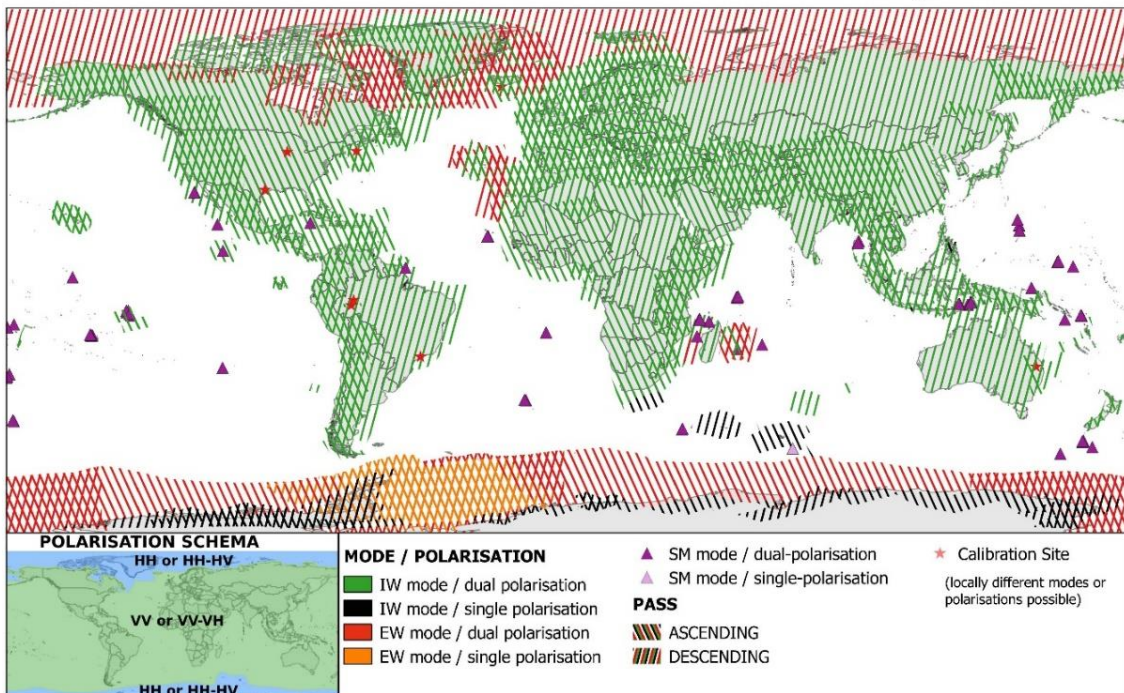


Figure 8. Sentinel-1 observation scenario: mode - polarization – acquisition geometry. Source: ESA.

In this research and more specifically, the VV channel taken from each acquisition date is exploited for the derivation of a surface water map. In fact, as visible in the example in Figure 9, derived processing Ground Range Detected (GRD) products, this polarization is strongly suitable for the purpose. In the reported SAR image that refers to 30.09.2020, water features appear dark while all other land cover types are characterized by a brighter gray tone, down to white for built up and vegetated areas.

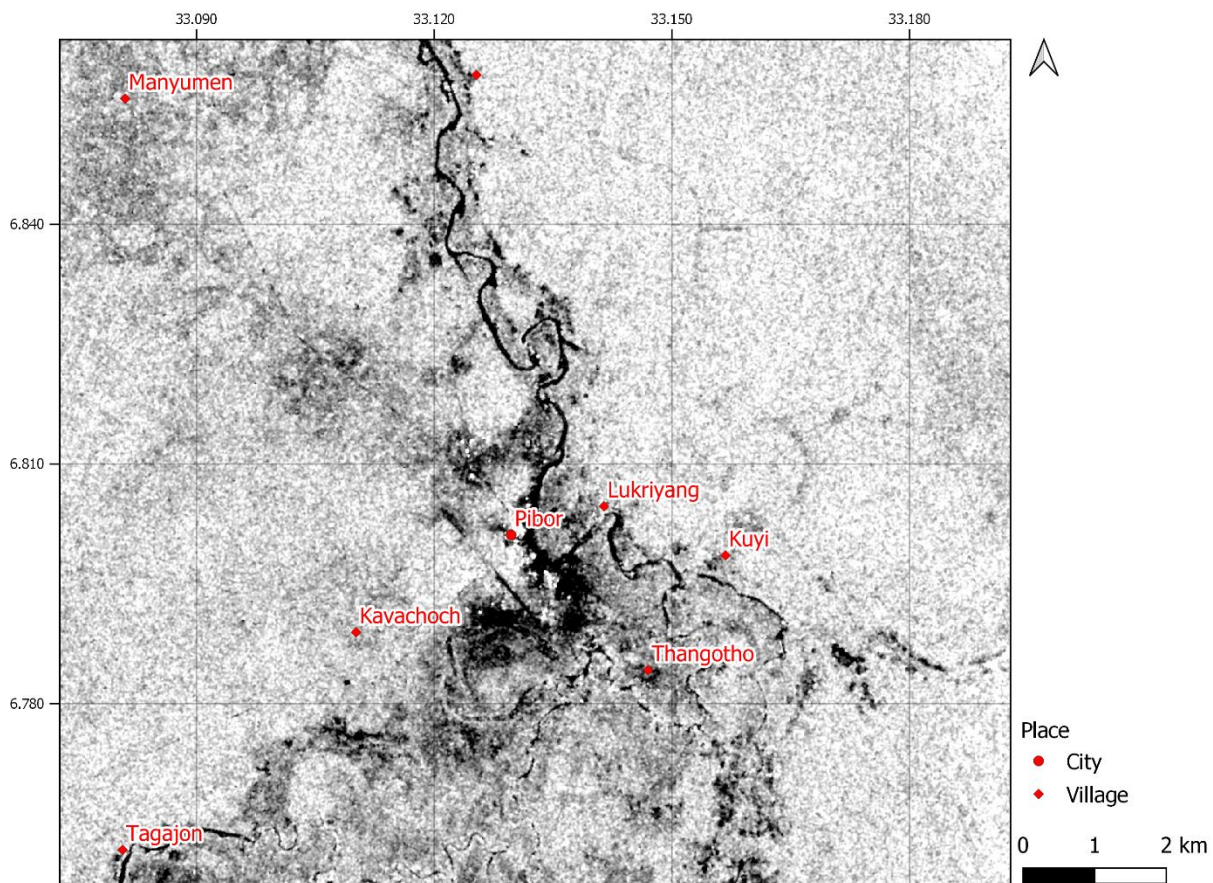


Figure 9. Region around Pibor, South Sudan. Sentinel-1 VV acquisition for 30.09.2020: dark features correspond to open water. Contains modified Copernicus Sentinel data [2020]. Places: © OpenStreetMap contributors.

Sentinel-2

The Copernicus Sentinel-2 mission comprises a constellation of two polar-orbiting satellites placed in the same sun-synchronous orbit, phased at 180° to each other.

Sentinel-2 aims at monitoring variability in land surface conditions, and its wide swath width (290 km) and high revisit time (10 days at the equator with one satellite, and 5 days with 2 satellites under cloud-free conditions which results in 2-3 days at mid-latitudes) will support monitoring of Earth's surface changes.

The payload instruments allow to acquire images in the spectral bands reported in Figure 10.

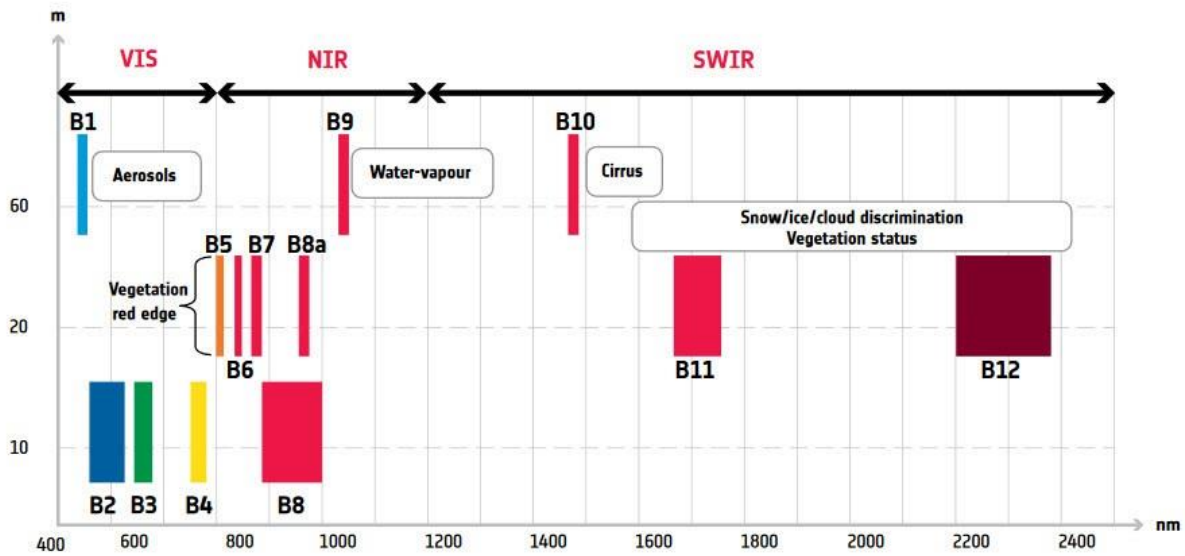


Figure 10. Sentinel-2 spectral bands. Horizontal axis shows the relevant wavelegths, vertical axis the corresponding spatial resolutions. Source: ESA.

The use of Sentinel-2 in this research is marginal. In fact, due to the expected massive cloud cover at the time of need, optical data often become unusable. For this reason, the methodology later discussed avoids their use.

However, Sentinel-2 is exploited as reference data, as for a selected date cross-validation points are extracted from an index calculated out of two of the spectral bands available.

Figure 11 shows a close-up of a Sentinel-2 optical and a Sentinel-1 SAR data acquired on the same date. As clearly visible, cloud cover hampers the correct capture of the Earth’s surface, preventing from any analysis. As microwaves can penetrate clouds, SAR can be effectively used in this case.

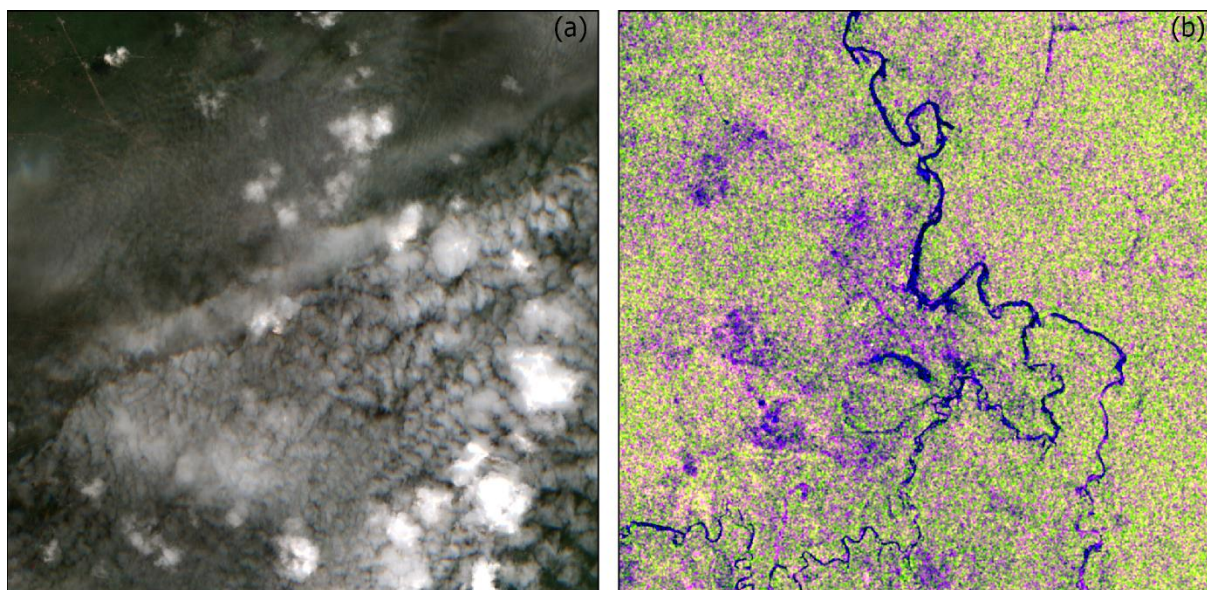


Figure 11. Pibor Post. Sentinel-2 true color composite (B4/B3/B2) (a) vs Sentinel-1 false color composite (VV/VH/VV-VH) (b). Data acquired on the same date, 07.08.2020. Contains modified Copernicus Sentinel data [2020].

WorldPop

WorldPop (University of Southampton, 2013) started in October 2013 to combine the AfriPop, AsiaPop and AmeriPop population mapping projects. It aims to provide an open access archive of spatial demography datasets for Central and South America, Africa and Asia to support development, disaster response and health applications.

The methods used (Alegana et al., 2015, Deville et al., 2014, Gaughan et al., 2013, Linard et al., 2012, Stevens et al., 2015, Tatem et al., 2007) are designed with full open access and operational application in mind, using transparent, documented and peer reviewed methods to produce easily updatable maps with accompanying metadata and measures of uncertainty.

Within the WorldPop project a range of open geospatial datasets are combined in a flexible regression tree framework to reallocate contemporary aggregated spatial population count data.

Statistical assessments suggest that the resultant maps are consistently more accurate than existing population map products, as well as the simple gridding of census data. Moreover, the 100m spatial resolution represents a finer mapping detail than has been produced at national extents, and the integration with household survey, microdata, satellite, and other data sources enables the production of more than simply population count estimates - age structures, births, pregnancies, poverty, and urban growth are all mapped, with further variables under production.

The data concerning population mapping, used in the present study, are generated using the methodological approaches described hereafter.

Land cover-based: through detailed mapping of settlements, and linkage of these settlement extents with gazetteer population numbers, most of the resident population can be mapped within settlements with good precision.

Mapping of the remaining minority rural populations follows different approaches. The settlement maps are used to refine land cover data, while local high resolution census data is exploited to identify typical regional per-land cover class population densities, which are then applied to redistribute census counts to map human population distributions.

This population mapping approach forms the basis of some older WorldPop datasets however is now being replaced by an alternative "Random Forest" mapping approach described below.

A visual comparison of WorldPop data (100m resolution) against Global Rural-Urban Mapping Project (GRUMP) (5 km) and Oak Ridge National Laboratory LandScan (1 km) is visible in Figure 12.

Random Forest: a new semi-automated dasymetric modeling approach has been built that incorporates census and a wide range of open access ancillary datasets in a flexible, "Random Forest" estimation technique (Stevens et al., 2015). A combination of widely available, remotely sensed, and geospatial datasets (e.g., settlement locations, settlement extents, land cover, roads, building maps, health facility locations, satellite nightlights, vegetation, topography, refugee camps) contribute to

the modeled dasymetric weights and then the Random Forest model is used to generate a gridded prediction of population density at ~100 m spatial resolution.

The prediction layer is then used as the weighting surface to perform dasymetric redistribution of census counts at a country level. The modelling process produces accompanying metadata for each output dataset, documenting input data sources and accuracy statistics. Outputs show marked improvements in mapping accuracies over the land cover-based approach outlined above and other population mapping approaches (Stevens et al., 2015).

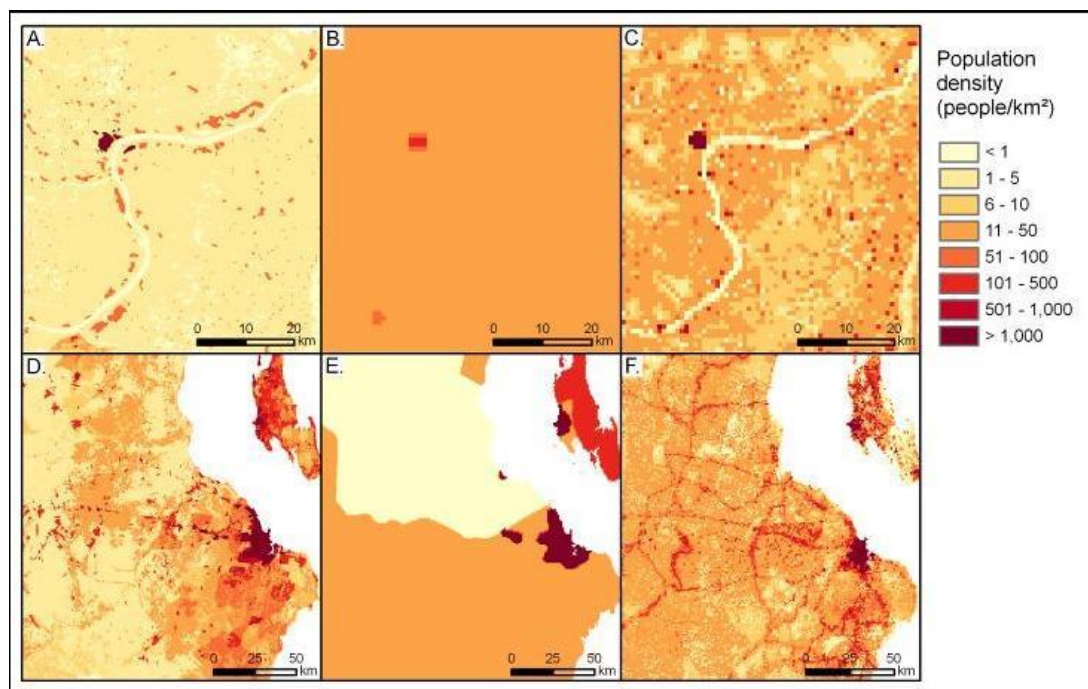


Figure 12. Visual comparison of WorldPop (A and D) with GRUMP (B and E) and LandScan (C and F) gridded population datasets. Upper line, North-East region of Guinea, along the Niger river. Lower line, region around Dar es Salaam, Tanzania (WorldPop, 2020b).

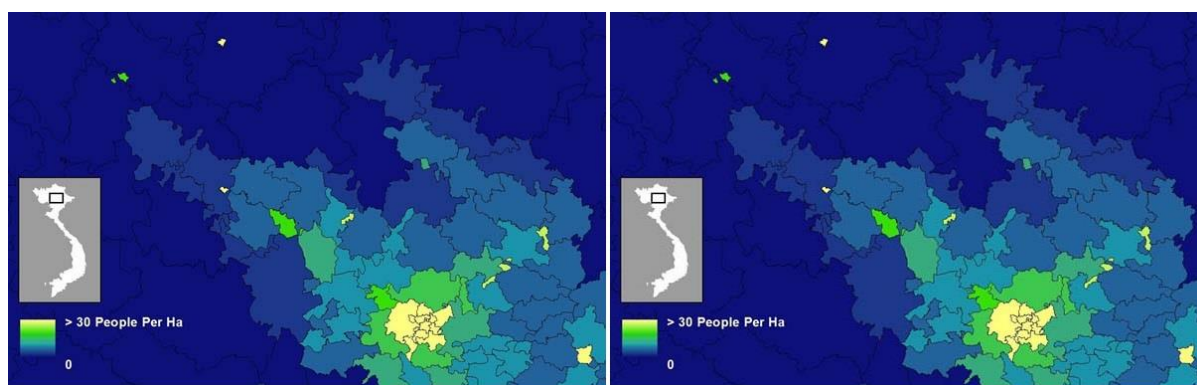


Figure 13. Left: input administrative level 2 census counts for northern Vietnam; right: output of random forest-based modelled population dataset. The layouts are showing perfect match. (WorldPop, 2020b).

Bottom-up population mapping: Where census data are outdated or unreliable, WorldPop has been collaborating with the Bill and Melinda Gates Foundation and Oak Ridge National Laboratories to develop approaches to estimating population distributions at high spatial resolution through a combination of satellite-derived

feature extractions and household surveys. Initial outputs are being completed and will be available on the WorldPop site in 2016, with some outputs already available for Nigeria in the country vaccination tracking system.

Intra-urban population mapping: to delineate, map, and predict urban extent and urban land uses in Africa, the Modelling and forecasting African Urban Population Patterns for vulnerability and health assessments (MAUPP) project (BELSPO, 2014) used by WorldPop in urban areas, works at two spatial resolutions: high resolution (HR ~30 m) and very high resolution (VHR < 5m).

The urban extent is defined as the Morphological Urban Area (MUA), which is the densely populated area where most of the economic and/or cultural activities occur and all the important transportation connections to other cities take place.

At the continent scale, a generalizable method based on statistical models to map the urban extension and the intra-urban densities of a set of ± 50 cities mainly using SAR and optical remote sensing as input data develop.

Information fusion extracted from the two datasets allows for the exploitation of redundancy, to increase global information and complementarity and to improve certainty and precision.

At the city scale, WorldPop develops a semi-automated method to map urban land covers and land uses to inform spatially detailed population distribution models. In addition, spatially detailed census data, historical maps, and existing population distribution datasets are used for modelling and predicting urban expansion and human population densities across Africa.

For this study, the dataset taken into consideration is the *constrained estimated age and sex structures of residential population per 100x100m grid square* (Google LLC, 2020), which makes use of a top-down constrained estimation approach summarized in Figure 14.

This approach has the advantage that where settlement/buildings are accurately mapped, the outputs present a more accurate population distribution, without small population numbers predicted in likely uninhabited areas. Additionally, the population is not distributed across a certain administrative unit.

However, it also has the following disadvantages:

- Relies on the accuracy of satellite-based settlement and building mapping.
- Where settlements or buildings are missed, this can lead to over-allocation of population to neighboring settlements. On the other hand, where settlements or buildings are incorrectly identified, under-allocation of population to neighboring settlements can occur.
- Datasets only currently available for 2020 – no multi-temporal data.

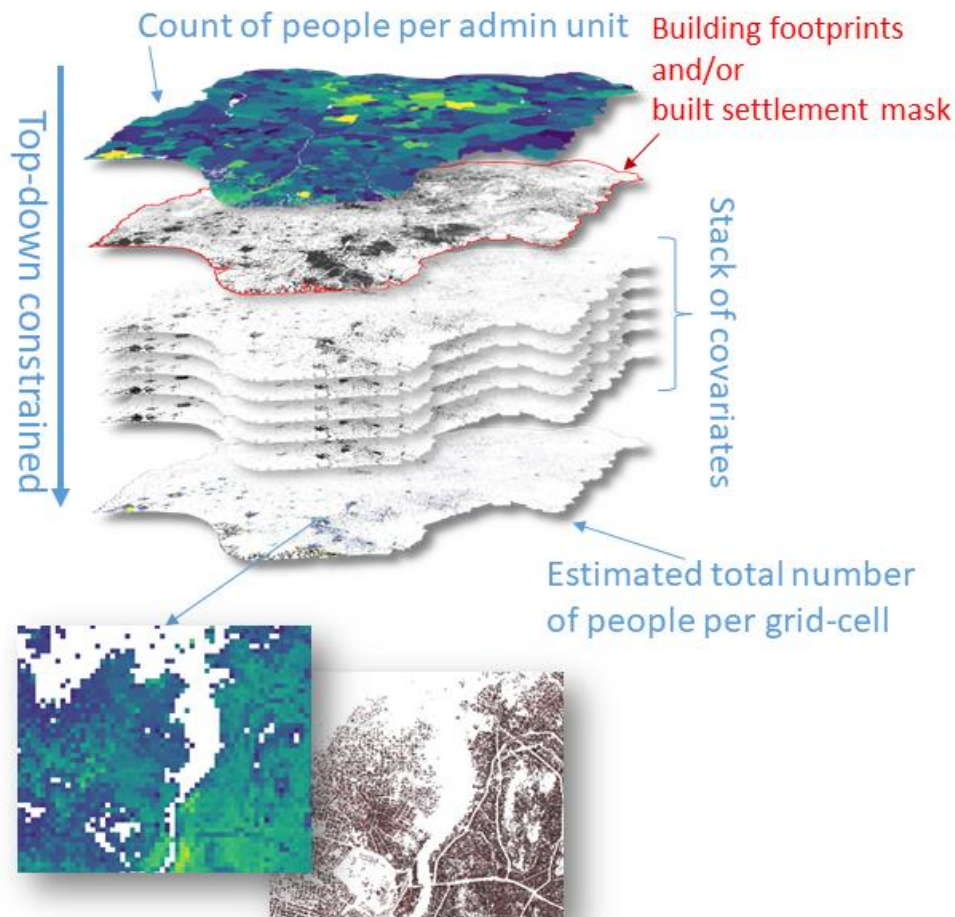


Figure 14. Top-down constrained population estimation approach adopted by WorldPop (WorldPop, 2020a).

In conclusion, the choice of this dataset is the only viable for this study, as in the area of interest all other datasets currently available do not feature a correct estimation in the most urbanized areas.

WorldPop data is therefore the one which is by default intersected with the surface water maps derived from Sentinel-1, in order to estimate the amount of population living in flooded areas. The decision of its use is made upon evaluation of available datasets present in the selected platform.

All other ready to use raster dataset present a lower spatial resolution, or, if accurate, do cover South Sudan (Meta, 2019) or don't feature a direct number of estimated people per cell (Sirko et al., 2021). A subset of the final dataset used, around the region of Pibor is shown in Figure 15.

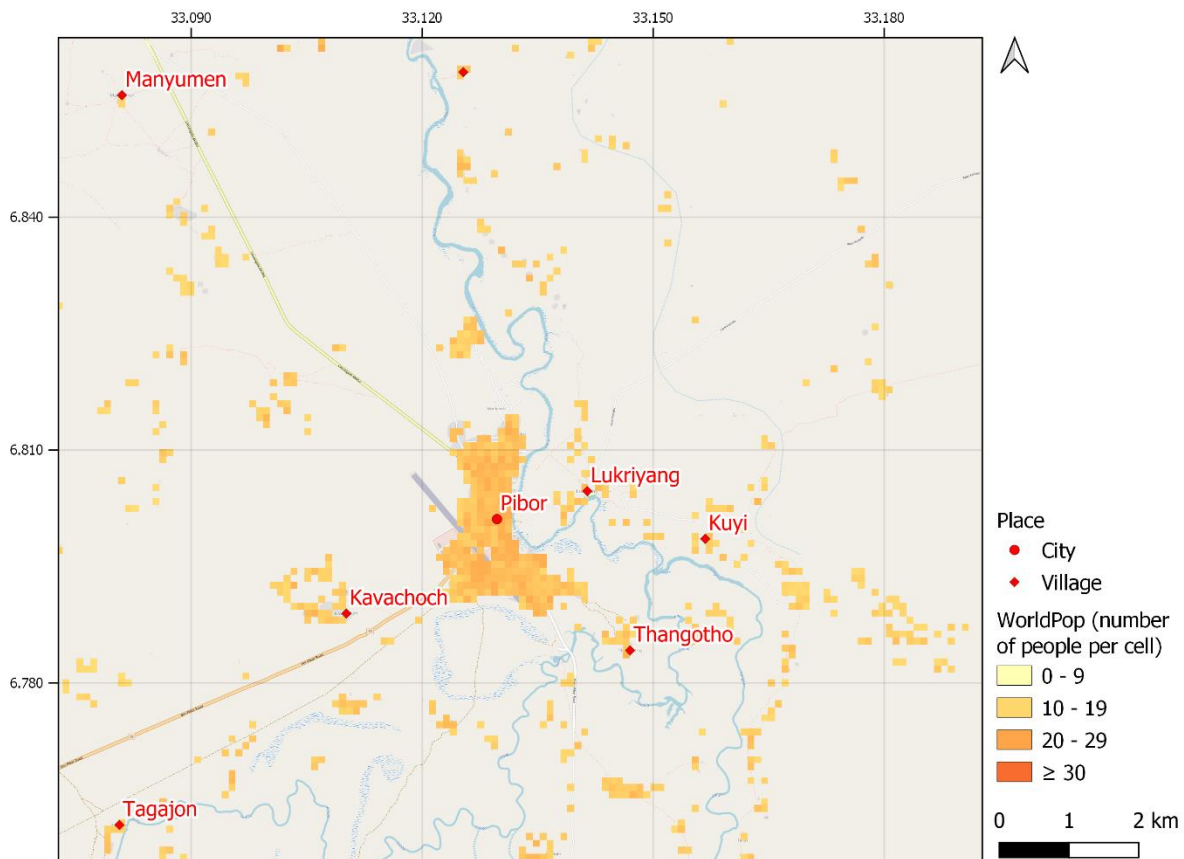


Figure 15. Region around Pibor, South Sudan. Estimated people per cell from World Pop data (www.worldpop.org, 2020). Background and places: © OpenStreetMap contributors.

OpenStreetMap

OpenStreetMap (OSM) (OpenStreetMap foundation, 2004) is a project to build a free geographic database of the world. Its aim is to ultimately have a record of every single geographic feature on the planet. While this started with mapping streets, it has already gone far beyond that to include footpaths, buildings, waterways, pipelines, woodland, beaches, postboxes, and even individual trees. Along with physical geography, the project also includes administrative boundaries, details of land use, bus routes, and other abstract ideas that aren't apparent from the landscape itself (Bennett, 2010).

The database is built by contributors, usually called mappers within OSM, who gather information by driving, cycling, or walking along streets and paths, and around areas recording their every move using Global Positioning System (GPS) receivers. This information is then used to create a set of points and lines that can be turned into maps or used for navigation.

Most mappers are volunteers working on the project in their spare time, although both commercial organizations and government bodies have started to contribute to the project.

The process of using groups of people to work on a task in this way, called crowdsourcing, is a recent phenomenon based around using the Internet to distribute tasks and gather the results. It's used by voluntary projects and commercial

organizations alike and has been particularly effective since broadband Internet connections became widely available.

Other data is gathered from out-of-copyright maps, public domain databases (ones with no copyright protection), or in some cases donations of proprietary databases by the companies that own them. In most cases, this needs further work to update and tidy the data, but it allows mappers to cover areas they can't get to, or for features that are difficult to survey.

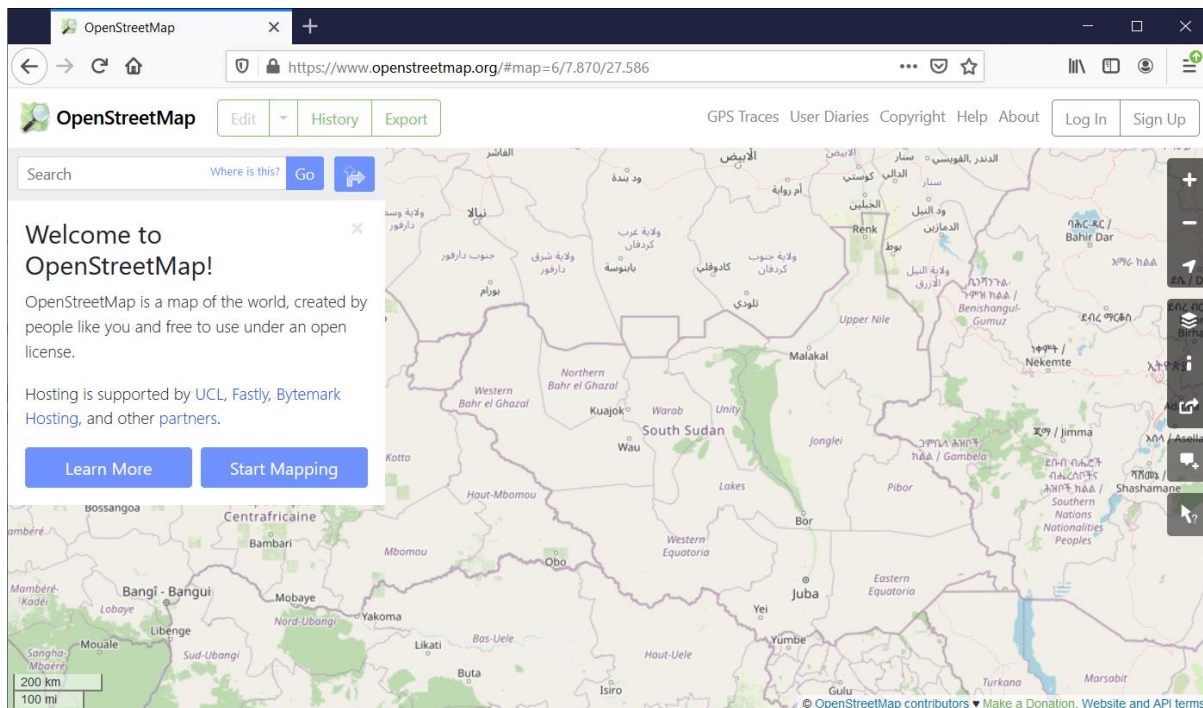


Figure 16. OpenStreetMap, main page. The canvas is centered on South Sudan. Data: © OpenStreetMap contributors.

For this study, OSM data is used for a twofold purpose: first, linear water features are buffered and optionally ingested in the image thresholding methodology, to drive water and other land covers sampling. Second, roads and places are used for visualization purpose, to demonstrate the possibility to derive the portion of road network flooded and add cartographic features to the results.

All the data from South Sudan have been gathered using Geofabrik download server (Geofabrik GmbH, 2007), which provides global daily dumps of the OSM database.

Figure 17 represents an example of OSM data ingested as asset in Google Earth Engine, where roads hampered by water are marked in purple and places are displayed as dots.

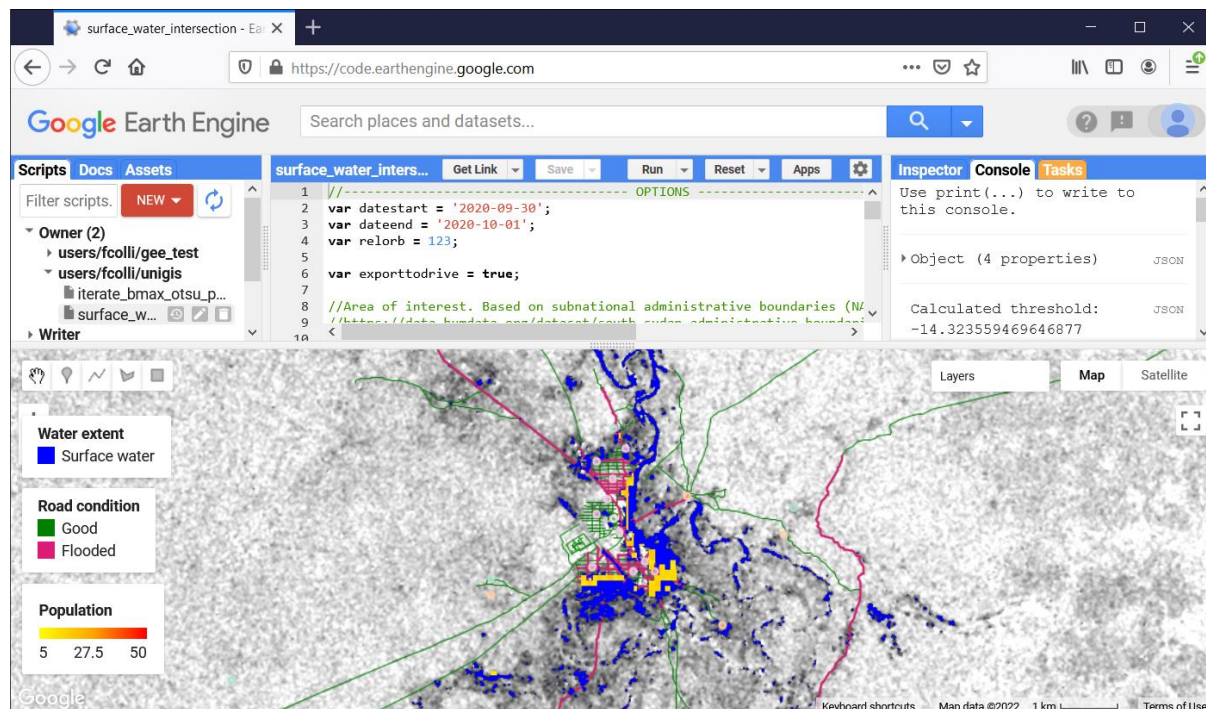


Figure 17. OSM data used as Google Earth Engine asset. Canvas: contains modified Copernicus Sentinel data [2020]. Population: WorldPop (www.worldpop.org 2020). Roads and places: © OpenStreetMap contributors.

ESRI land cover

Recently, a robust approach to derive a global and updated LULC map using Deep Learning (DL) and Copernicus Sentinel-2 data has been developed (Karra et al., 2021). The dataset produced has been made freely available by the Environmental Systems Research Institute (ESRI), which has also supported its development (ESRI, 2020) and has a resolution of 10m, which is suitable for the intersection with Sentinel-1 data.

The ESRI global LULC map features an overall accuracy of 85% across 10 classes:

1. Water.
2. Trees.
3. Grass.
4. Flooded vegetation.
5. Crops.
6. Scrub/shrub.
7. Built Area.
8. Bare ground.
9. Snow/Ice.
10. Clouds.

The majority of these classes are not relevant in this study. In fact, for the purpose of surface water extraction only the class water (1) and flooded vegetation (4) are potentially of interest and only water (1) can be used with a satisfactory level of confidence. A subset of the product covering the region of Pibor is depicted in Figure 18.

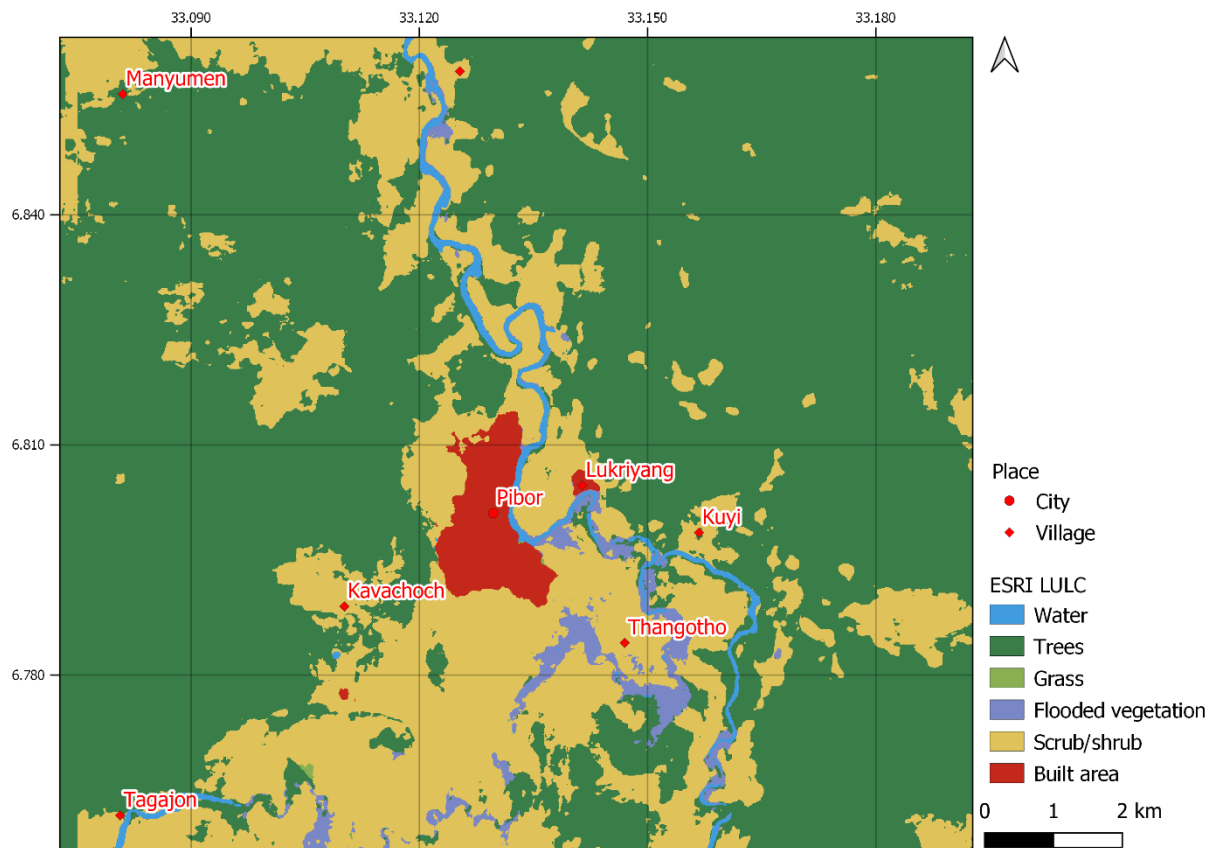


Figure 18. ESRI 10m global LULC (ESRI, 2020) around Pibor, South Sudan. Places: © OpenStreetMap contributors.

In the presented methodology, ESRI global LULC is the dataset of election to be used for the histogram sampling part in the adaptive-constrained thresholding approach.

Google Earth Engine

Google Earth Engine (GEE) (Gorelick et al., 2017) consists of a multi-source analysis-ready data catalog combined with a high-performance fast and parallel computation service. It is accessed and controlled through a web-based Application Programming Interface (API) associated with an Interactive Development Environment (IDE) that enables rapid prototyping and visualization of results.

The data catalog hosts a large repository of publicly available geospatial datasets, including acquisitions from a variety of satellite and aerial imaging systems in both optical and non-optical bands, environmental indicators, weather and climate data, land cover, topographic, and socio-economic datasets. All this data is preprocessed to a ready-to-use in an information-preserving form that allows efficient access and mitigates the barriers associated with data management.

Users can access, query, and analyze data from the public catalog as well as their own private assets using a library of operators provided by the embedded API. These operators are implemented in a large parallel processing system that automatically subdivides and distributes computations, providing high-throughput analysis capabilities. Users access the API either through a lightweight client library or through the web-based IDE mentioned earlier. Standard language for the web-based IDE is JavaScript, while the client library features API for Python as well.

Users can sign up for access at the GEE homepage (Google LLC, 2010), and access the user interface, the user's guide, and a large number of tutorials, examples, training videos, function references, and educational curricula. While prior experience with GIS, remote sensing and scripting makes it easier to get started, they are not strictly required, and the user's guide is oriented towards domain beginners. Accounts come with a quota for uploading personal data and saving intermediate products, and any inputs or results can be downloaded for offline use or exported to the user's drive.

While an example of the web-based IDE has been already introduced above in Figure 17, a simplified architecture of GEE is visible in Figure 19.

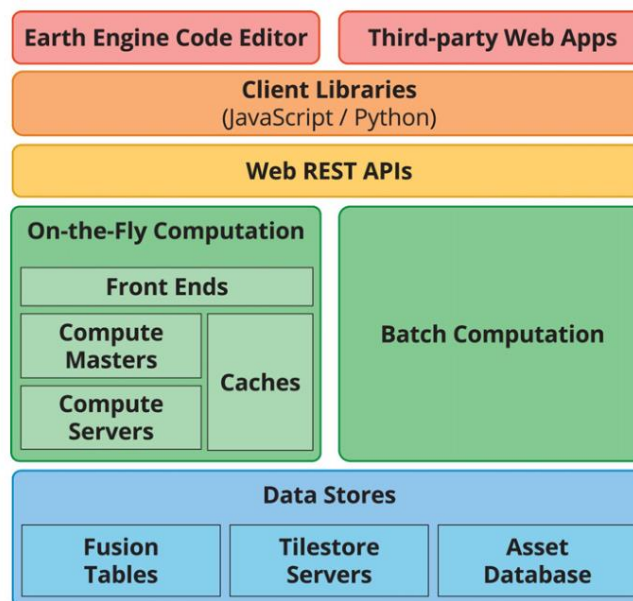


Figure 19. Simplified GEE architecture (Gorelick et al., 2017).

This research makes exclusive and extended use of GEE to embed the presented methodology. In addition, the methodology itself was formulated considering the datasets available in the platform, so that a ready-to-use solution with maximum portability and efficiency could be realized.

For data not directly available, they were uploaded and managed in the form of assets. The GEE platform is also exploited in the batch computation part, making easy and efficient the processing of the whole Sentinel-1 data archive for the extraction of surface water maps and the export of the corresponding statistics.

Methods

This section describes the methodology implemented to extract the information defined in the aims and objectives section at page 9. The entire methodological flow is summarized in Figure 20. For the processing and conditional blocks, implementing the pipeline core and for the cross validation of surface water maps, further subsections are presented to additionally detail the corresponding processes.

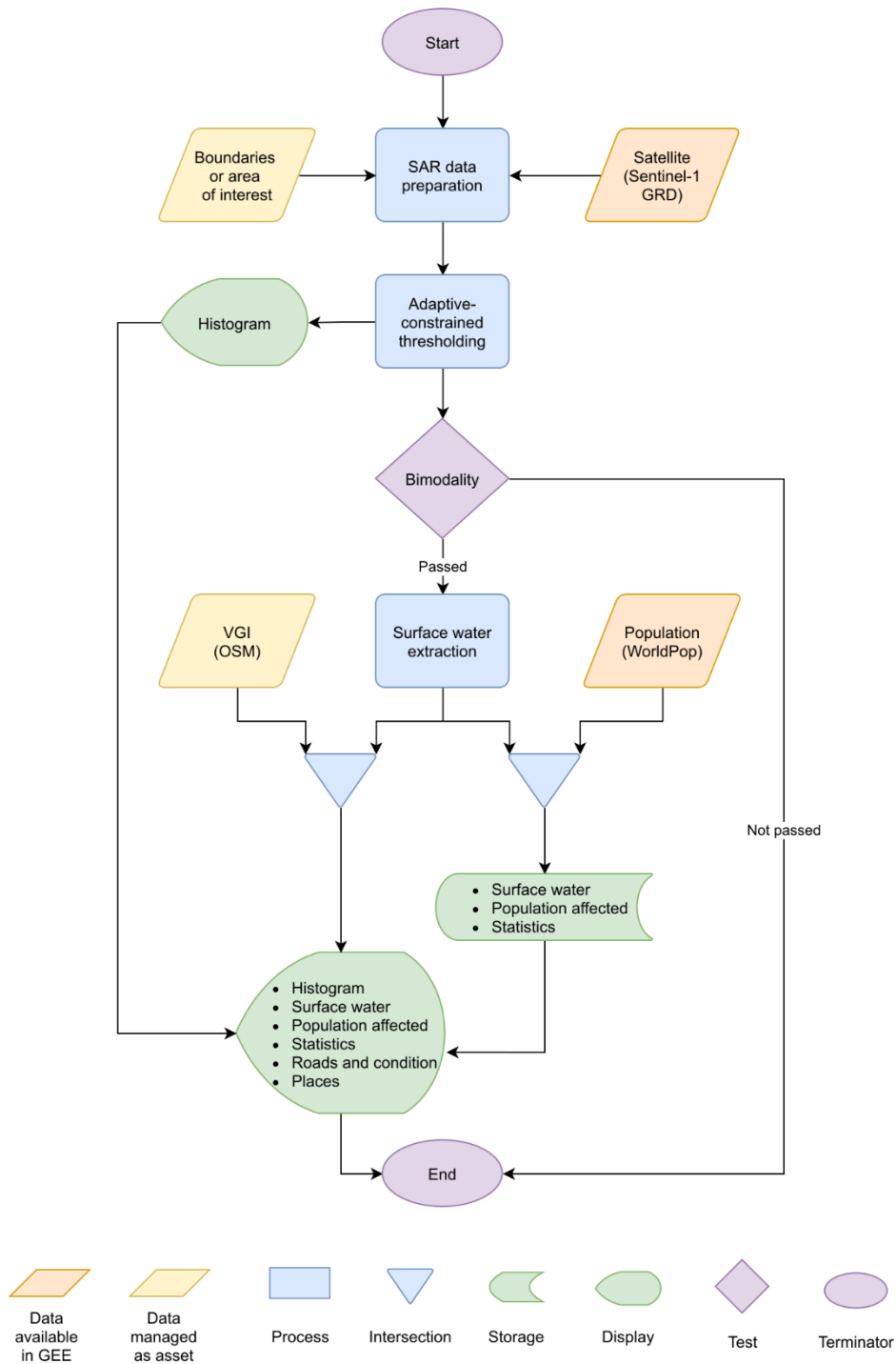


Figure 20. Methodology flow chart.

SAR data preparation

The imagery used in this study is available in GEE as *ImageCollection*, identified as *COPERNICUS/S1_GRD*. This collection consists of Level-1 GRD scenes processed to backscatter coefficient (σ°) in decibels (dB). The backscatter coefficient represents target backscattering area (radar cross-section) per unit ground area. Because it can vary by several orders of magnitude, it is converted to dB as $10 \cdot \log_{10}(\sigma^\circ)$. It measures whether the radiated terrain scatters the incident microwave radiation preferentially away from the SAR sensor (dB < 0) or towards the SAR sensor (dB > 0). This scattering behavior depends on the physical characteristics of the terrain, primarily the geometry of the terrain elements and their electromagnetic characteristics.

GEE uses the following pre-processing steps (ESA, 2014) to derive the backscatter coefficient in each pixel:

- Apply orbit file. Updates orbit metadata with a restituted orbit file (or a precise orbit file if the restituted one is not available).
- GRD border noise removal. Removes low intensity noise and invalid data on scene edges.
- Thermal noise removal. Removes additive noise in sub-swaths to help reduce discontinuities between sub-swaths for scenes in multi-swath acquisition modes.
- Radiometric calibration. Computes backscatter intensity using sensor calibration parameters in the GRD metadata.
- Terrain correction (orthorectification). Conversion of data from ground range geometry, which does not take terrain into account, to σ° , using different reference elevation models depending on the latitude.

The steps above allow the data to be ready-to-use. However, some additional steps can be further applied in order to have an optimal Analysis Ready Data (ARD) (Mullissa et al., 2021).

Beyond the data selection step, which does not manipulate in any way the input data, radiometric slope correction and speckle filtering are applied in this research.

Selection

This preparation step is limited to the selection of the correct data to be analyzed. However, its description is relevant as the current Sentinel-1 collection is already large and its size is foreseen to grow even more with the launch of more satellites (Jutz, 2020).

For the selection, when the correct collection is filtered, the following additional criteria are applied:

- *Geography*. An area of interest is ingested, or alternatively an administrative unit is indicated (e.g., Pibor).

- *Relative orbit.* Each Sentinel-1 overpass is labelled with a relative orbit number, extracted from the absolute orbit number, which guarantees the same acquisition configuration to be labelled with unique identifier over time, irrespective from the spacecraft. This is limiting the data to a single viewing geometry, decreasing the revisiting time. This filter is not mandatory. However, the results of this study are generated taking into account a single relative orbit (123) in order to reduce the large amount of data processed.
- *Acquisition date.* Only the data belonging to a specific time frame is selected. In the case of a single date, an interval of acquisition date + 1 day is specified.

Figure 21 graphically shows two out of three criteria: The data is filtered according to a specific relative orbit number and, within the returned set, only the frames intersecting with a specific geometry are selected for the analysis.

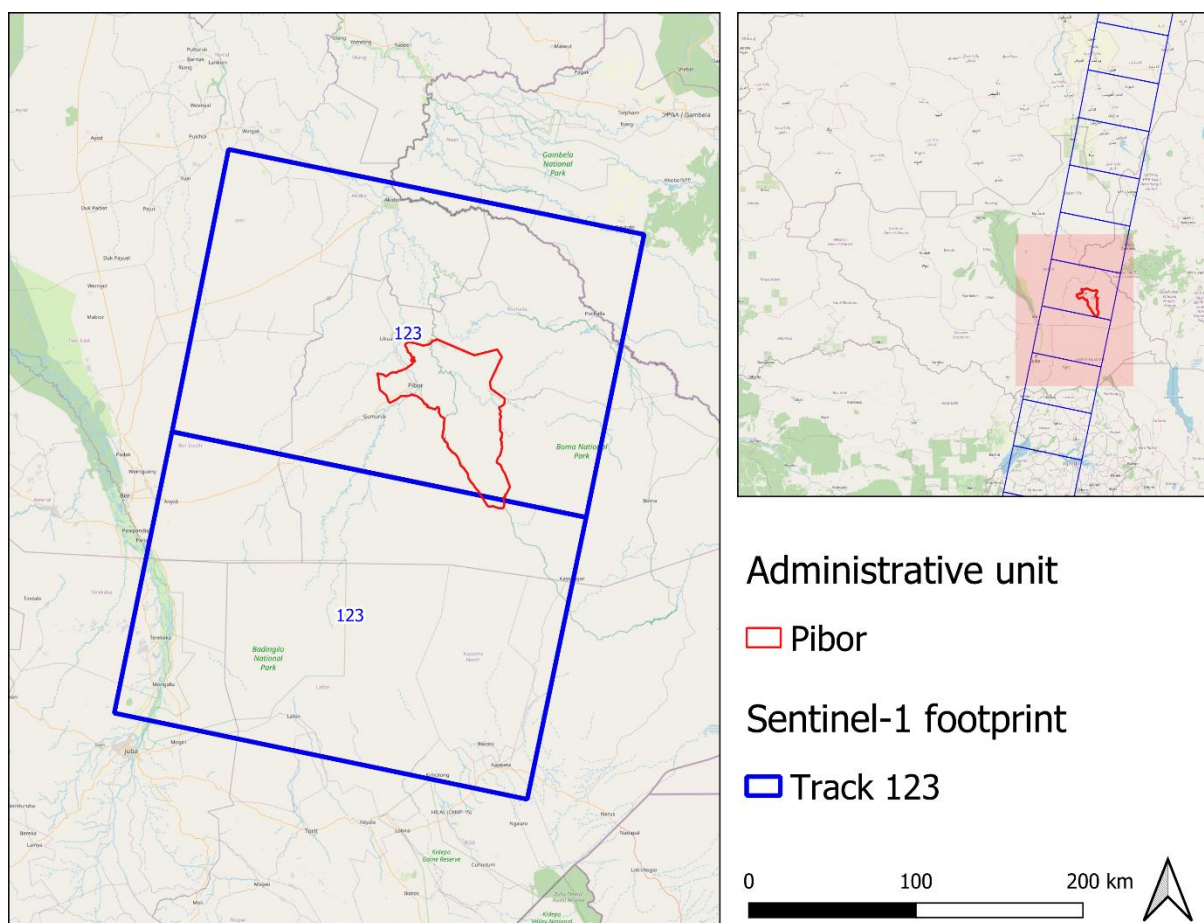


Figure 21. Pibor administrative unit and intersecting portion of Sentinel-1 stripe including 2 frames. The overview shows the full Sentinel-1 acquisition segment among South Sudan and beyond. Background: © OpenStreetMap contributors.

Radiometric slope correction

At the present time, there is no unique consensus on the specification of ARD products for SAR backscatter. However, to comply with the ARD data standards for land product, the Committee on Earth Observation Satellites (CEOS) (Committee on Earth

Observation Satellites (CEOS), 2021) suggests performing a radiometric slope correction.

For this reason, the present methodology makes use of a recently developed slope correction technique (Vollrath et al., 2020) in order to mitigate the undesired topography contribution particularly in mountainous regions, where the distortion effects introduced by the SAR viewing geometry (side-looking) are stronger.

The implemented correction is largely based on previously established physical reference models (Hoekman and Reiche, 2015) which depend on the angular relationships between the SAR image and the terrain geometry, that is extracted by using a given Digital Elevation Model (DEM). For the present research, the reference DEM used is Shuttle Radar Topography Mission (SRTM), which is widely used and featuring a satisfactory resolution along with good vertical accuracy (Farr et al., 2007).

Since the target area of this study is mostly flat, the effect of any slope correction algorithm would not be much noticeable. However, a mountainous region is present in South Sudan in the easternmost part of the country bordering Ethiopia and the present methodology is designed to be replicable across different areas. Hence, an example of slope correction is presented for the Boma area, approximately 150 km Southeast from Pibor, with the aim of demonstrating its effectiveness.

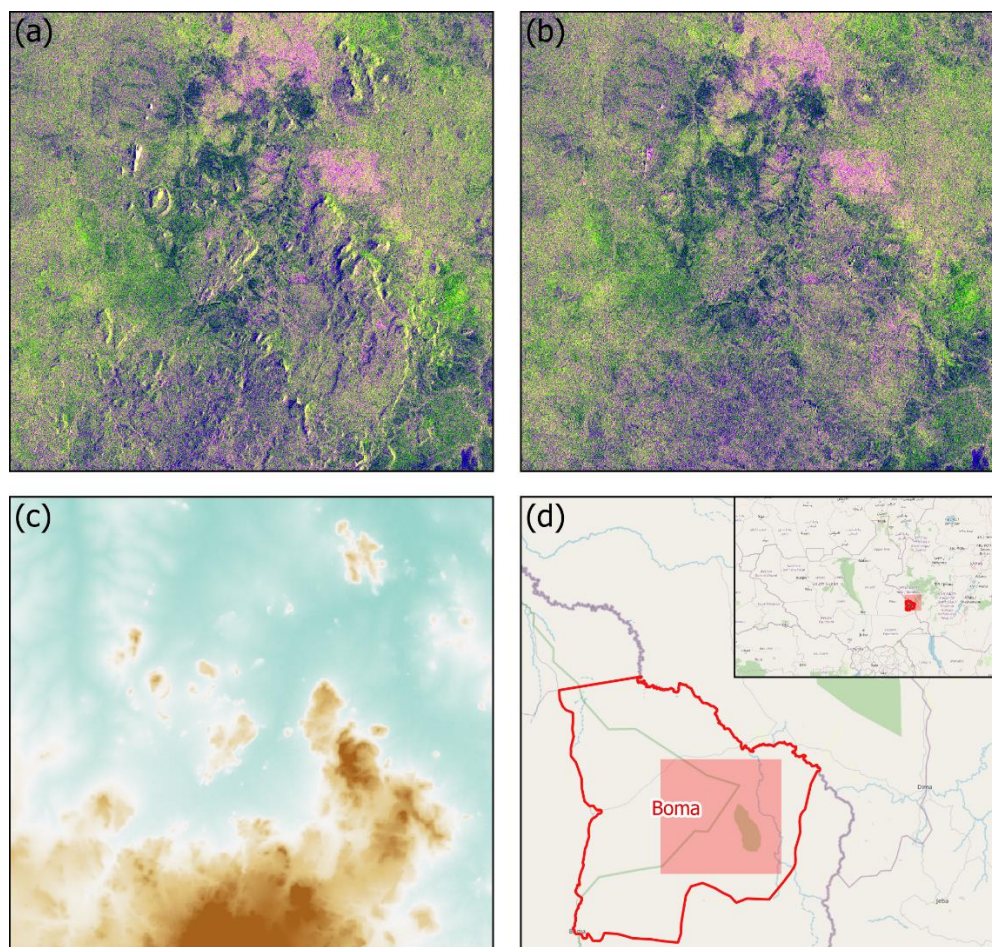


Figure 22. Effect of radiometric slope correction on Sentinel-1 image over Boma (19.09.2020). VV-VH color composite before (a) and after (b) the correction. SRTM DEM 30m (c) and map inset (d). Contains modified

Copernicus Sentinel data [2020]. Inset background: © OpenStreetMap contributors. SRTM data courtesy of NASA.

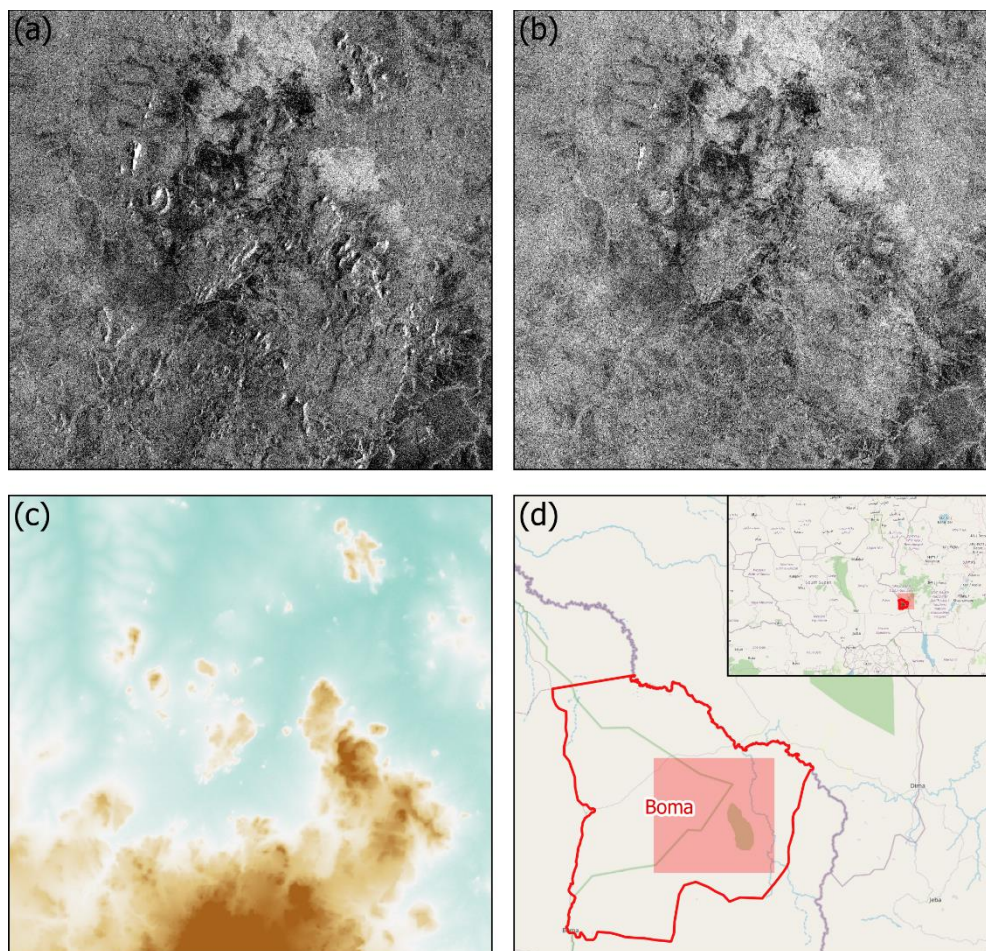


Figure 23. Effect of slope correction on Sentinel-1 image over Boma (19.09.2020). VV channel before (a) and after (b) the correction. SRTM DEM 30m (c) and map inset (d). Contains modified Copernicus Sentinel data [2020]. Inset background: © OpenStreetMap contributors. SRTM data courtesy of NASA.

Figure 22 and Figure 23 show the effect of the radiometric slope terrain correction over Boma, the DEM used and the location of the area. In both figures it appears clear how the distortion over mountainous (i.e., slopy) areas is mitigated by this processing step.

The removal of dark shadow features, particularly in the VV channel used in the present study, is the most important characteristic that this correction aims to achieve, to avoid misclassification of water in the latter stages of the methodology pipeline.

Speckle filtering

Speckle, appearing in SAR images as granular noise, is due to the interference of waves reflected from many elementary scatterers. Speckle in SAR images complicates the image interpretation problem by reducing the effectiveness of image segmentation and classification. To alleviate deleterious effects of speckle, various ways have been devised to suppress it (Lee et al., 1994).

For the present research, the implementation of different filtering methods has been evaluated.

Multi-temporal filtering approaches (De Grandi et al., 1997) are discarded at the outset, as each acquisition and each layer derived is meant to be analyzed individually. In fact, such a filtering approach could interfere with the real value of the single image, possibly mixing radiometry coming from different dates. It follows that this situation should be avoided, to always lead to a real estimation of the surface water for the corresponding date and prevent large commission or omission errors.

For what concerns single date denoising techniques, Lee (Lee, 1980), Kuan (Kuan et al., 1985) and Frost (Frost et al., 1982) filters can offer acceptable performances. However, all these filters assume the noise to be stationary across the whole image. Oppositely, the Gamma Maximum A-Posteriori (MAP) filter (Lopes et al., 1993), assumes that speckle is not stationary within the processing window (Maître et al., 2009).

This filter is therefore applied in this research after the radiometric slope correction on all the Sentinel-1 images considered, as appears to be the most suitable for multipurpose applications, especially in challenging regions where a multitude of landcovers are present. In addition, in the study area, the landscape can change abruptly due to natural events such as flooding, justifying even more the implementation of a flexible filtering technique.

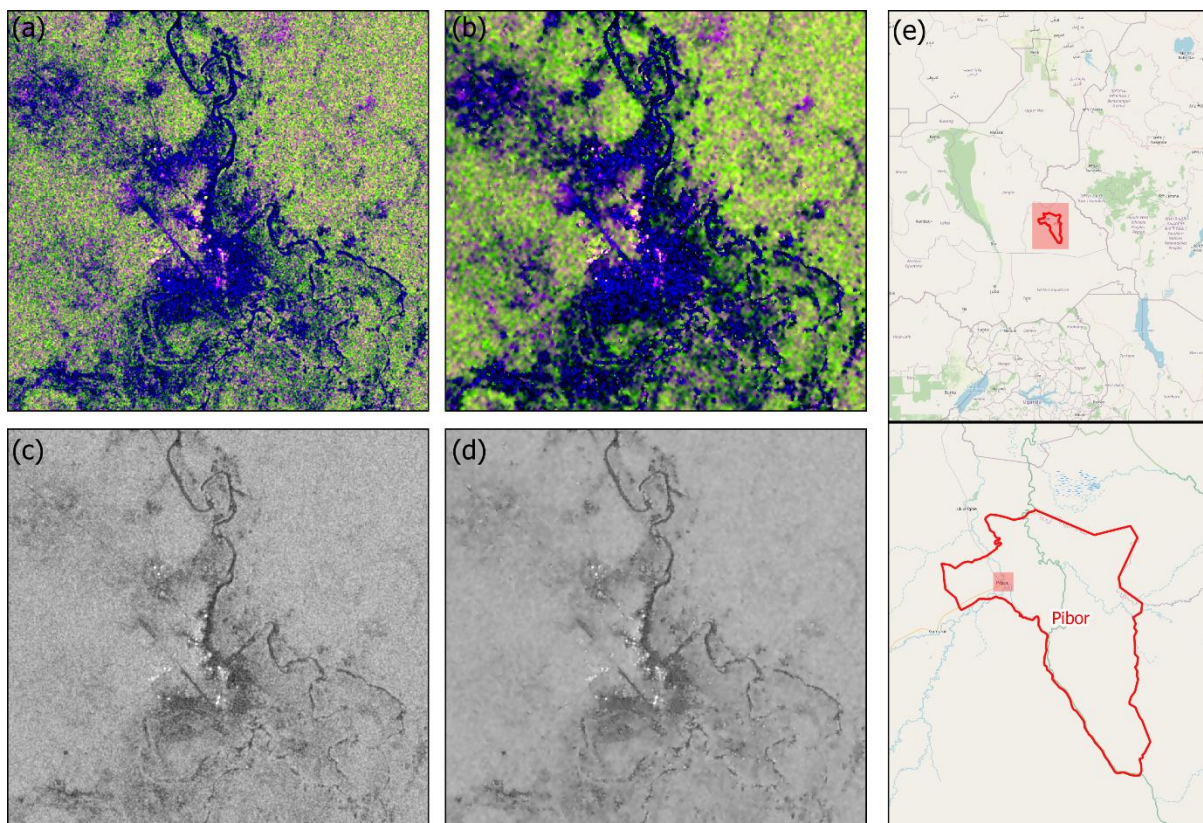


Figure 24. Effect of Gamma-MAP speckle filtering on Sentinel-1 slope corrected image around Pibor (30.09.2020). VV-VH color composite before (a) and after the filtering (b) and VV channel before (c) and after (d) the filtering, map inset (e). SRTM DEM 30m (c) and map inset (d). Contains modified Copernicus Sentinel data [2020]. Inset background: © OpenStreetMap contributors.

Figure 24 shows the effect of the Gamma-MAP filtering on a closeup around Pibor. It is visible, especially on the VV channel image, that the multiplicative noise affecting the original data is efficiently removed.

Surface water extraction

This section presents the core methodology for surface water maps extraction. The step, which occurs downstream of data preparation, uses exclusively SAR imagery and an existing land cover map as auxiliary data.

Surface water extraction is performed, for each individual acquisition, by means of VV channel thresholding. Particularly, a maximum is determined so that any pixel that is associated with a backscatter value lower than it is considered water. The threshold is derived automatically by evaluating a histogram, computed only using the areas falling within specific classes of the land cover map.

In addition, the applications developed for this research allow to apply the classical methodology where the determination of the threshold takes place on the histogram derived from the entire image, or to define custom areas of extraction by polygon.

For completeness, two use cases relevant to the same acquisition date are reported: the first, ordinary thresholding, using no delineation area; the second, adaptive-constrained thresholding, first introduced in this research, using the ESRI global land cover as auxiliary data.

Ordinary adaptive thresholding

Given a grey level image, this method (Otsu, 1979) is based on finding a threshold which minimizes the weighted within-class variance, that corresponds to maximize the between-class variance for a certain number of classes.

As the aim of surface water extraction is separate two classes, this technique can be effectively applied on SAR images, since water features appear darker than everything else. Namely, generally referring to grey-level images, this is equivalent to separate a background from a foreground.

The method has the advantage to operate directly on a grey level histogram. Once the histogram is computed, the weighted within-class variance can be defined as:

[1]

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t)$$

Where the class probabilities are estimated as:

[2]

$$q_1(t) = \sum_{i=1}^t P(i) \quad q_2(t) = \sum_{i=t+1}^I P(i)$$

In addition, the class means are given by:

[3]

$$\mu_1(t) = \sum_{i=1}^t \frac{iP(i)}{q_1(t)} \quad \mu_2(t) = \sum_{i=t+1}^I \frac{iP(i)}{q_2(t)}$$

While the individual class variances are:

[4]

$$\sigma_1^2(t) = \sum_{i=1}^t [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \quad \sigma_2^2(t) = \sum_{i=t+1}^I [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)}$$

The total variance can be expressed as:

[5]

$$\sigma^2 = \sigma_w^2(t) + q_1(t)[1 - q_1(t)][\mu_1(t) - \mu_2(t)]^2$$

Since this quantity remains constant and independent of t , the effect of changing the threshold is merely to move the contributions of the two terms back and forth. It follows that, minimizing the within-class variance is the same as maximizing the between-class variance, which is defined as:

[6]

$$\sigma_B^2(t) = q_1(t)[1 - q_1(t)][\mu_1(t) - \mu_2(t)]^2$$

And that can be easily computed recursively as t is moved within a certain range.

Even though is always possible to find a threshold t which satisfies the maximum between-class variance criteria, this method is effective only under some assumptions:

- The distribution in histogram used is bimodal.
- Spatial coherence or any other notion of object structure is ignored.
- Statistics are stationary across the application area.
- Illumination is uniform.

Particularly the first point, that assumes bimodality, it is likely to be not satisfied in many cases. For the area of interest, where large water bodies are not present and the surface water reaches its peak of extent with seasonal trend, the application of the method without proper caution might lead to incorrect results.

For instance, if the Sentinel-1 acquisition from 30.09.2020 is taken, the histogram computed and the threshold which maximizes the between-class variance derived, the result would be the one shown by the plot in Figure 25.

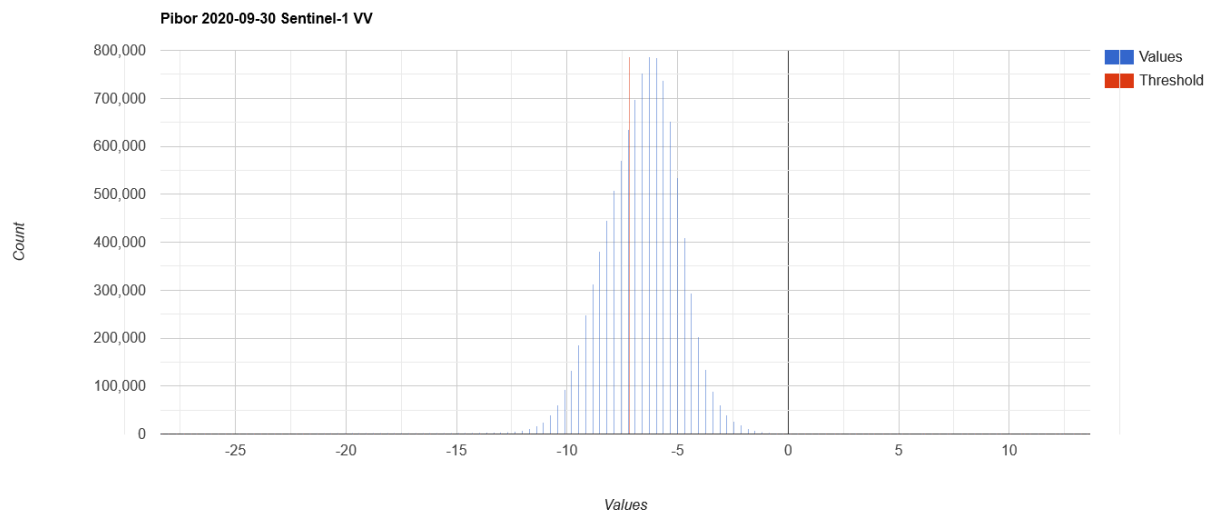


Figure 25. Histogram computed for the Sentinel-1 VV acquisition of 30.09.2020 and relevant threshold derived using the ordinary thresholding method.

In this case, the threshold of -7.197 dB found on a distribution that appears unimodal, leads to large overestimation of all the variables as show in the application overview in Figure 26.

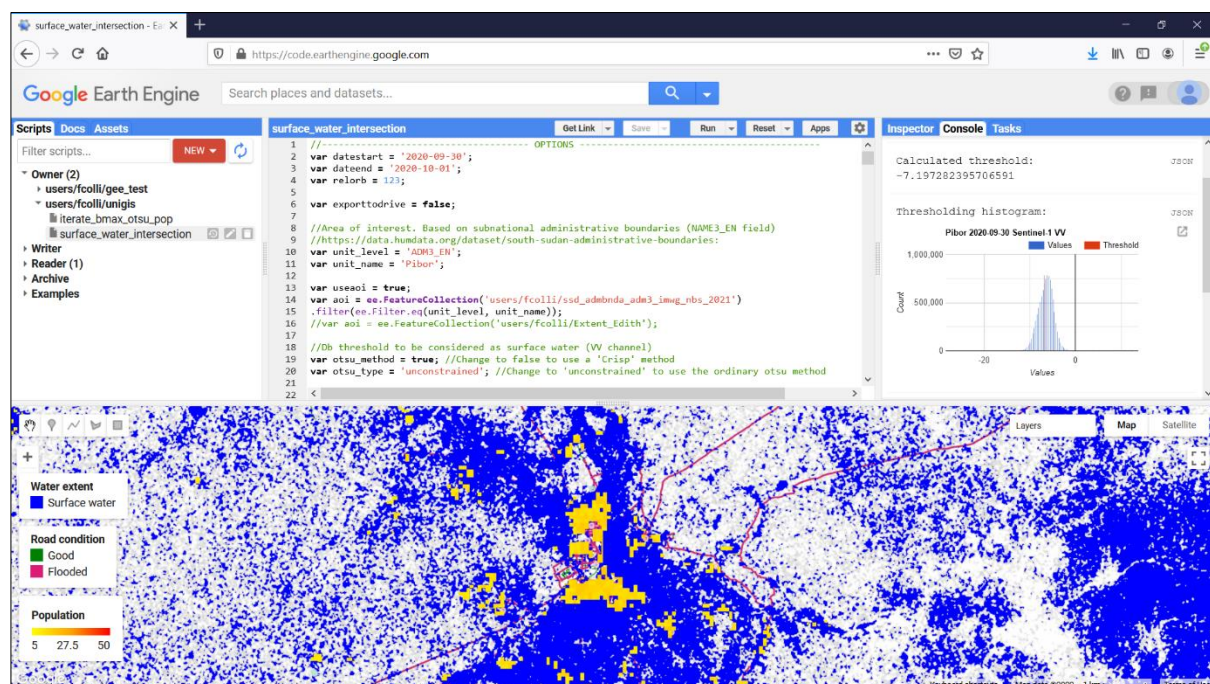


Figure 26. Ordinary thresholding in GEE. The threshold derived on the whole image leads to a large commission error, visible in the map canvas. Histogram used, threshold found, and additional information are always redirected to the console for supervision purposes.

Adaptive-constrained thresholding

To overcome the limitations of the ordinary thresholding method described above, some approaches have been proposed.

Beyond an appropriate bimodality test, implemented a-posteriori in this research, two constraining methods known as Bmax-Otsu and Edge Otsu (Markert et al., 2020) have been recently developed.

Both of the methods largely rely either on bimodal-like distributions across the whole image or on the detection large water bodies, identified using an edge-detector (Canny, 1986) or by the application of anisotropic diffusion (Perona and Malik, 1990), which diverges into a binary mask highlighting outlines. However, for the study area no of the boundary condition set by these methodologies are satisfied.

Therefore, as adaptive-constrained thresholding it is intended a method which preserves the adaptive feature of the ordinary, constraining the calculation of the thresholding histogram to areas delineated by means of auxiliary data. This auxiliary information can come either from an existing raster dataset (i.e., a land cover), ingested by the user (arbitrary polygons), or from a line asset (e.g., VGI buffered hydrological networks).

For the same acquisition considered above, the thresholding histogram is extracted on the areas overlapping with the class “1. Water” of the product described above. This specific class, isolated from the others, is visible in Figure 27 at variable scales over three locations in the Pibor region.

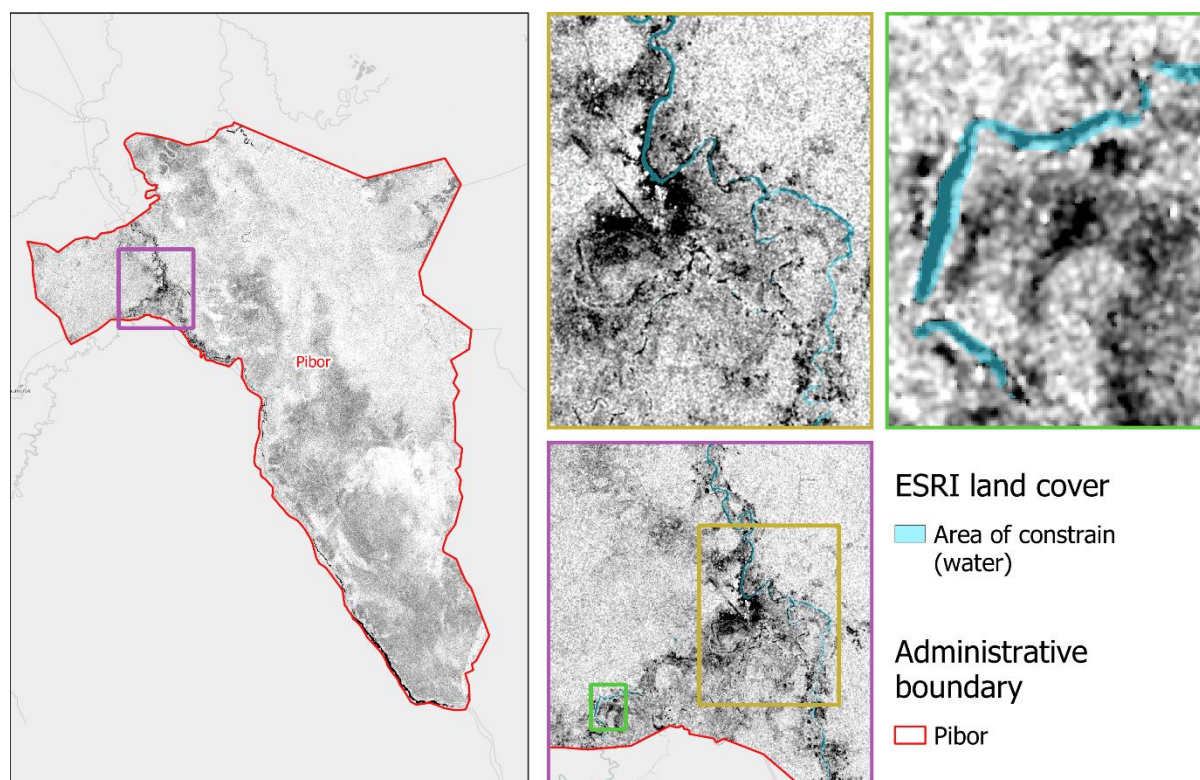


Figure 27. Pibor, Sentinel-1 VV acquisition from 30.09.2020, slope corrected, and Gamma-MAP filtered. The insets are showing the intersection with the “1. Water” class of the ESRI land cover (ESRI, 2020). The samples falling within the area of constrain are used for the computation of the thresholding histogram. Contains modified Copernicus Sentinel data [2020]. Background: © OpenStreetMap contributors.

When the Sentinel-1 VV imaged is sampled, the threshold t maximizing the between-class variance in equation [6] can be determined on a bimodal distribution. In fact, even though the class extracted from the ESRI land cover is denominated “1. Water”, the layer has been spatially filtered and aggregated. It follows that “1. Water” includes other land covers as well, particularly flooded vegetation for which the accuracy is not satisfactory (Karra et al., 2021), making possible to sample both water and not water in the evaluated set of values.

Figure 28 shows the thresholding histogram, with the derived value of -14.324 dB at the center of it. This threshold corresponds to a relative minimum separating the two classes (water/other), for which the in-between class variance is maximized; this threshold is 7 dB lower than the one derived using the ordinary method. In addition, the absolute value results in line with the expectation of water reflectance for SAR VV images at C-band (Hong et al., 2015).

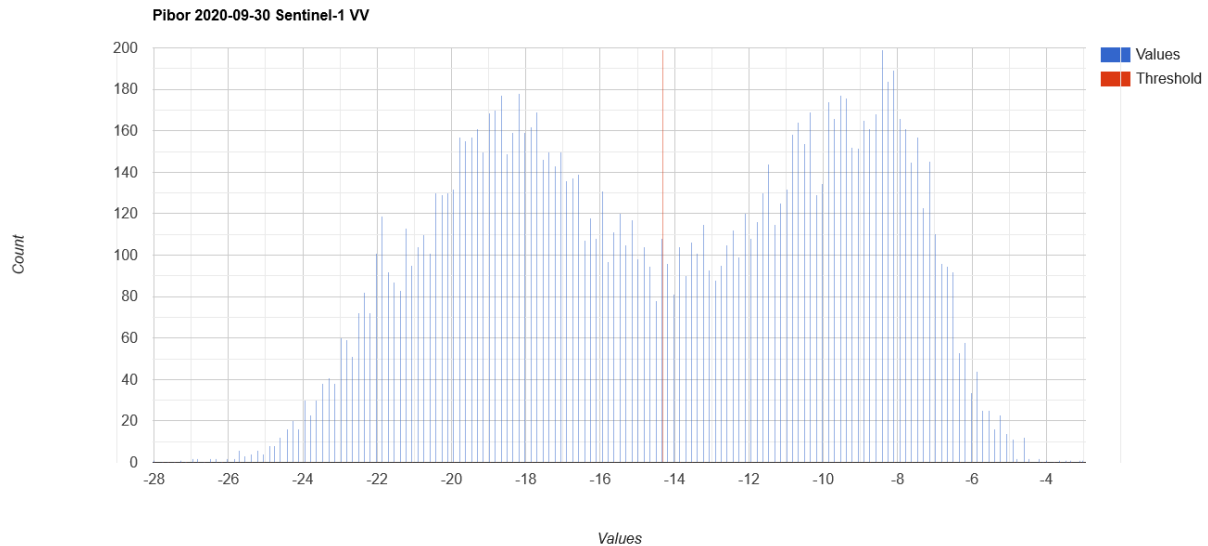


Figure 28. Histogram computed for the Sentinel-1 VV acquisition of 30.09.2020 and relevant threshold derived using the adaptive-constrained thresholding method.

After the histogram evaluation, thresholding is applied on the whole image. The overview, visible in Figure 29, shows the estimation of surface water, which is covering a reasonable extent. Dark grey features are referring to moisty, not completely submerged, areas.

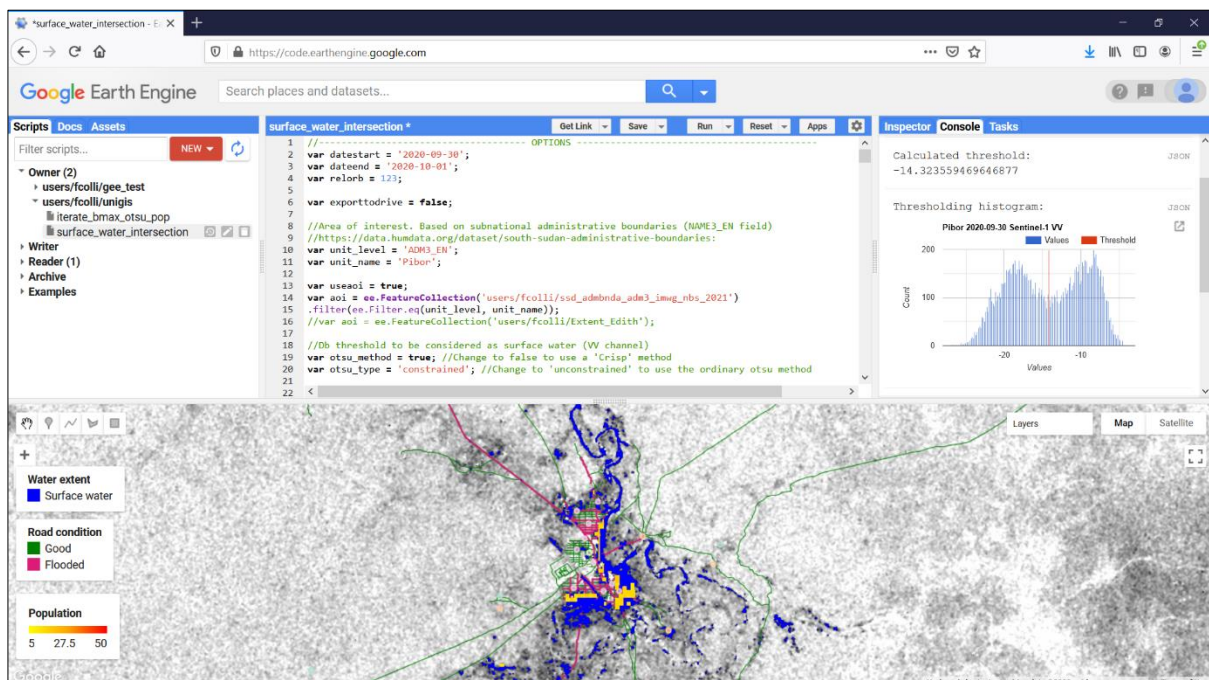


Figure 29. Adaptive-constrained thresholding in GEE. The threshold derived leads to reasonable water extent estimation.

Bimodality test

As introduced above, adaptive thresholding methods are performing efficiently when applied on sets of data featuring bimodal distribution. Therefore, a bimodality test is implemented to control the validity of the threshold found.

In this research the test is performed a-posteriori, as the method is based on the computation of variances, which must be calculated in any case during the thresholding step.

A diagnostic indicator of bimodality can be found in the normalized between-class variance, defined as:

[7]

$$B(t) = \frac{\sigma_B^2(t)}{\sigma^2}$$

Where $\sigma_B^2(t)$ is computed for the threshold as the maximum of [6] and σ^2 as [5]. The normalized quantity $B(t)$ can be easily evaluated, and should be greater than 0.75 for general distributions to be bimodal (Demirkaya and H. Asyali, 2004).

After processing the Sentinel-1 archive from 2016 to 2021, the distribution of $B(t)$ for all acquisitions appears as depicted in the graph in Figure 30.

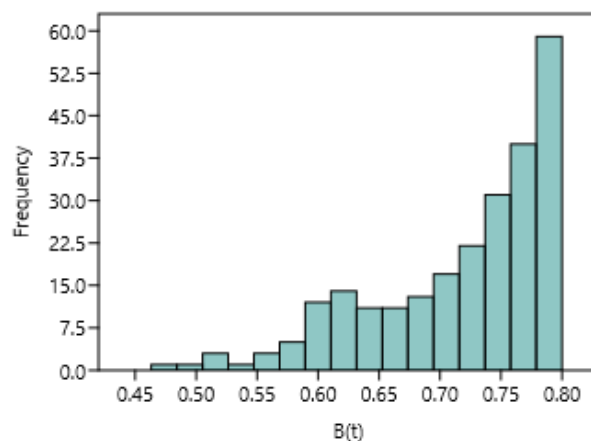


Figure 30. Distribution of $B(t)$.

While the relevant summary statistics are reported in Table 3.

Table 3. Summary statistics for $B(t)$.

| | |
|-----------------|----------|
| N | 244 |
| Min | 0.477957 |
| Max | 0.821005 |
| Sum | 175.5958 |
| Mean | 0.719655 |
| Std. error | 0.004767 |
| Variance | 0.005546 |
| Stand. dev | 0.074469 |
| Median | 0.743372 |
| 25th-percentile | 0.668607 |
| 75th-percentile | 0.776792 |

| | |
|------------|----------|
| Skewness | -0.91954 |
| Kurtosis | 0.166251 |
| Geom. mean | 0.71551 |
| Coeff. var | 10.34792 |

$B(t)$ presents therefore a skewed minus distribution, with maximum frequencies above 0.75. These ranges of values are the ones where the bimodality confidence is the highest. Once only the samples belonging to this interval are selected, the distribution of the relevant thresholds derived is plotted, resulting in the histogram in Figure 31.

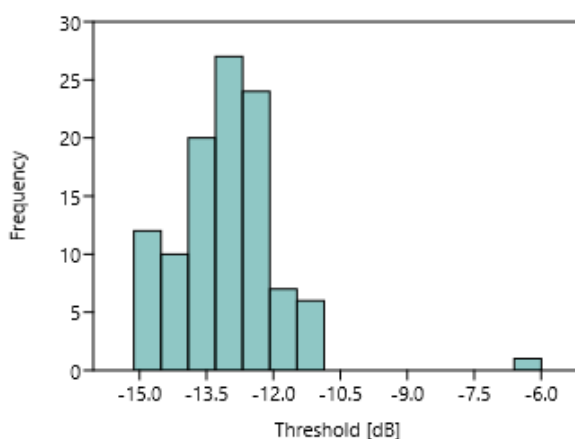


Figure 31. Distribution of threshold [dB], after the application of the bimodality test.

The corresponding summary statistics for threshold [dB] are then summarized in Table 4.

Table 4. Summary statistics for thresholds derived for the acquisition where the bimodality test is passed.

| | |
|-----------------|----------|
| N | 107 |
| Min | -15.123 |
| Max | -5.99683 |
| Sum | -1394.43 |
| Mean | -13.0321 |
| Std. error | 0.1156 |
| Variance | 1.429887 |
| Stand. dev | 1.195779 |
| Median | -13.0348 |
| 25th-percentile | -13.6765 |
| 75th-percentile | -12.4029 |
| Skewness | 1.754079 |
| Kurtosis | 10.19782 |
| Geom. mean | 0 |
| Coeff. var | -9.17568 |

From the statics above, appears clear that the bimodality test is effective in most of cases. However, some outliers featuring high values of threshold persist. Even though these samples represent less than 30% out of the total, and that their assessment could not be straightforward except for very high values (the maximum of -5.996 dB is related to an image with radiometric problems) it reasonable for this study to operate in the most conservative way.

Therefore, along the bimodality test, only the images returning a threshold lower than -12.5 dB, which is 0.1 dB lower than the 75th-percentile boundary (Table 4), are considered. Undertaking this cautionary measure, the final set of surface water mask is extracted on a set of images featuring the $B(t)$ and threshold distribution depicted by the histograms in Figure 32.

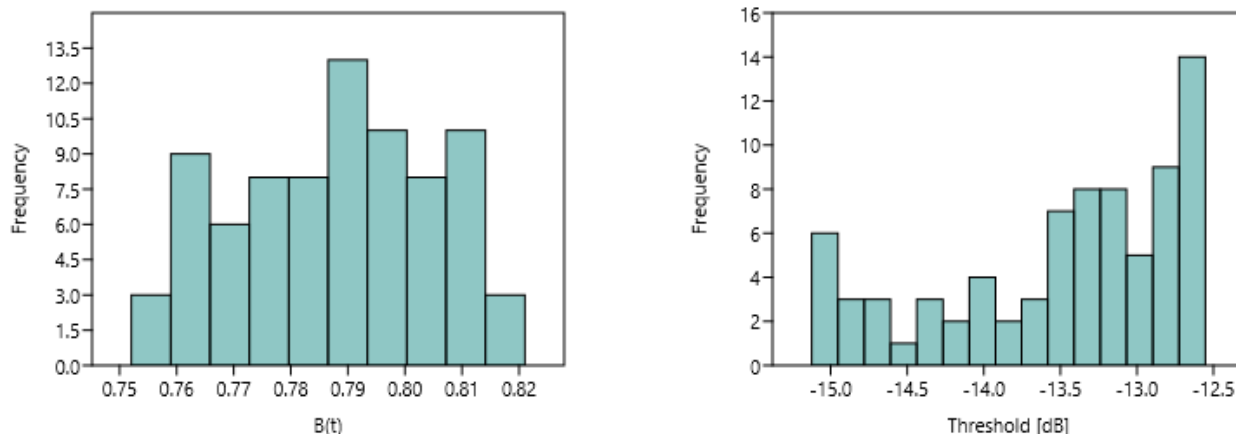


Figure 32. $B(t)$ and threshold distribution of the final set.

To which correspond the summary statistics in Table 5.

Table 5. Final dataset. Summary statistics of $B(t)$ and threshold [dB].

| | $B(t)$ | Threshold [dB] |
|-----------------|----------|----------------|
| N | 78 | 78 |
| Min | 0.752077 | -15.123 |
| Max | 0.821005 | -12.5536 |
| Sum | 61.45537 | -1053.65 |
| Mean | 0.787889 | -13.5083 |
| Std. error | 0.002022 | 0.086862 |
| Variance | 0.000319 | 0.588506 |
| Stand. dev | 0.017854 | 0.767141 |
| Median | 0.787662 | -13.3592 |
| 25th-percentile | 0.773712 | -14.0028 |
| 75th-percentile | 0.802866 | -12.8775 |
| Skewness | -0.11097 | -0.66601 |
| Kurtosis | -1.00885 | -0.64258 |
| Geom. mean | 0.787689 | 0 |
| Coeff. var | 2.266061 | -5.67902 |

It is worth to note that the filtering criteria can be tuned and fully customized in the applications developed for this research. In fact, both $B(t)$ value can be accepted lower, for instance considering the range above the 25-percentile (Table 3) or ingesting an arbitrary value. For the threshold of -12.5 dB, assessed by expert knowledge and explorative data analysis, the same could be applied or the condition could be removed. Both tunings, however, should be avoided when the user intends to operate in the safest domain, and it is highly recommended to implement restrictive controls especially when evaluating large amount of data.

Cross-validation

Validation of products derived using remote sensing is a necessary practice. However, ground data are frequently lacking for areas such as the one of interest in this study. Therefore, in the present research, the surface water maps accuracy assessment is carried out by means of reference points extraction from optical data acquired by Sentinel-2.

Sentinel-2 archive has been previously analyzed in search of acquisitions made on the same date of those suitable for surface water mapping using Sentinel-1. The query resulted in the list summarized in Table 6.

Table 6. Sentinel-2 acquisitions on the same date of those considered for Sentinel-1 and relevant cloud cover over the entire tile, as returned by the relevant acquisition metadata.

| Date | Cloud cover (%) |
|------------|-----------------|
| 22/10/2017 | 40.7379 |
| 21/11/2017 | 0.02 |
| 21/12/2017 | 0 |
| 11/11/2019 | 38.2343 |
| 11/12/2019 | 92.1575 |
| 10/01/2020 | 8.0569 |
| 09/02/2020 | 25.5672 |
| 08/06/2020 | 6.2516 |
| 07/08/2020 | 42.1543 |
| 05/11/2020 | 17.641 |
| 05/12/2020 | 62.0856 |
| 04/01/2021 | 90.8222 |
| 02/08/2021 | 80.5717 |

As expected, only few acquisitions are available among which 21.11.2017 and 21.12.2017 are completely or almost cloud free and can be used for the purpose.

To avoid a time consuming and error-prone photo interpretation process, the decision is to use an index calculated using the available spectral bands. To detect water, the best index would be the Normalized Difference Water Index (NDWI) (Gao, 1996). However, the index is computed using band 8 (Near Infrared) and band 12 (Medium Infrared), where band 12 has a spatial resolution of 20 meter only.

Therefore, the use of Normalized Difference Vegetation Index (NDVI) (Pearson and Miller, 1972), that can be generated using band 8 (near Infrared) and band 4 (red), both at 10 m resolution seems to be an acceptable proxy. This index, in fact, is effective to detect water as normally its value is negative or weakly positive for this land cover type (Fan et al., 2020).

To produce the validation sets, random sampling of points on pixels featuring values of NDVI lower than 0.05 is performed. This sampling is done over a subset of 10000x10000 pixels including Pibor Post.

As the corresponding surface water layer coming from the same date has a class share (water vs. other) which is strongly unbalanced, the number of samples for each class should not be even. For 21.12.2017, its number is therefore computed using the quantities in Table 7.

Table 7. Variables used for number of validation samples calculation for 21.12.2017.

| Class | PixelSum | Percentage (%) | Area [sqm] | W_i | S_i | $S_i * W_i$ |
|-------|----------|----------------|-------------|---------|-------|-------------|
| Other | 453297 | 95.64499079 | 44701774.57 | 0.95645 | 0.2 | 0.19129 |
| Water | 20640 | 4.35500921 | 2035408.633 | 0.04355 | 0.5 | 0.021775 |
| Total | | | | | | 0.213065 |

W_i and S_i are respectively the percentage of each class divided by 100 and a quantity assessed by conjecture, which should be proportional to the expected accuracy for the relevant class. If S_0 is defined as the expected standard deviation for overall accuracy, and assumed as 0.01, the number of samples N can be calculated as (Olofsson et al., 2014):

[8]

$$N = \left(\sum_{i=1}^C \frac{W_i S_i}{S_0} \right)^2$$

Which returns, for the current case, $N = 453$. The distribution of samples for each class can be then calculated as the mean of even distributed sets and weighted sets, as summarized in Table 8.

Table 8. Validation samples distribution across classes for 21.12.2017.

| Class | Weighted | Equal | Mean |
|-------|----------|-----------|------|
| Other | 434.1968 | 226.98353 | 330 |
| Water | 19.77031 | 226.98353 | 123 |

It follows that 330 and 123 samples are collected for each class. After the number of samples of each class is determined, random selection is done within the NDVI image, returning the points depicted in Figure 33 and Figure 34.

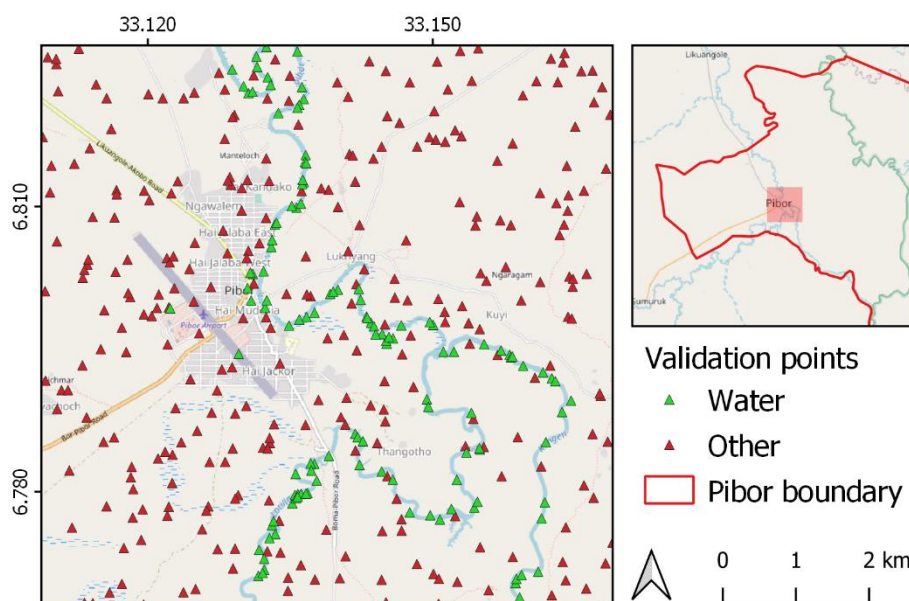


Figure 33. Validation site and validation set for 21.12.2017. Background: © OpenStreetMap contributors.

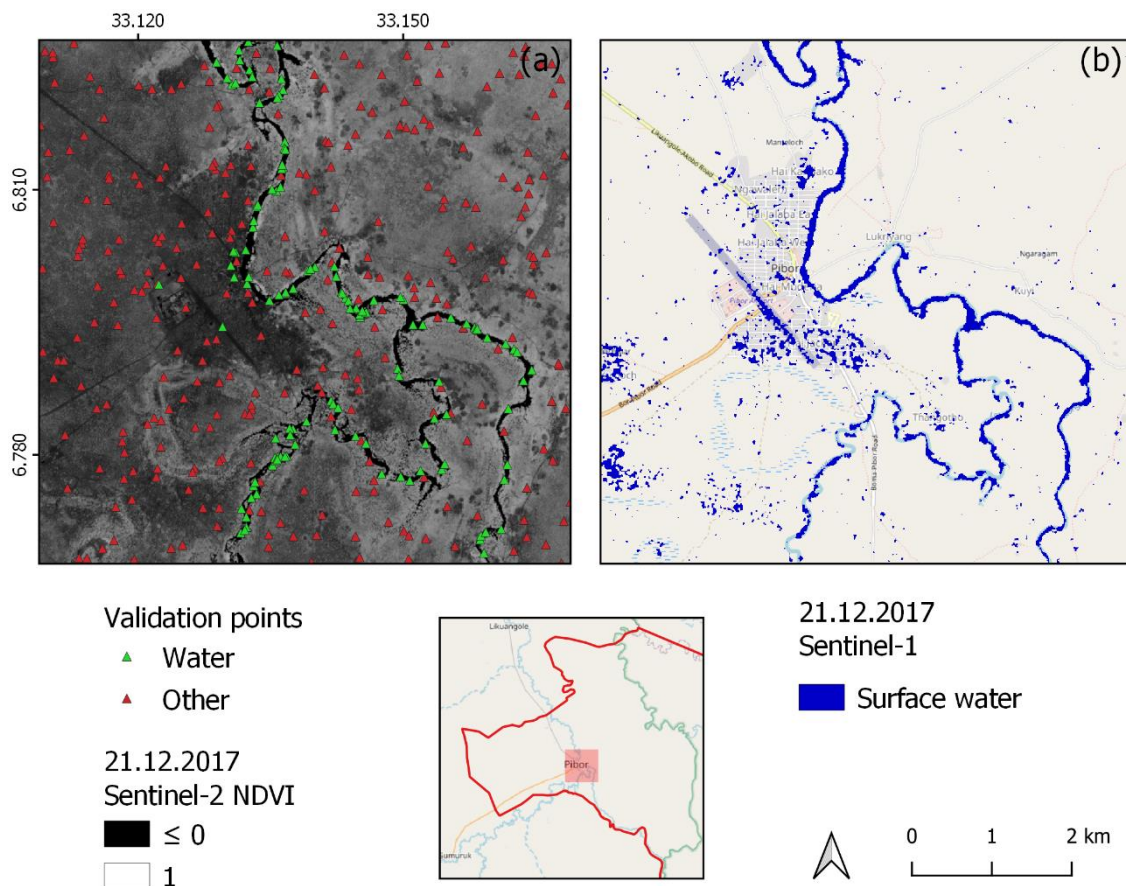


Figure 34. Validation set overlaid to Sentinel-2 NDVI, date 21.12.2017 and surface water map extracted from Sentinel-1 for the same date (b). Contains modified Copernicus Sentinel data [2017]. Background: © OpenStreetMap contributors.

The validation points are used to compute the confusion matrix in Table 9, where:

- Accuracies of the classification are reported for each class, that represent the fraction of correctly classified pixels regarding all pixels of that ground truth class. Overall accuracy is reported in the yellow background cell.
- Reliabilities of classes are shown as the fraction of correctly classified pixels regarding all pixels classified as that class in the classified image.
- Green cells report the correctly classified samples, while the red summarize the omission and commission errors.

Table 9. Confusion matrix for the surface water layer of 21.12.2017 validated against Sentinel-2 NDVI from the same date.

| | | Predicted class | | Accuracy |
|--------------|-------|-----------------|--------|----------|
| | | Water | Other | |
| Actual class | Water | 110 | 13 | 89.43% |
| | Other | 38 | 292 | 88.48% |
| Reliability | | 74.32% | 95.74% | 88.74% |

The procedure above is repeated for the second most cloud free date, 21.11.2017. As the class shares change, the number of samples to extract is computed again using the quantities in Table 10.

Table 10. Variable used for number of validation samples calculation for 21.11.2017

| Class | PixelSum | Percentage (%) | Area [sqm] | W_i | S_i | $S_i * W_i$ |
|-------|----------|----------------|------------|----------|-------|-------------|
| Other | 459576 | 96.96985 | 45957600 | 0.969699 | 0.2 | 0.19394 |
| Water | 14361 | 3.03014 | 1436100 | 0.030301 | 0.5 | 0.015151 |
| | | | | | Total | 0.20909 |

After applying [8], the number of samples returned is $N = 378$ with the shares shown in Table 11.

Table 11. Validation samples distribution across classes for 21.11.2017.

| Class | Weighted | Equal | Mean |
|-------|----------|----------|------|
| Other | 376.1261 | 189.2108 | 282 |
| Water | 2.295437 | 189.2108 | 96 |

A map layout, in Figure 35, shows the validation set resulting from random selection overlaid to the relevant NDVI and the corresponding surface water map.

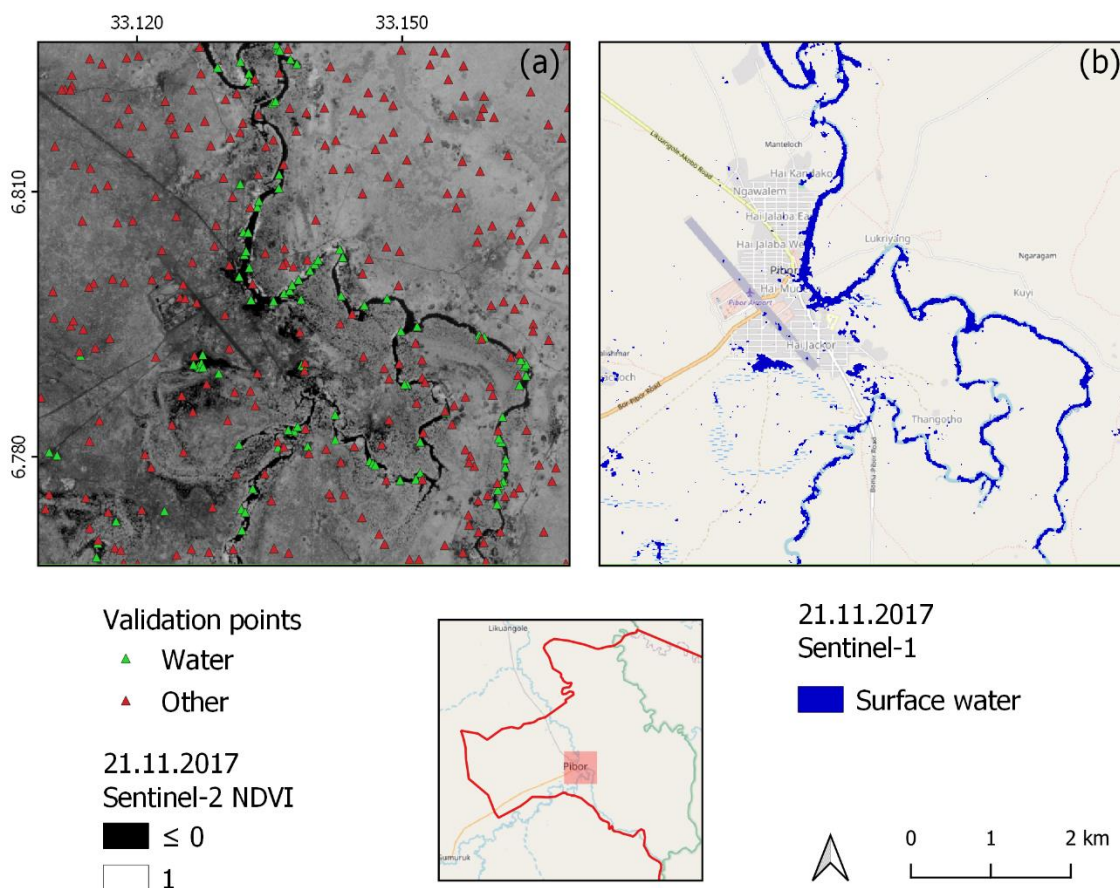


Figure 35. Validation set overlaid to Sentinel-2 NDVI, date 21.11.2017 and surface water map extracted from Sentinel-1 for the same date (b). Contains modified Copernicus Sentinel data [2017]. Background: © OpenStreetMap contributors.

A confusion matrix is computed and summarized in Table 12.

Table 12. Confusion matrix for the surface water layer of 21.11.2017 validated against Sentinel-2 NDVI from the same date.

| | | Predicted class | | Accuracy |
|--------------|-------------|-----------------|--------|----------|
| | | Water | Other | |
| Actual class | Water | 96 | 15 | 86.49% |
| | Other | 33 | 282 | 89.52% |
| | Reliability | 74.42% | 94.95% | 88.73% |

The production of a third cross validation set, using the acquisition of 08.06.2020 featuring a cloud cover of 6.25% over the full Sentinel-2 tile is considered. However, the acquisition is affected by haze and cirrus, which might hamper the correct NDVI extraction as depicted in Figure 36. In addition, the lack of a thermal band in the Sentinel-2 instrument would make challenging a proper cloud and shadow masking. The decision is therefore to avoid the use of 08.06.2020.

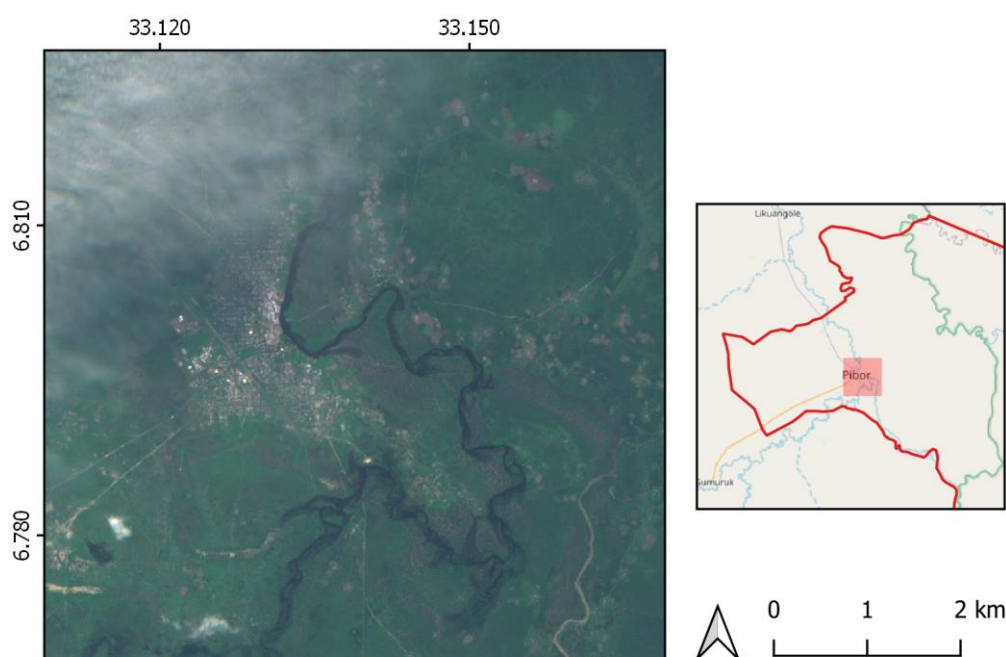


Figure 36. Sentinel-2 natural color composite over Pibor, 08.06.2020. Contains modified Copernicus Sentinel data [2017]. Overview: © OpenStreetMap contributors.

The two cross-validations performed show an overall accuracy of 88.74% and 88.73%. Accuracies for water class range between 86.49% and 89.43%. The reason of lower reliability, around 74% for the water class has to be found in the unbalanced class share, which is although mitigated by the uneven selection of the number of samples.

In summary, the accuracies look satisfactory for a remote sensing product automatically derived from a single data source. Further cross validations could be undertaken in future to better assess the validity of the product.

Population data intersection

For each surface water extent layer derived by means of the methodology described in the sections above, a relevant population layer is extracted. The process, which consist in an intersection between the two layers, allows to estimate the population living in flooded areas for each Sentinel-1 satellite overpass, suitable for surface water mapping as per the conditions explained in the Bimodality test section.

As Sentinel-1 is available at a pixel size of 10 m, while WorldPop data has a coarser resolution of 100 m, a nearest neighbor resampling is first applied on the surface water layer. Then, the obtained layer is used as mask for WorldPop data. This approach is similar to the one adopted by the United Nations for flood mapping and impact assessment (United Nations, 2020), which uses a combination of change detection between subsequent dates and thresholding. However, in this research, higher resolution population data is used, and a single-date mapping strategy is adopted. In fact, the use of a single acquisition date appears more appropriate, as it allows to detect persistent water through different acquisition, potentially enabling to map flood duration in the case of long and slow events.

At the end of the intersection, the resulting masked population layer can be summarized, to extract a single figure representing the total estimated population living in flooded areas. Figure 37 represents the intermediate steps of the process, from the original surface water map to the final estimated population layer.

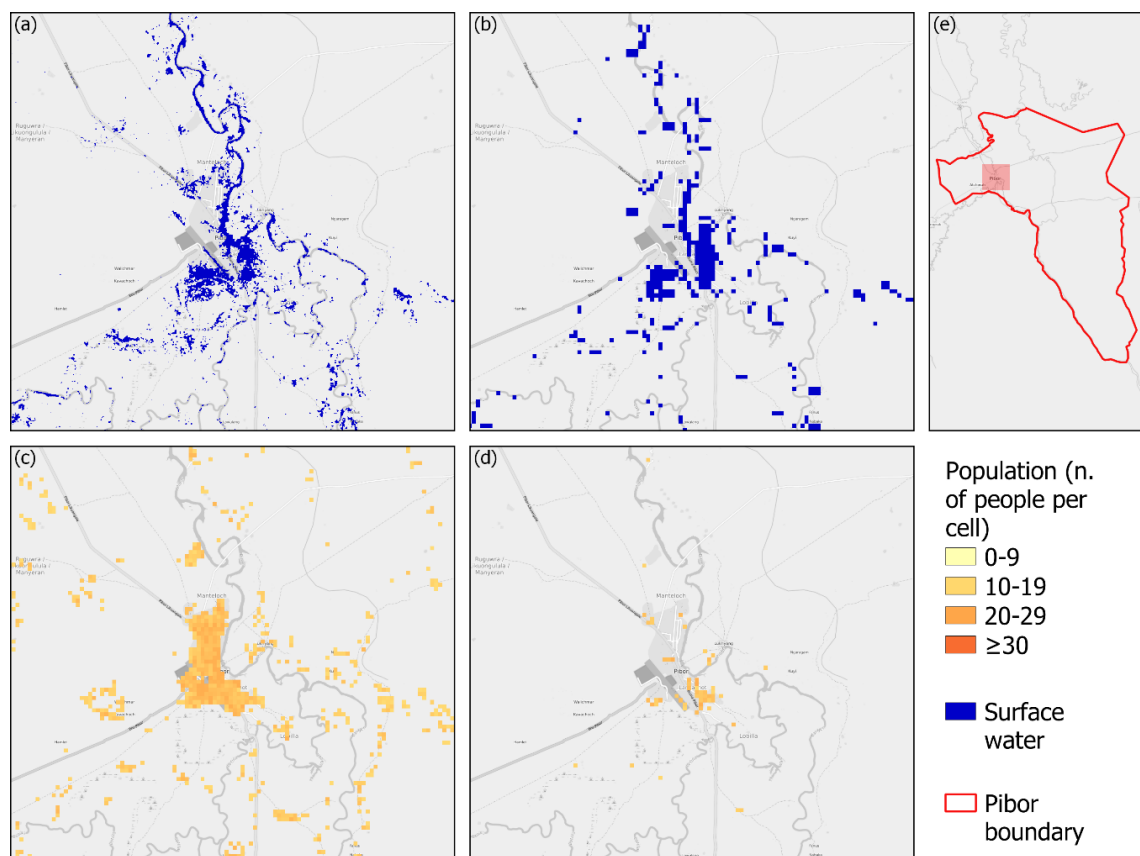


Figure 37. Surface water – population intersection process. Surface water extent at 10 m (a), resampled surface water extent at 100 m (b), WorldPop full layer (c), estimated population affected layer (d) and map overview (e). Contains modified Copernicus Sentinel data. Population: WorldPop (www.worldpop.org 2020). Background: © OpenStreetMap contributors.

Some limitations and source of mistake exist for the proposed approach. First, the resampling method, nearest neighbor, utilized by default (Google LLC, 2021) might be suboptimal. Since in this case the resolution is reduced, a resampling strategy which considers the majority value of the pixels (i.e., mode) falling within the coarser resolution cell would be appropriate. In this study, the maintenance of default parameters in the GEE core functions has been privileged. However, the change of the reprojection method can be explored in future research.

The second major drawback is given by the precision and the frequency update of WorldPop. In fact, the constrained population layer, which offers the best features in spatial terms, is present in GEE for the year 2020. It follows that the population amount is considered static across all the years. However, the methodology remains valid also due to fact that these demography layers are intrinsically uncertain. In addition, the offered ready-to-use solution can be possibly improved at no cost, when a more updated/better layer will be available.

The intersection process has been graphically described using offline layers, exported by means of the GEE export functionality at fixed scale. When the application is used online, the population affected is rendered in the canvas and summarized in the console on-fly as depicted in Figure 38.

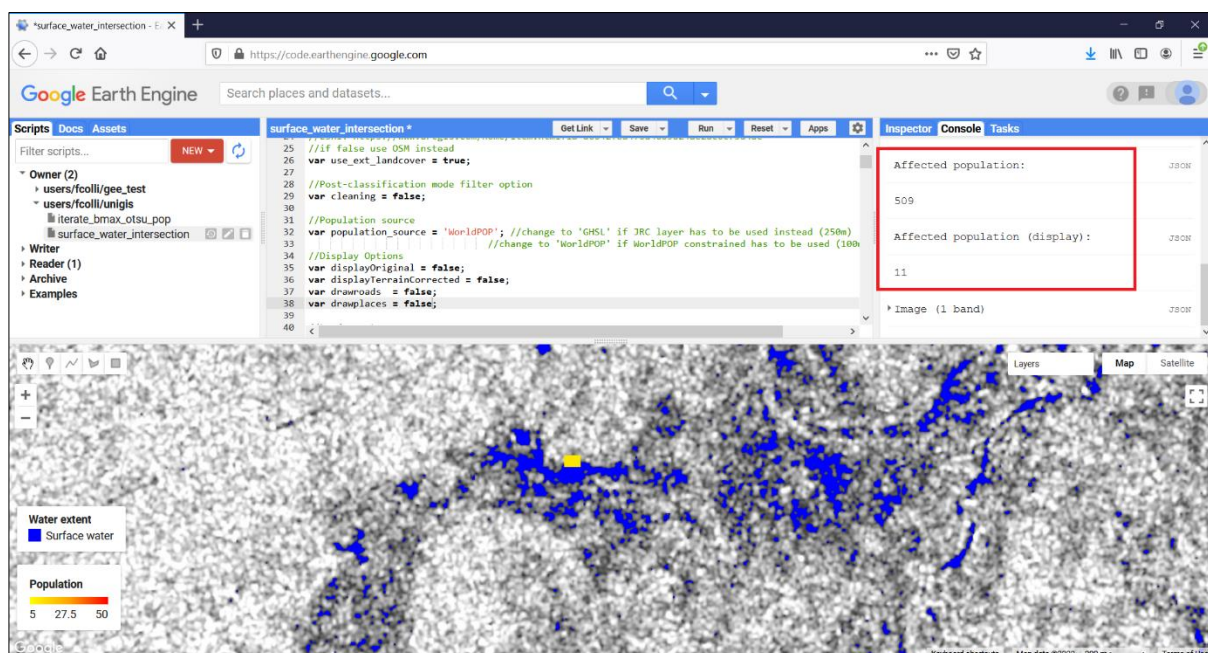


Figure 38. GEE single-date application. When run over a specific region, the console reports the amount of population affected on the whole area of interest and the subset relevant to the current canvas (red box). In the example, for 30.09.2020, 509 people are estimated to live in flooded areas in the whole Pibor, while only 11 are estimated within the boundaries of the current canvas (automatically derived by the application).

VGI intersection

VGI is exploited in this research to demonstrate its possible use in the intersection process with water extent maps. However, main aim of this study is to provide information about surface water itself and population, for which final products are generated and optionally exported for offline use. Therefore, VGI data are used at visualization level, within the web-based interface of GEE.

Particularly, two types of layers are intersected with water surface maps, both gathered from OSM and managed as external asset:

- Roads, having multiline feature type.
- Places (populated settlements), having point feature type.

For roads and beyond visualization, an intersection process actually occurs. In fact, to those segments of the multiline feature which are intersecting with water, a particular rendering style is associated.

To determine the segments of intersection, a conversion of the resampled surface water raster (100 m resolution) to polygon is first performed. A direct intersection using raster data is theoretically possible. However, GEE features native clip, intersect and select routines which operate using exclusively geometry features. The opposite process, rasterization of roads, has been discarded due to the difficult choice of the associated width.

After the polygonization, the intersecting segments from the roads layer are gathered by means of a selection process. Finally, the feature returned are depicted in purple (flooded condition), while the remaining ones are left in green color (good condition). A graphical example of the process carried out is visible in Figure 38.

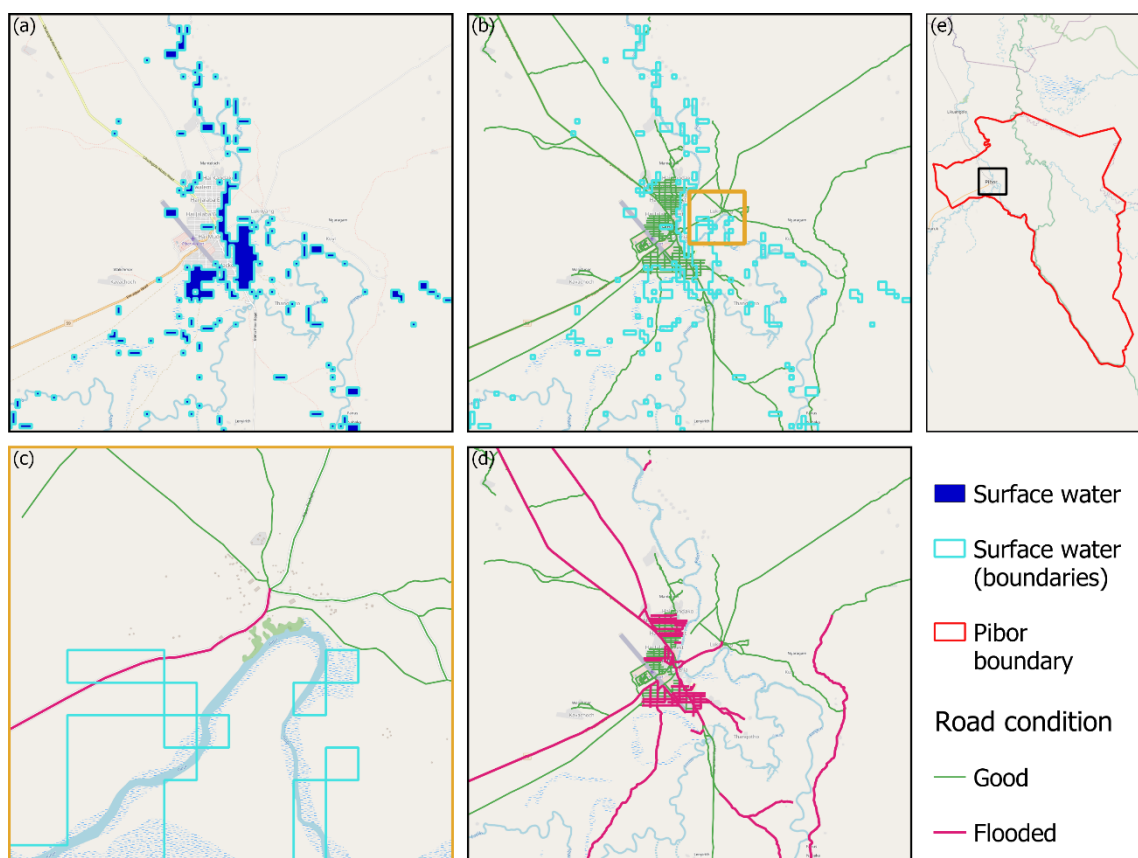


Figure 39. Resampled (100m) surface water extent and polygonised surface water extent (a), original road dataset as gathered from OSM and overlaid polygonised water extent (b), detail of an intersection (c), final result (d) and overview (e). Contains modified Copernicus Sentinel data. Background and roads: © OpenStreetMap contributors.

The usage of the resampled layer instead of the original layer at 10 m resolution is necessary as polygonization processes are using a large amount of memory and result in a massive number of features, which rapidly become unmanageable. Therefore, even though suboptimal from the spatial accuracy point of view, this seems the only strategy to avoid frequent exceptions, given mostly by user memory limit exceeded. For what concerns places, their use is restricted to visualization. These point features are optionally drawn in the GEE canvas, as point features, or using point features in combination with labels.

The selection of features to render is performed on the “fclass” field of the place asset gathered from OSM. Namely, the features having “fclass” equal to “City”, “Suburb”, “Town”, “Village” and “Hamlet”. The relevant labels are taken from the “name” field.

Figure 40 reports an overview of the GEE interface, with places and corresponding labels plotted on the canvas.

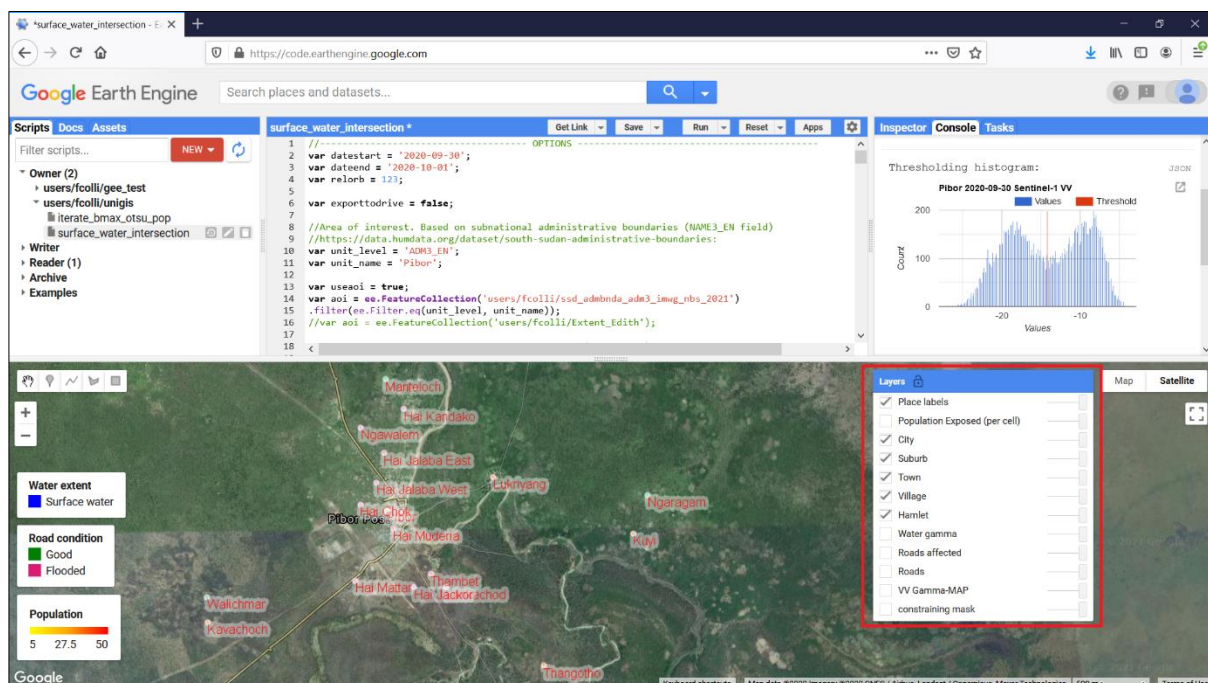


Figure 40. Rendering of places from OSM and relevant labels. All the other layers have been disabled in order to highlight the feature of interest (red box).

The current use of place features is therefore restricted. However, it seemed reasonable to provide the possibility to plot this information, in order to increase the orientation within the maps and to add cartographic features to the web-based canvas. In fact, even though some places are already present and provided by Google when the street map is selected, the area of interest features also unofficial or irregular settlements, which might be important in the humanitarian context. These features are often present in OSM, as they are digitized by volunteers.

Future extensions and research can embed intersection between flooded areas and places, for instance including places of interest like airports, hospitals, or other types. In such case, a complete management of input data and a set of examples is made already available by the applications developed for the purpose of the present study.

Results

In this chapter, the research results are presented. The section is divided into three parts: the beginning part concerns surface maps and relevant population layers. The second one, water occurrence, describes the information derived from the full stack of surface water layers. In the last part, an overview of the developed applications with particular focus on the customizable parameters is given.

Due to the large amount of data processed and results generated, the beginning of the chapter provides graphical results only for the zone of western Pibor (where most of the population live) for the ten dates featuring the largest population impact. Full set of graphics for Pibor Post encompassing the whole observation period and a complete table for the Pibor administrative area are provided in Annex 1.

Tabular results and charts are also presented for the same ten dates considered, calculated across the entire administrative area level 3 of Pibor. Graphs covering the entire time frame are shown.

A restriction applies to the applications part as well. In fact, since the code length is about 900 lines for each, it seems ineffective to report them completely. Therefore, the decision is to describe the main input parameters that may be useful for future use, while the complete code in JavaScript is provided in Annex 2.

Surface water and population

As per each surface water map a layer estimating the population living in flooded areas is computed, the relevant figures extraction is one of the main outcomes of the present research. Table 13 summarizes the first ten date where the multitemporal analysis has shown the greatest amount of population estimated to live in flooded areas. The dates are sorted in time order, from the oldest to the newest.

Table 13. Dates with larger estimated amount of population exposed (first ten).

| Date | Population affected |
|------------|---------------------|
| 21/12/2017 | 662 |
| 27/12/2017 | 701 |
| 24/10/2019 | 491 |
| 05/11/2019 | 644 |
| 17/11/2019 | 450 |
| 22/01/2020 | 399 |
| 21/02/2020 | 379 |
| 04/03/2020 | 535 |
| 30/09/2020 | 509 |
| 05/11/2020 | 400 |

As introduced, the graphical results are given for the dates above over a specific area, shown in Figure 41. In this layout, depicting the original WorldPop dataset, it is clearly visible that the main map area is the one where most of the population lives within Pibor. After that and in the next pages, the products from the ten selected dates are presented. The insets refer to particularly interesting hotspots.

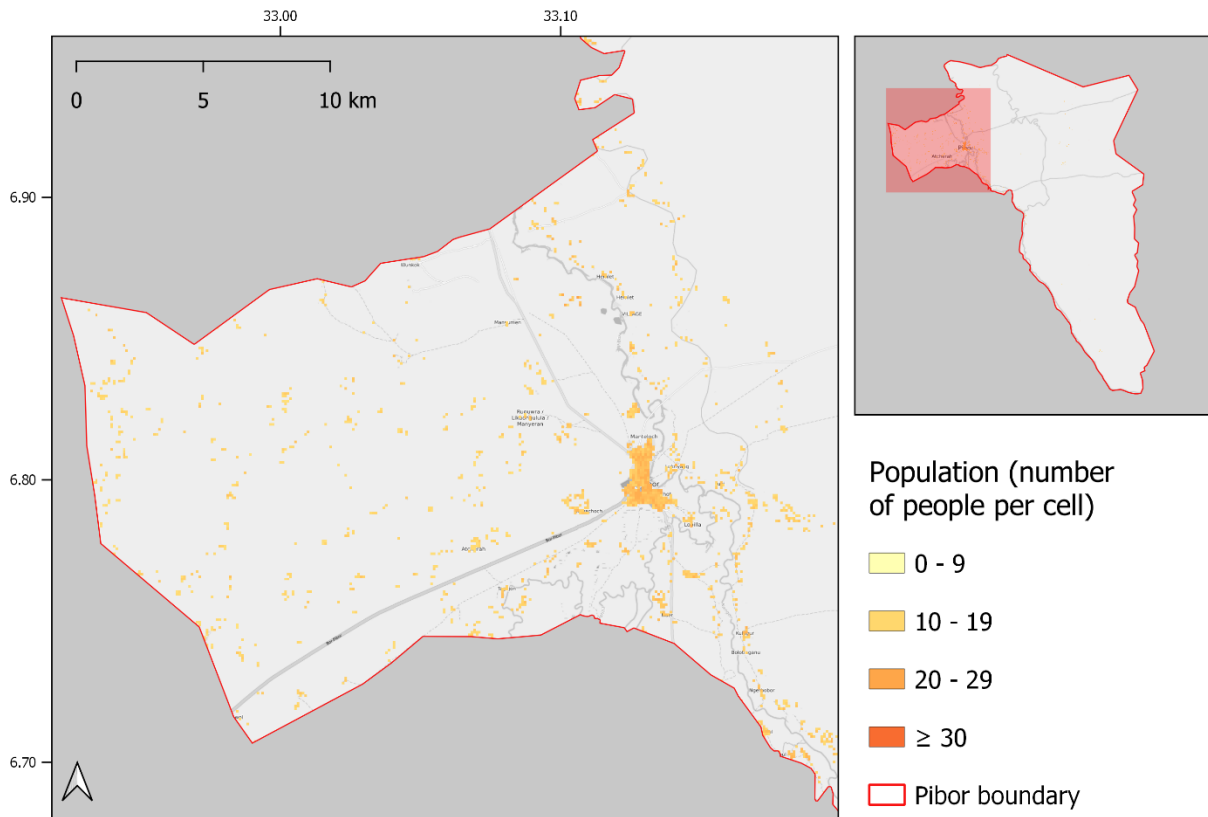


Figure 41. Area selected for the presentation of the graphical results. WorldPop clearly shows that this is the most populated area within Pibor. Population: WorldPop (www.worldpop.org 2020). Background: © OpenStreetMap contributors.

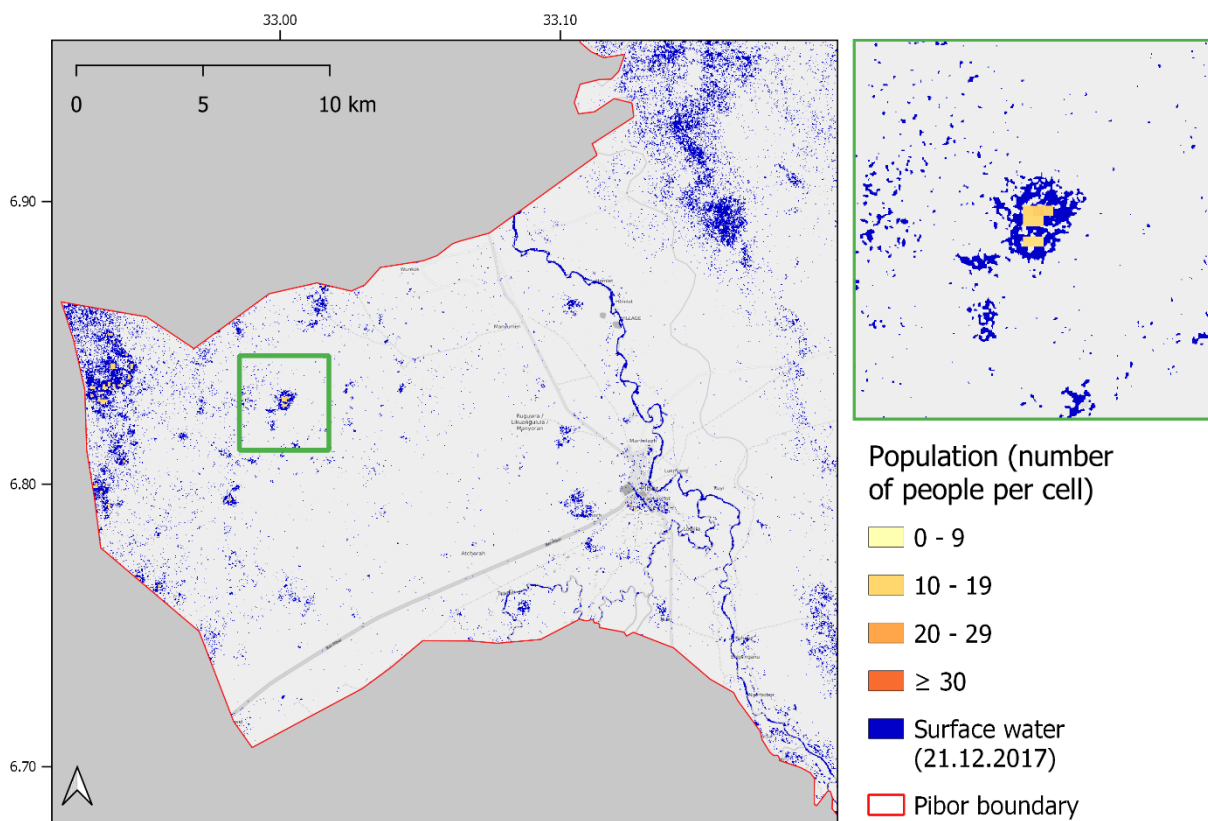


Figure 42. 21.12.2017 products.

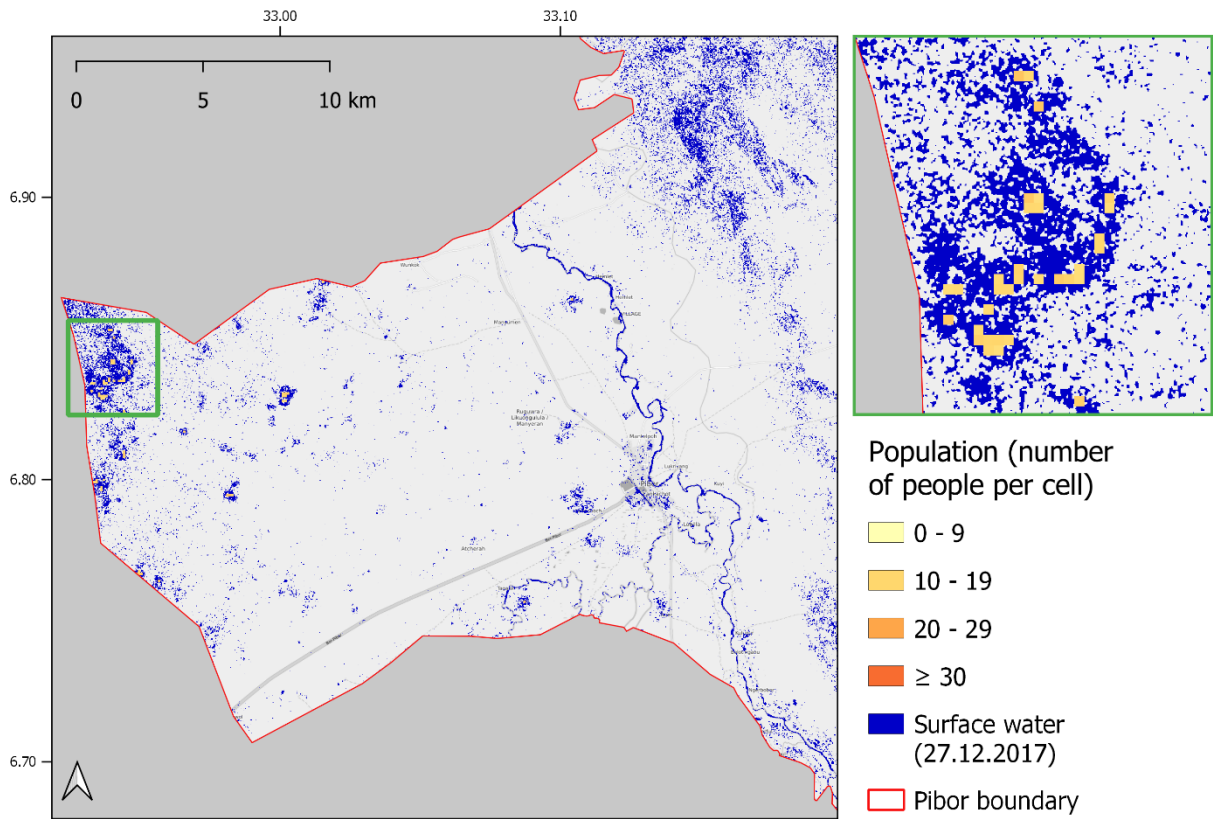


Figure 43. 27.12.2017 products.

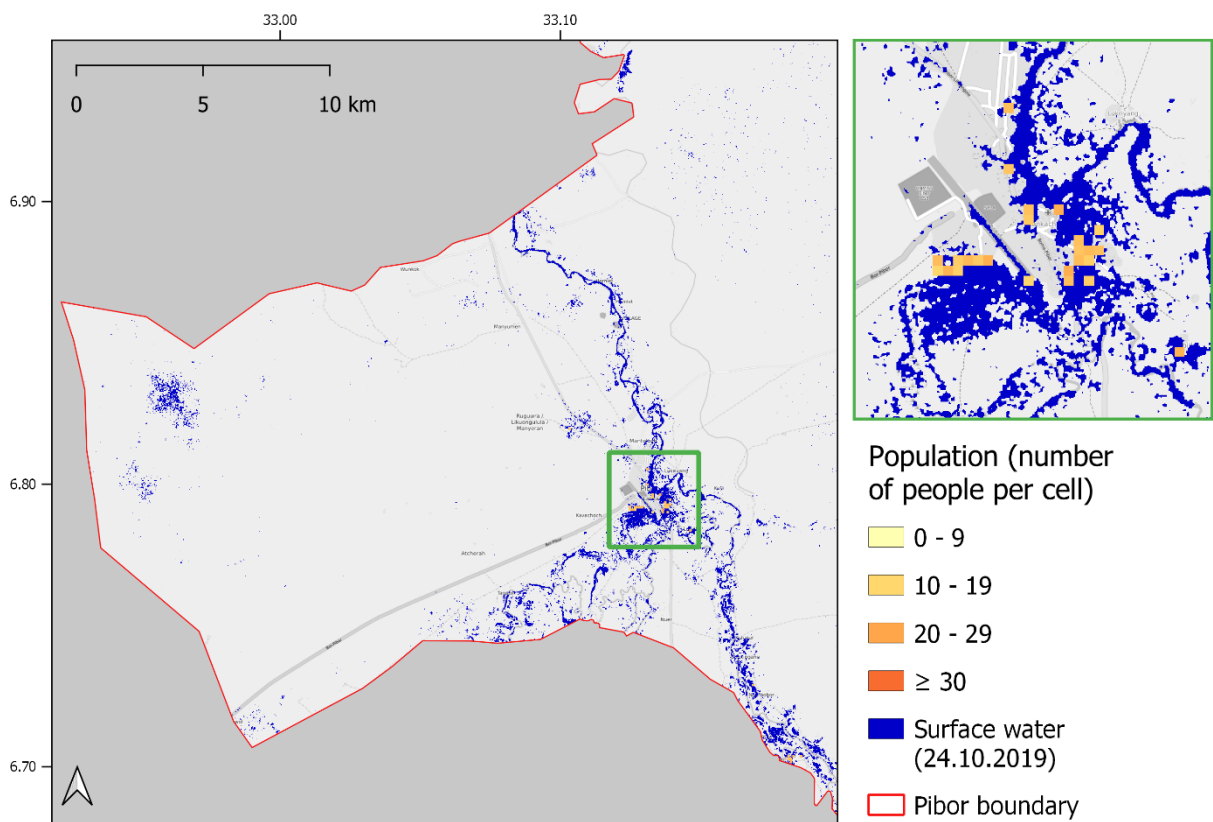


Figure 44. 24.10.2019 products.

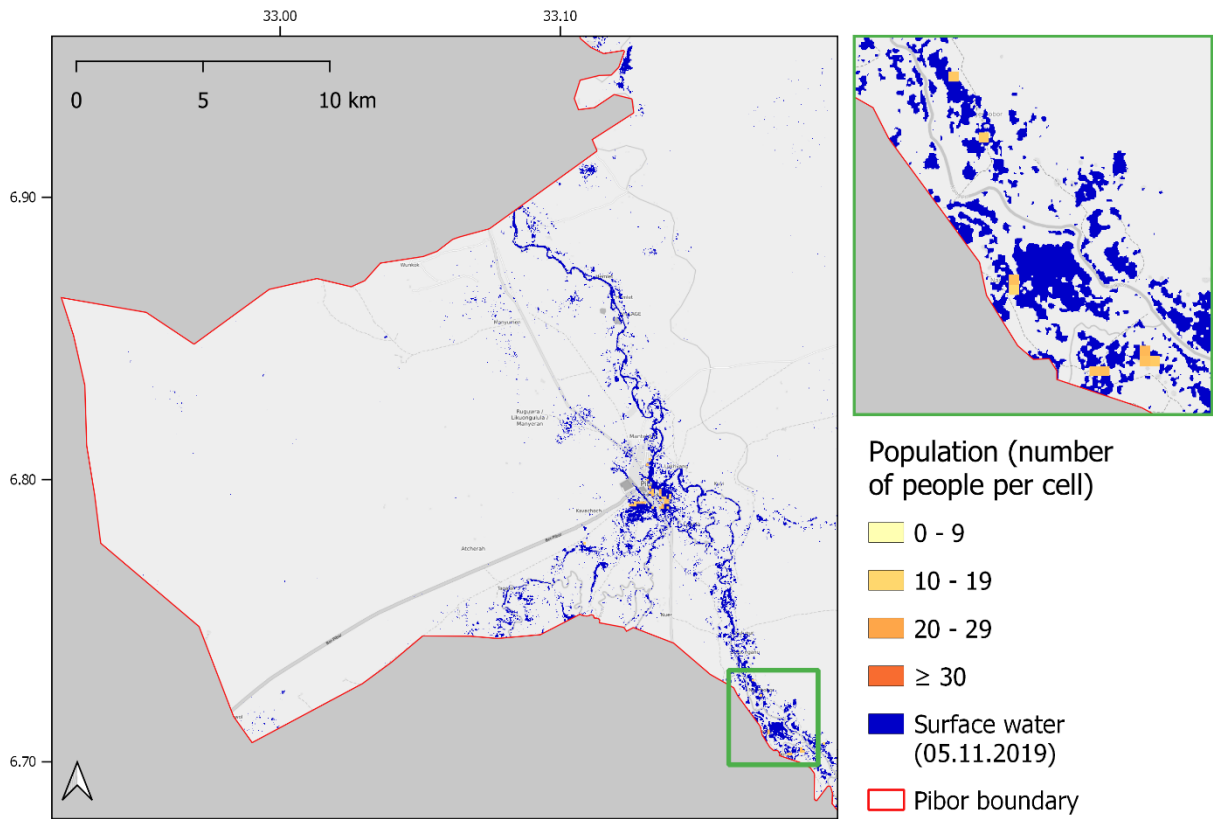


Figure 45. 05.11.2019 products.

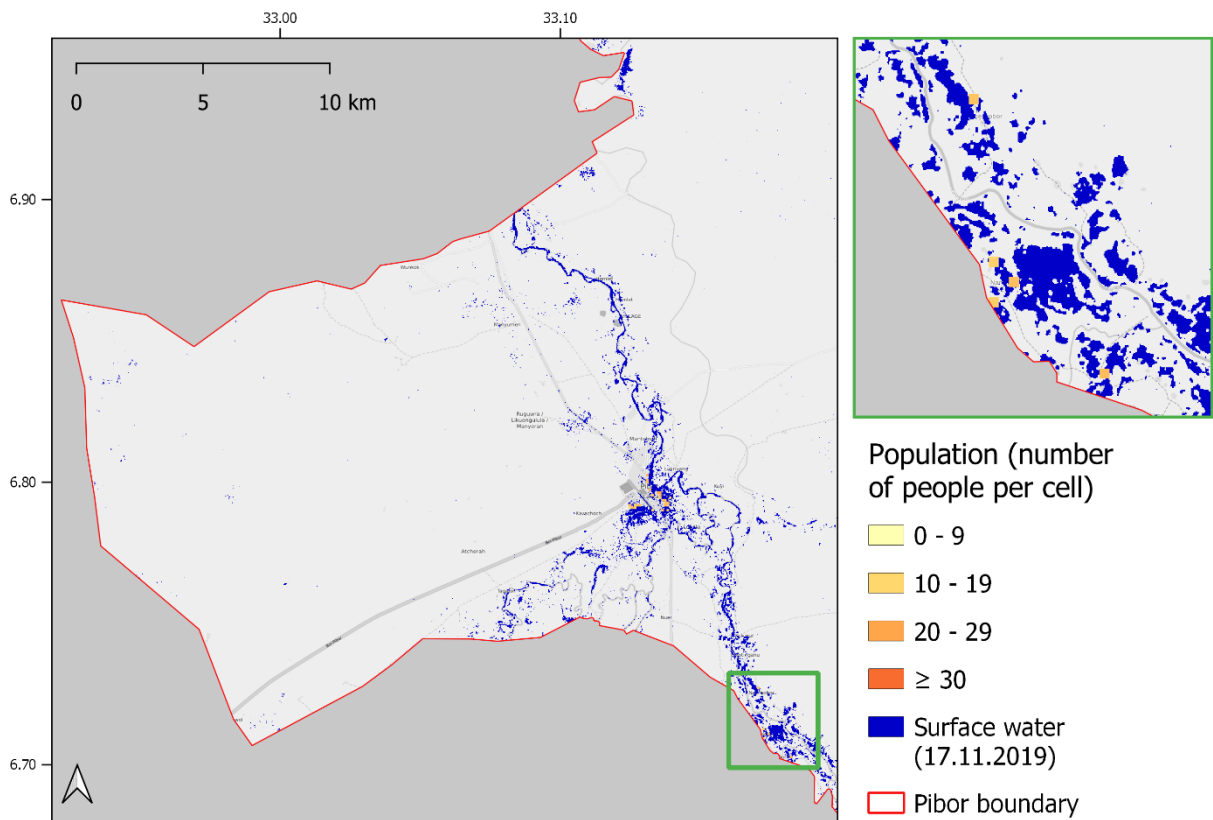


Figure 46. 17.11.2019 products.

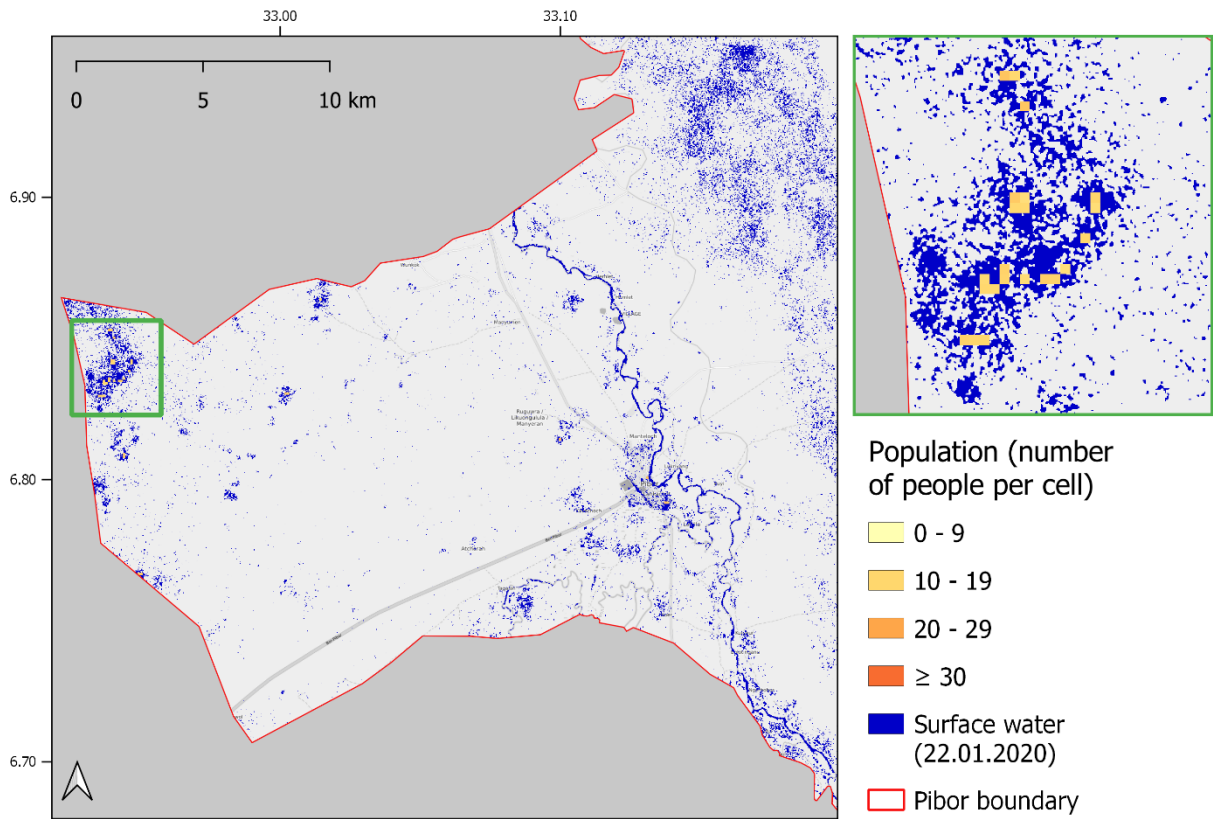


Figure 47. 22.01.2020 products.

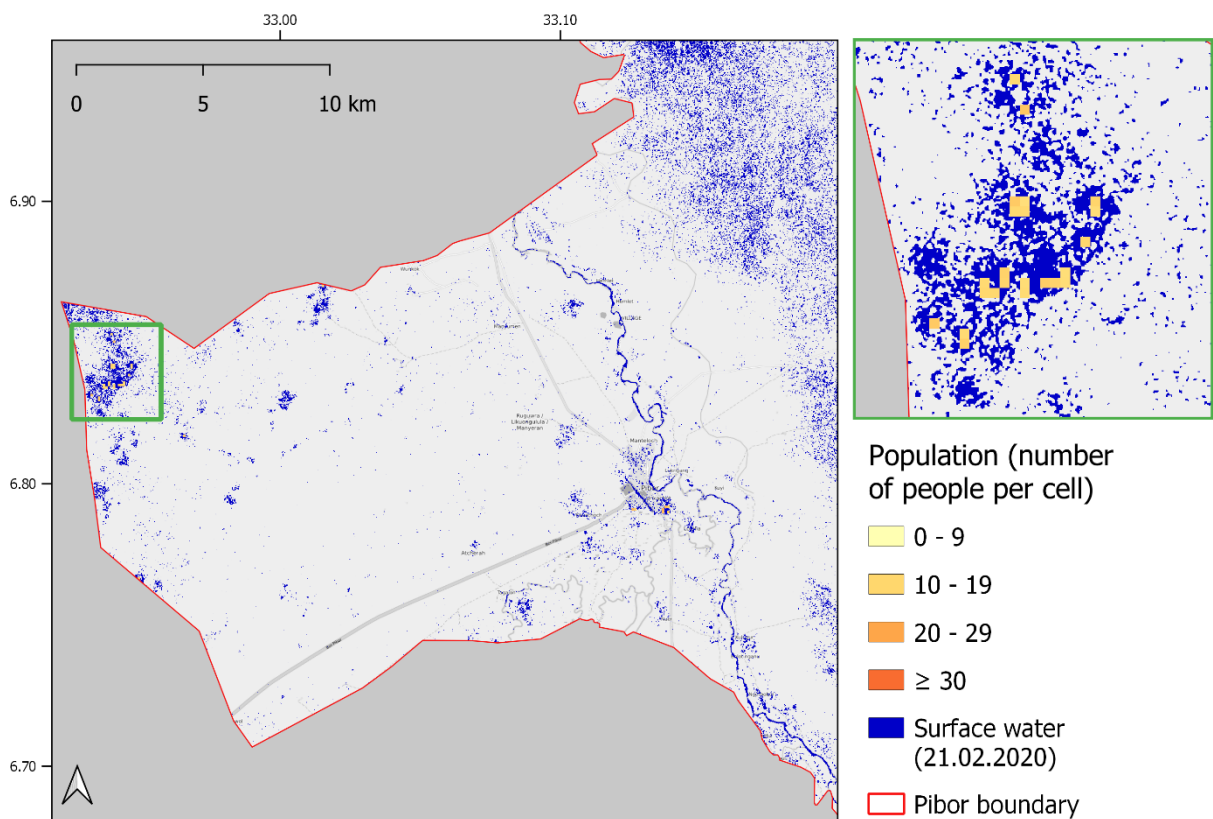


Figure 48. 21.02.2020 products.

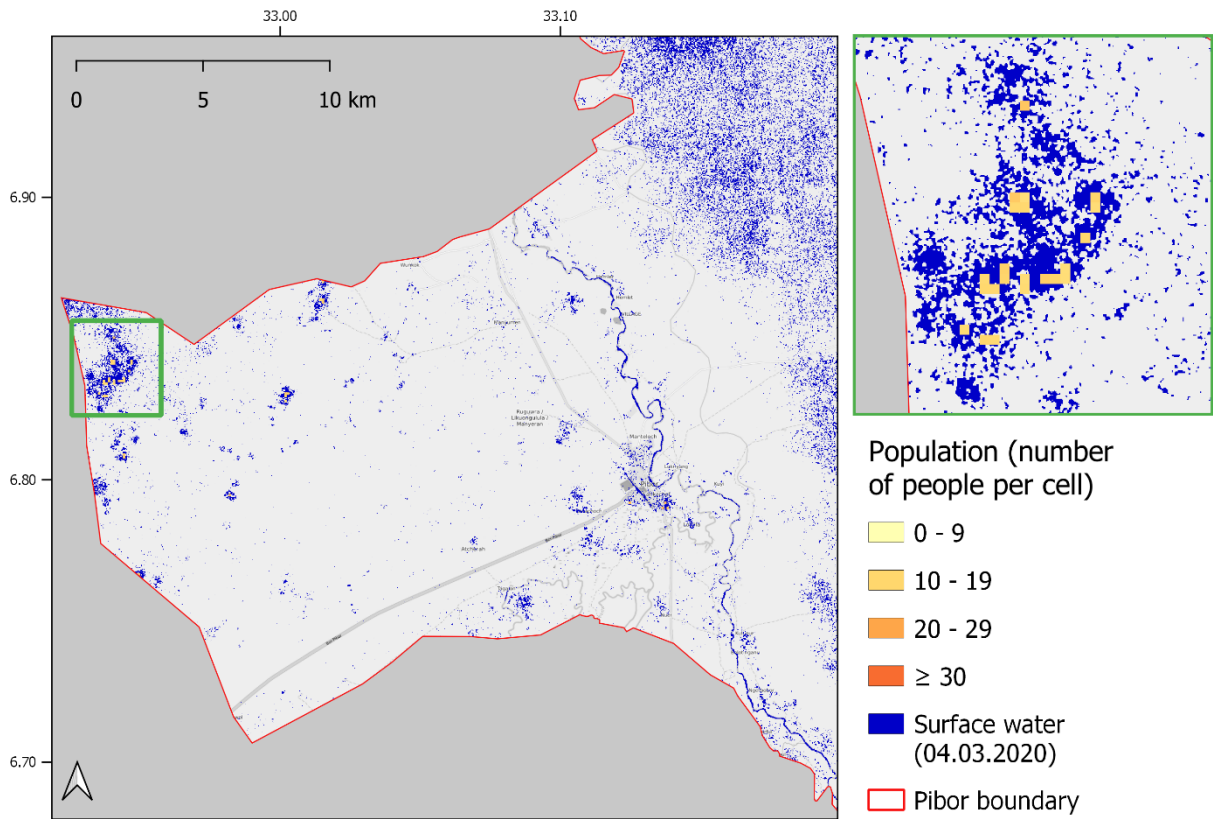


Figure 49. 04.03.2020 products.

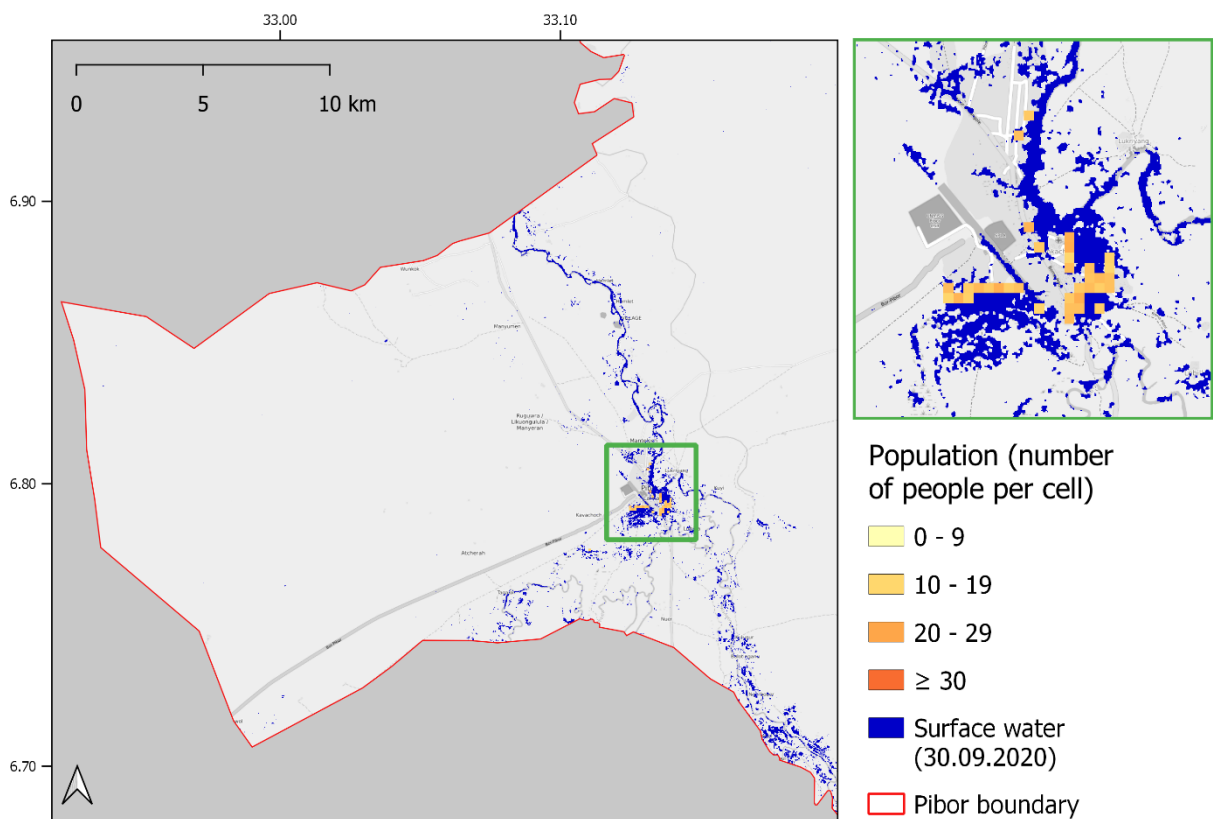


Figure 50. 30.09.2020 products.

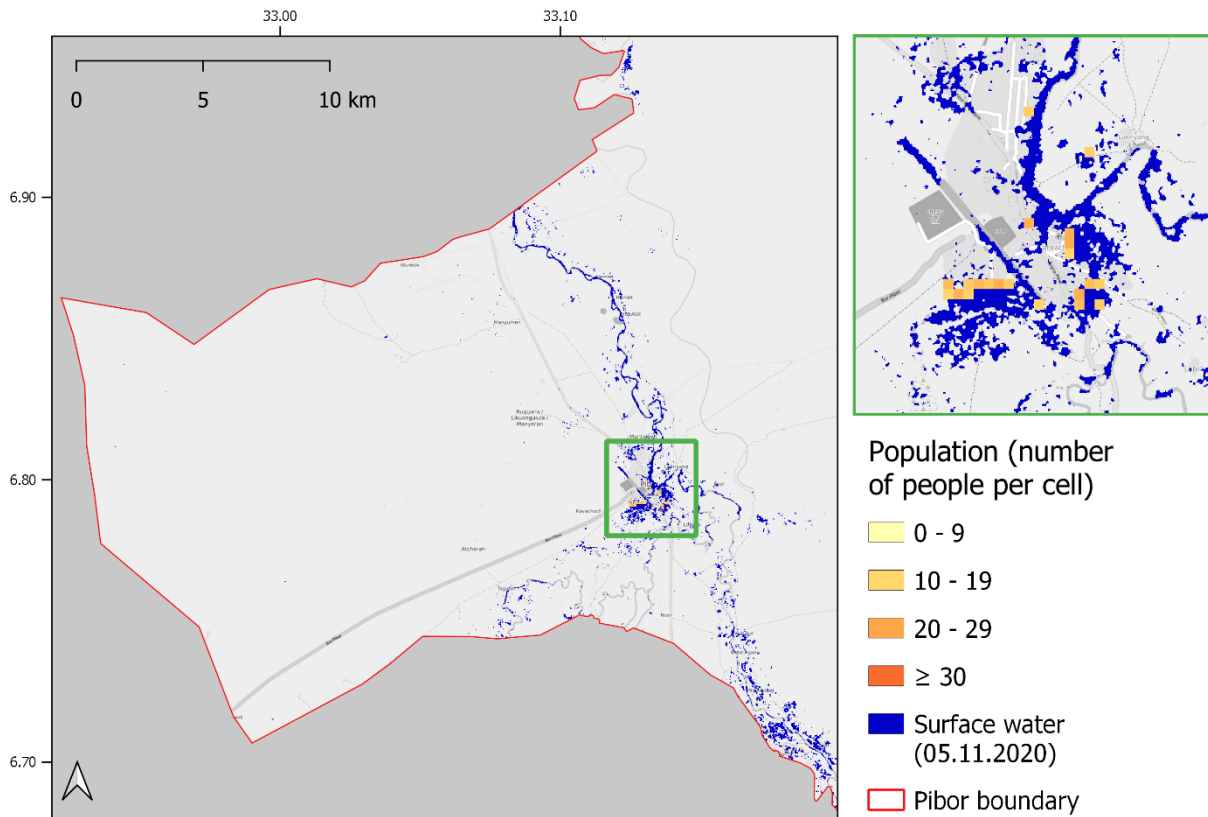


Figure 51. 05.11.2020 products.

Maps in Figure 42, Figure 43, Figure 44, Figure 45, Figure 46, Figure 47, Figure 48, Figure 49, Figure 50 and Figure 51 depict the surface water extent and the population living in flooded areas for the relevant date. All the layouts contain modified Copernicus Sentinel data [2017, 2019, 2020] and show population data from WorldPop (www.worldpop.org). Backgrounds are © of OpenStreetMap contributors.

Beyond the statistics and the graphic reporting the ten most interesting dates, some bar charts are presented hereafter.

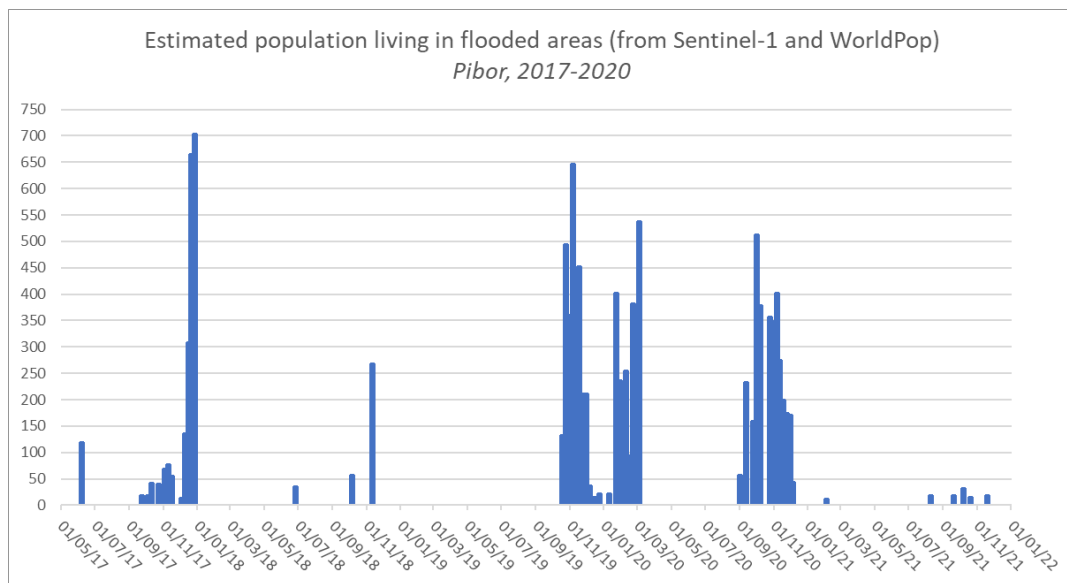


Figure 52. Chart showing the population affected over the whole observation period (evenly spaced horizontally).

Figure 52, reports the entire observation period. Due to the challenges in visualizing long time series with a natural time axis, a compressed chart is also reported in Figure 53. This chart does not feature evenly spaced marks in the horizontal axis. However, it increases the readability of the quantities and helps to isolate groups of dates with consistent population impact.

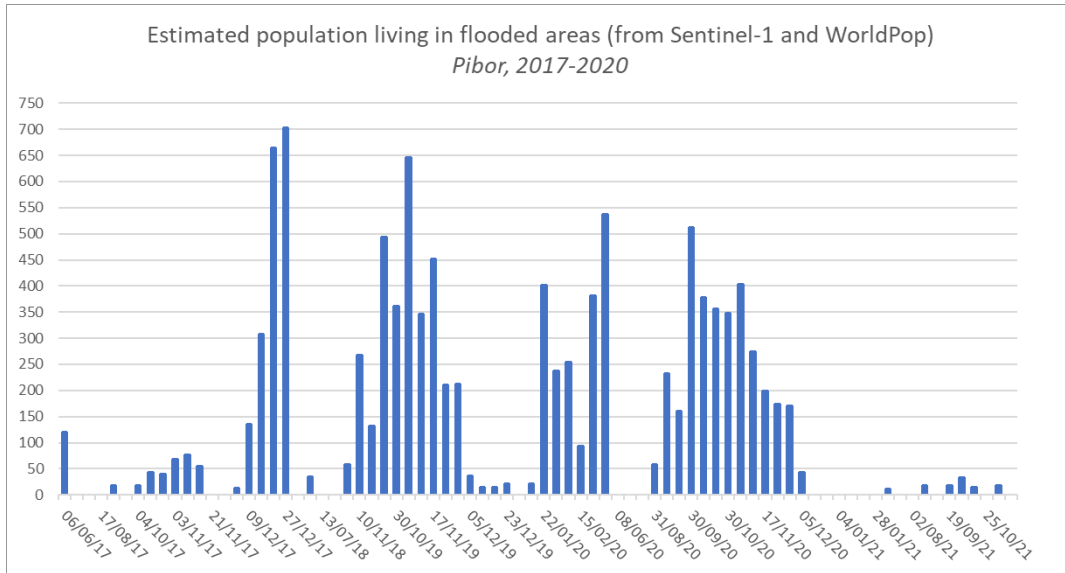


Figure 53. Chart showing the population affected over the whole observation period (compressed horizontally).

From the charts above, becomes clear that 2020 is the year for which the greatest number of surface water maps has been generated. For this reason, a specific chart for this period is shown in Figure 55.

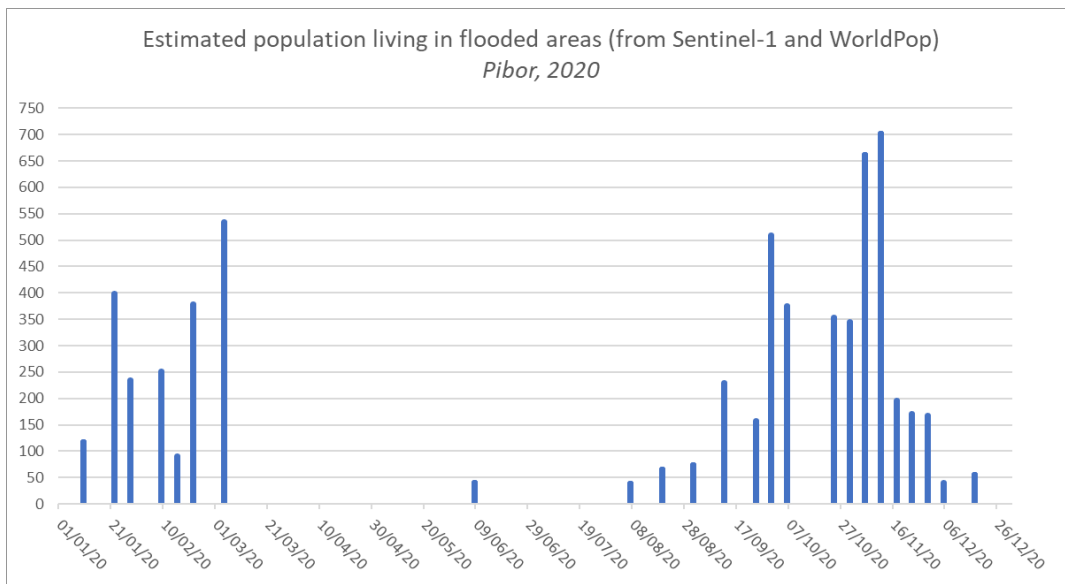


Figure 54. Chart showing the population affected for the year 2020.

Water occurrence

The complete surface water layer dataset presented above provides multi-temporal coverage over multiple years, using the dates in Table 16 and excellent resolution. It follows that it is suitable for the generation of a condensed layer, showing the occurrence of water relative to each pixel. Thus, a product called water occurrence is derived and presented in this part.

Specifically, the layer is generated by considering the percentage of observations in which it is detected as water, out of the entire number of observations. For each pixel, W_{occ} is computed as:

[9]

$$W_{occ} = \frac{W_{obs}}{T_{obs}}$$

Where W_{obs} refer to the water observations and T_{obs} is the total number of valid observations as per the conditions explained in the Bimodality test section of this research.

The resulting layer, that is an exclusive output of the Multitemporal application, is classified using a blue color ramp and looks as depicted in Figure 55.

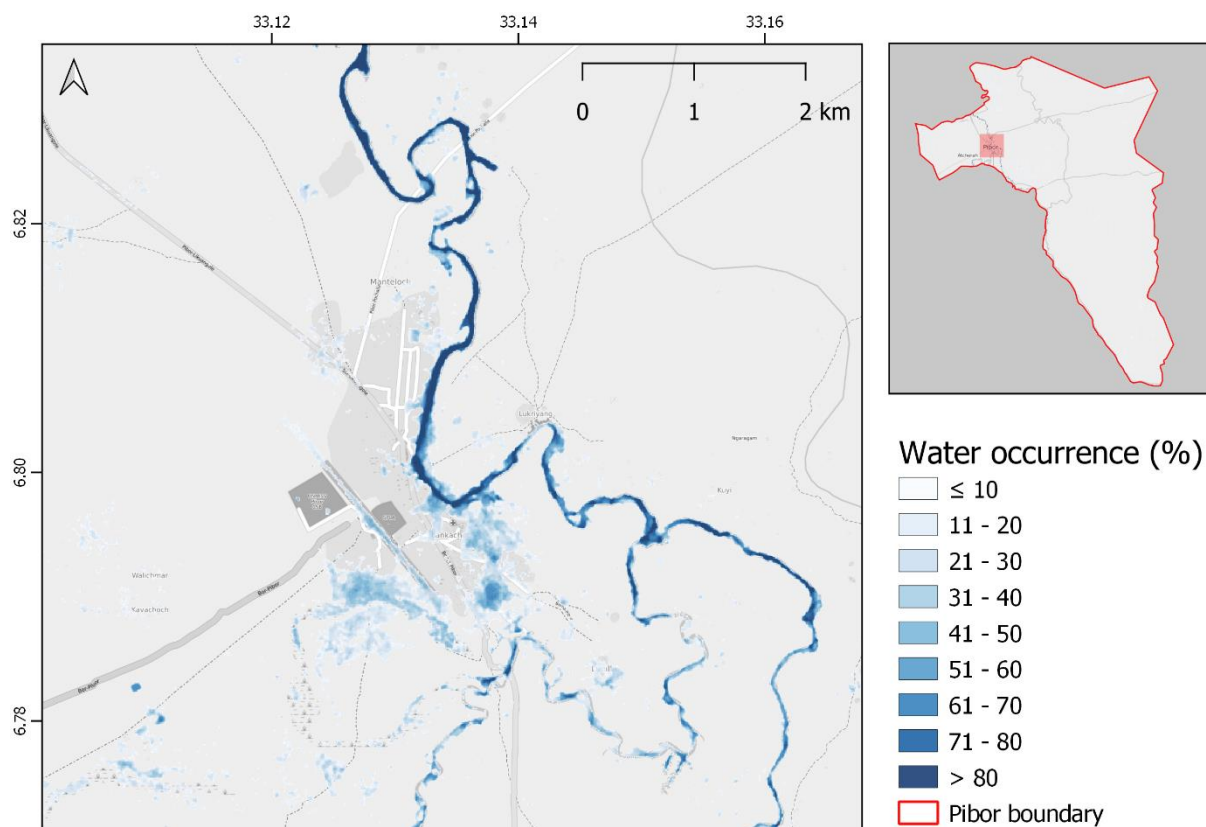


Figure 55. Water occurrence, Pibor. For each pixel, the percentage of water observation is shown. Contains modified Copernicus Sentinel data [2017 -2021]. Background: © OpenStreetMap contributors.

The study area is traversed by the Pibor river and features wetlands and several flood prone areas. For this reason, a map layout showing insets relevant to particularly interesting zones at variable scales is presented in Figure 56.

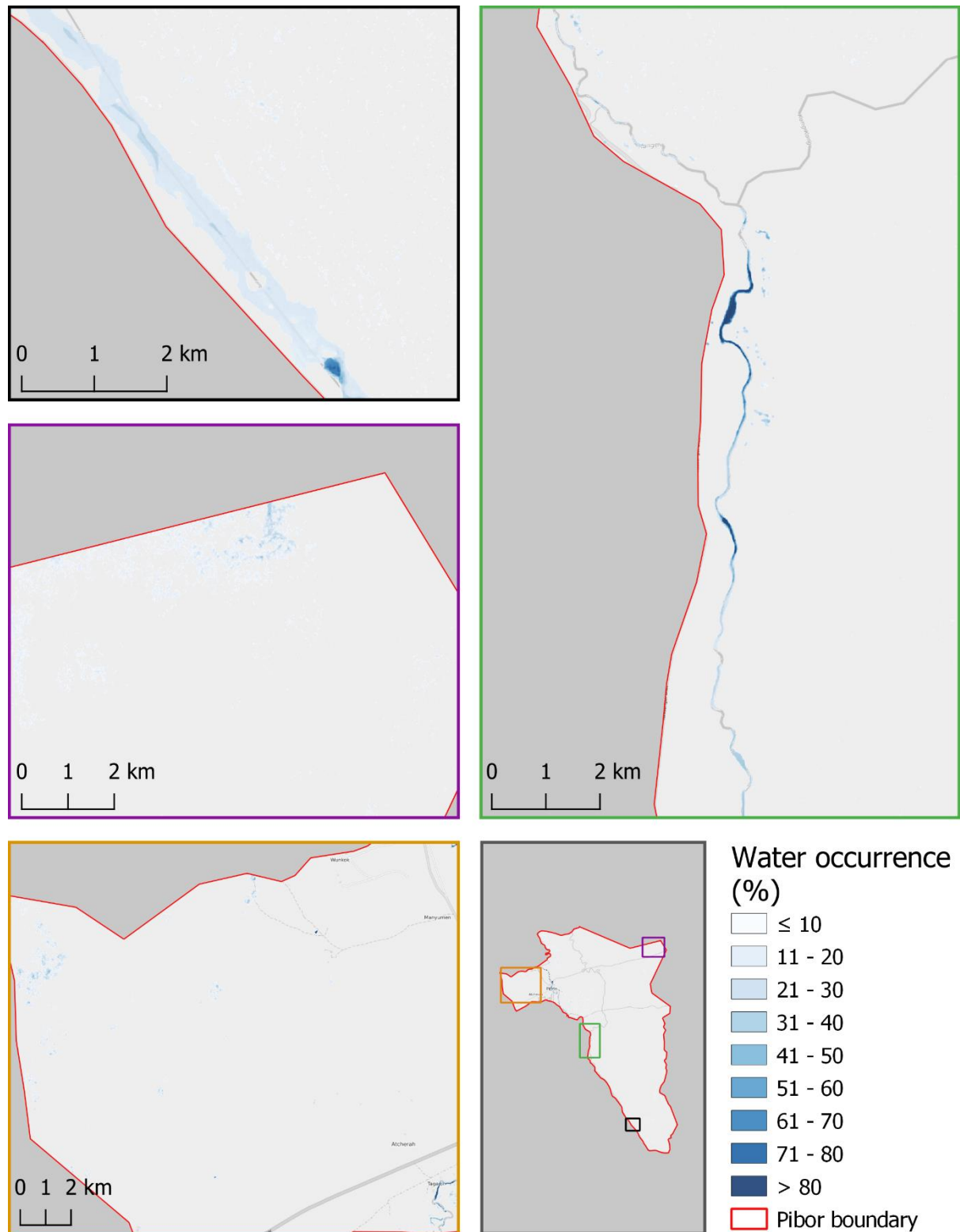


Figure 56. Water occurrence. Different views at different scales. Contains modified Copernicus Sentinel data [2017 -2021]. Background: © OpenStreetMap contributors.

From the presented layer, appears clear that some river zones do not show high occurrences all the time. This happens for several reasons. First, there are some places where the riverbed is so narrow that the resolution does not allow a proper detection. This is especially true during the dry season, when the water level is so shallow that the effect of the neighbouring vegetation, combined with the slight blurriness introduced by the de-speckle filtering increases the backscatter over the the river itself. Second, some parts of the riverbed might completely dry in some dates. Third, the classification might be inaccurate in some locations for specific dates.

However, the main purpose of this product is to integrate and improve data of existing solutions. In fact, the occurrence map is similar to the one found in Global Surface Water Explorer (EC - JRC, 2020). A visual comparison between the two products over two locations is shown in Figure 57.

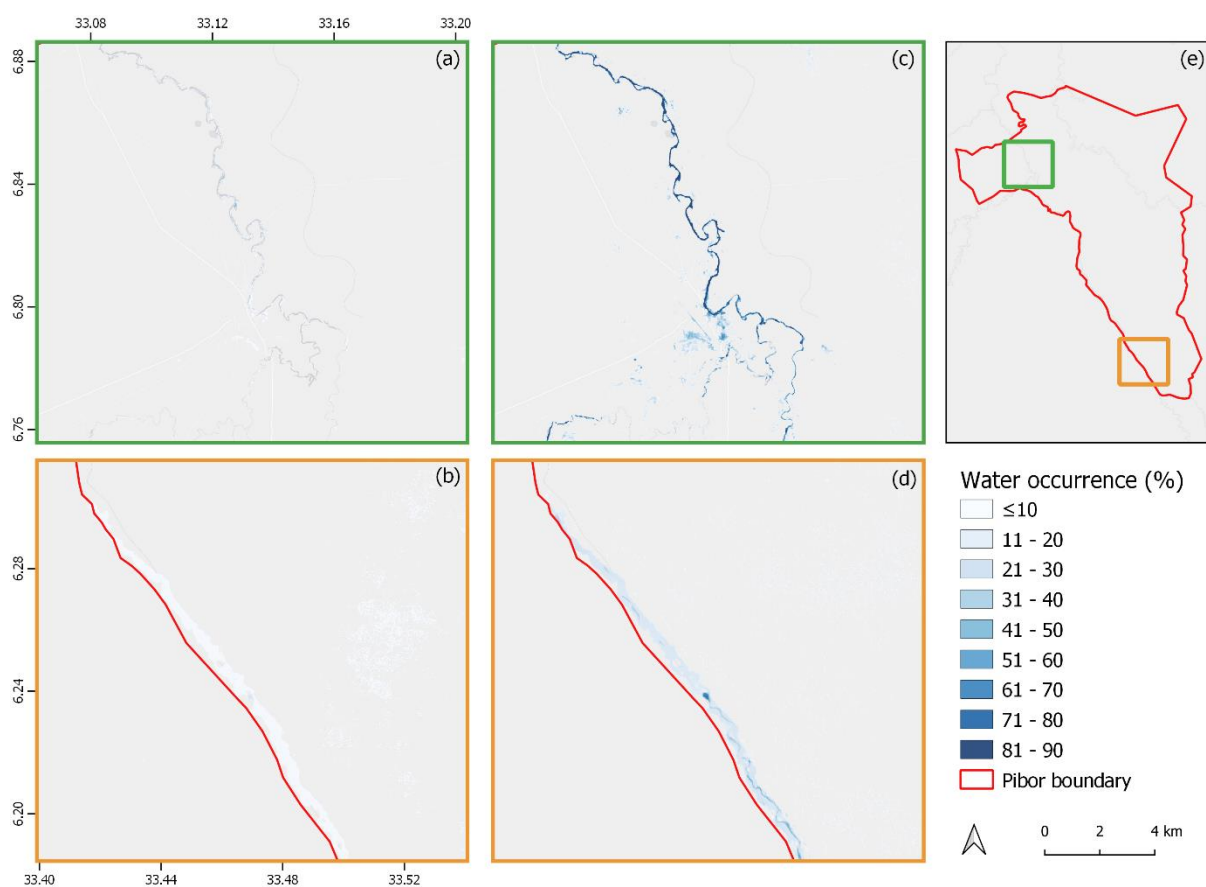


Figure 57. Water occurrence: EC-JRC (a and b) vs layer derived in this research (c and d). a and b source: EC-JRC/Google. c and d: contains modified Copernicus Sentinel data [2017 -2021]. Background: © OpenStreetMap contributors.

Even without appropriate validation of the product, which is beyond the scope of this research, it seems clear that the existing EC-JRC map is underestimated in some areas. In addition, the spatial resolution of the map hosted by Global Surface Water explorer is 30 m, versus 10 m for the data produced. The layer water occurrence therefore proves to be a viable alternative or interesting as a complementary use alongside existing solutions.

Applications

As extensively introduced, all results obtained by applying the methodology of this research are obtained using Google Earth Engine as processing and export tool. Particularly, two applications have been developed: one analyzing single date data and one considering multitemporal data.

The two applications implemented for the purpose, however, are also to be considered partially integral to the result. In fact, their use allows to satisfy one of the objectives of this research, namely, to make the methodology applicable in different areas. In addition, they are a starting point for further extensions and refinement of the methodology. Both applications make use of several SAR-specific and general image processing routines, not described in detail and for which the full code is available in the Utility and processing functions section of Annex 2.

An overview of the main customizable parameters of the applications is given in this section, with the aim to provide a sort of documentation to facilitate possible future users or researchers. Two parts are used to fulfill the documentation requirements: the first concerning the single date application and the second describing the multitemporal one.

Single date

The single date application allows to extract a surface water layer from a Sentinel-1 acquisition and to intersect it with population and VGI. The two layers (surface water and involved population) can be fully exported at native resolution for offline use. VGI information and figures are displayed in the console. The input parameters, tunable from the header of the application, are summarized in Table 14.

Table 14. Single date application main parameters.

| Name | Type | Default | Notes |
|-------------------------|-------------------|--------------------------------|---|
| datestart | string | "2020-09-30" | Date of the analysis. |
| dateend | string | "2020-10-01" | Date of the analysis + 1 day. |
| relorb | integer | 123 | Sentinel-1 relative orbit. |
| exporttodrive | boolean | true | Export the results to Gdrive. |
| unit_level | string | "ADM3_EN" | Administrative level. |
| unit_name | string | "Pibor" | Name of the administrative unit. |
| useaoi | boolean | true | Use AOI specified above or process the whole Sentinel-1 image. |
| aoi | FeatureCollection | ssd_admbnda_adm3_imwg_nbs_2021 | Asset containing the administrative areas. |
| otsu_method | boolean | true | Apply adaptive thresholding. If false, a crisp threshold (-15) is applied. |
| otsu_type | string | "constrained" | "constrained" refers to the "adaptive-constrained" thresholding in this research. |
| use_ext_landcover | boolean | true | Use an external land cover instead of a specified asset. By default, ESRI landcover is used. The dataset source can be changed in the application code. |
| cleaning | boolean | false | Apply a morphological filter (experimental). |
| population_source | string | "WorldPop" | Population data source. Other value is "GHSL". |
| displayOriginal | boolean | false | Add the original Sentinel-1 acquisition to the map canvas. |
| displayTerrainCorrected | boolean | false | Add the slope corrected Sentinel-1 acquisition to the map canvas. |

| | | | |
|------------|---------|-------|--|
| drawroads | boolean | true | Add roads (from VGI) in the map canvas. |
| drawplaces | boolean | true | Add places (from VGI) in the map canvas. |
| addlabels | boolean | false | Add place labels to the map canvas. |

The full code in JavaScript, quite detailed commented, can be found in the Single date application part of Annex 2.

Multitemporal

In the multitemporal application, embedding by default only the methodology of adaptive-constrained thresholding, less parameters can be tuned. In fact, its use is more intended for intensive processing rather than for research purpose. For this reason, there are three main differences compared to the single date application.

First, results are by default exported only in tabular form. However, the spatial layers computed for the encompassed period can be sent to the export queue by changing the relevant parameter. This results in a long series of tasks. Second, the VGI data is not used as limited to visualization and intersected with a single date only. Finally, a water occurrence layer, obviously missing in the first application, can be optionally visualized.

A detailed list of the customizable parameters is given by Table 15.

Table 15. Multitemporal application main parameters.

| Name | Type | Default | Notes |
|------------------|-------------------|--------------------------------|---|
| datestart | ee.Date | "2021-08-01" | Start date of the analysis. |
| dateend | ee.Date | "2021-12-31" | End date of the analysis. |
| relorb | integer | 123 | Sentinel-1 relative orbit. If set to 'none' all the relative orbits intersecting the area of interest are considered. |
| datetoshow | string | "none" | The date specified, within the datestart-dateend range, is visualized in the canvas. |
| showsum | boolean | false | If change to true, the water occurrence layer is visualized in the canvas. |
| exporttodrive | boolean | true | Export the results in tabular form to Gdrive. |
| storeexportbatch | boolean | false | Exports the full results (spatial layers) to Gdrive. |
| bmxtresh | float | 0.75 | Threshold used for the bimodality test. |
| unit_level | string | "ADM3_EN" | Administrative level. |
| unit_name | string | "Pibor" | Name of the administrative unit. |
| useaoi | boolean | true | Use AOI specified above or process the whole Sentinel-1 image. |
| aoi | FeatureCollection | ssd_admbnda_adm3_imwg_nbs_2021 | Asset containing the administrative areas. |

The full code of the application can be found in the Multitemporal application part of Annex 2.

Discussion

The outcome of this research primarily consists in the setup of a robust methodology to extract surface water maps from remote sensing data and in its demonstration by means of the generation of a multitemporal series of products. The set is derived making exclusive use of free of charge SAR data and auxiliary layers, indicating the capability to operate irrespective from hour of the day and weather present at the time of need.

The methodology designed is based on a well-known image thresholding technique (Otsu, 1979) and it is inspired by existing approaches (Markert et al., 2020) that do not meet the applicability requirements for the area of interest. Oppositely, the adaptive-constrained thresholding presented exploits sampling on precise subsets, delineated by an auxiliary layer in order to increase effectiveness over particularly challenging areas like South Sudan.

The statistics computed show particular success in the application of the methodology and its embedded tests for automatic threshold calculation for SAR images. The cross-validation performed over the most interesting part of the study area shows an overall accuracy of more than 88% for the two classes of interest (water/other).

The 78 surface water layers enable the generation of the same number of products estimating the number of people living in flooded areas. In addition, a water occurrence map is derived. These results fulfill two important objectives: first, to provide information about possibly displaced people (Lang et al., 2018). Second, to propose another source of data enhancing existing products (Pekel et al., 2016). The two objectives fulfilled are framed in the domains of humanitarian action for the first one and in that of risk analysis for the second one.

In quantitative terms, the results confirm the dramatic involvement of the study area in catastrophic flooding events in recent years (MSF, 2019, MSF, 2020, MSF, 2021, MSF, 2022). Especially for 2019 and 2020, in fact, a large number of surface water maps are generated with associated indication of high number of possible people involved. However, the date for which the largest number of people were affected appears to belong to the year 2017. In this year, on 27.12.2017 a total number of 701 people is estimated to live in flooded areas. In addition, according to the water occurrence layer, 48% of the land is flood prone. The water occurrence has a promising effectiveness as integration or replacement of existing incomplete solutions.

As the possible surface water layers intersection with other data source is demonstrated, the products can be efficiently used to assess the impact on infrastructures and places, of critical importance in the planning and response activities of NGOs such as MSF (Lang et al., 2018). For this information, however, no quantitative assessment is performed.

Important outcome of the research are the JavaScript applications for GEE. Their use enables a time-relevant and reliable information extraction, with near to no human intervention. GEE applications feature a high-power and large storage infrastructure exploitation (Gorelick et al., 2017), allowing the user to save local resources. The large possibility of parameters customization offer, in addition, a large portability over different areas beyond the one targeted in the present research.

Conclusions and outlook

This research highlights the effectiveness of using remote sensing data acquired in the microwave spectrum in the extraction of surface water maps, particularly in South Sudan. In fact, the methodology proposed using Sentinel-1 data is specifically designed to operate independently of weather conditions present, making it operational during the rainy season, when and where flood events are more likely to occur.

The generated set of layers demonstrate a multi-year applicability and satisfactory robustness. However, more investigation would be ideal regarding accuracy, that is promising from the first assessment, since with current data a cross-validation is possible for few dates.

Concerning the methodology itself, future expansions may concern downstream filtering at the classification step, as the maps sometimes appear to be affected by residual noise. These steps, however, are to be considered with great caution, to avoid missing the advantage offered by the excellent resolution of the data taken into consideration.

The temporal frequency of the surface water layers, still limited to a repetition time of 12 or 6 days depending on the Sentinel-1 schedule, is in perspective of increase, as the Copernicus program foresees the launch of two additional satellites with SAR sensors as payload (Sentinel-1C and Sentinel-1D).

Regarding other data sources, the use of optical images is mostly excluded. The analysis of the Sentinel-2 archive, in fact, confirms the scarce existence of cloud free images for the main periods of need. In any case, the synergistic use of SAR and optical should be favored in the few cases where user need and availability of optical data match.

As the use case demonstrations highlight, the demographic data analyzed show some shortcomings, such as the resolution, which is much lower than in remote sensing data, and the lack of multi-temporal series. Beside these limitations, they remain a good proxy indicator for the estimation of possible displaced people and the use of WorldPop seems to be the best current choice in this sense.

Water occurrence, derived as side product by condensing the surface water maps into a single layer, allows to detect the most flood prone areas, in qualitative and quantitative terms. Oppositely, the intersected VGI are limited to visualization in a web interface. However, their use is possible, and they can already be evaluated with the proposed pipelines.

The developed applications offer a useful tool to support the work of NGOs like MSF. Their release is a starting point for the refinement of the presented methodology and for its further expansion. A first refinement could for instance concern the consideration of accessibility to places, or the estimation of people displaced within a buffer.

In conclusion, this research proposes a user-friendly solution ready for further study and product generation, which can bring both current benefit and be of support and inspiration for further information delivery systems in the context of humanitarian action.

References

- ACHARYA, T. D., LEE, D. H., YANG, I. T. & LEE, J. K. 2016. Identification of water bodies in a Landsat 8 OLI image using a J48 decision tree. *Sensors*, 16, 1075.
- ALEGANA, V. A., ATKINSON, P. M., PEZZULO, C., SORICHETTA, A., WEISS, D., BIRD, T., ERBACH-SCHOENBERG, E. & TATEM, A. J. 2015. Fine resolution mapping of population age-structures for health and development applications. *J R Soc Interface*, 12.
- ASIAN DISASTER PREPAREDNESS CENTER - NASA JPL. 2022a. *Historical Flood Analysis Tool* [Online]. Available: <https://servir.adpc.net/tools/historical-flood-analysis-tool> [Accessed 27.01 2022].
- ASIAN DISASTER PREPAREDNESS CENTER - NASA JPL. 2022b. *SERVIR Mekong* [Online]. Available: <https://servir.adpc.net/> [Accessed 27.01 2022].
- BELSPO. 2014. *MAUPP: Modeling and forecasting African Urban Population Patterns* [Online]. Available: <https://maupp.ulb.ac.be/> [Accessed 25.01 2022].
- BENNETT, J. 2010. *OpenStreetMap*, Packt Publishing Ltd.
- BIORESITA, F., PUISSANT, A., STUMPF, A. & MALET, J.-P. 2018. A Method for Automatic and Rapid Mapping of Water Surfaces from Sentinel-1 Imagery. *Remote Sensing*, 10.
- BJORGO, E. 2001. Supporting humanitarian relief operations. *RAND-PUBLICATIONS-MR-ALL SERIES-*, 403-432.
- BRAUN, A. & HOCHSCHILD, V. 2017. Potential and Limitations of Radar Remote Sensing for Humanitarian Operations. *GI Forum*, 1, 228-243.
- BRAUN, A., LANG, S. & HOCHSCHILD, V. 2016. Impact of refugee camps on their environment a case study using multi-temporal SAR data. *Journal of Geography, Environment and Earth Science International*, 4, 1-17.
- CANNY, J. 1986. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, 679-698.
- CAO, H., ZHANG, H., WANG, C. & ZHANG, B. 2019. Operational Flood Detection Using Sentinel-1 SAR Data over Large Areas. *Water*, 11.
- CGIAR. 2018. *SRTM DEM Digital Elevation Database* [Online]. Available: <https://srtm.csi.cgiar.org/> [Accessed 25.01 2022].
- CHECCHI, F., STEWART, B. T., PALMER, J. J. & GRUNDY, C. 2013. Validity and feasibility of a satellite imagery-based method for rapid estimation of displaced populations. *International journal of health geographics*, 12, 1-12.
- COMMITTEE ON EARTH OBSERVATION SATELLITES (CEOS). 2021. *Analysis Ready Data For Land (CARD4L)* [Online]. Available: https://ceos.org/ard/files/PFS/NRB/v5.0/CARD4L-PFS_Normalised_Radar_Backscatter-v5.0.pdf [Accessed 06.01 2022].
- CONGALTON, R. G. & GREEN, K. 2019. *Assessing the accuracy of remotely sensed data: principles and practices*, CRC press.
- DE GRANDI, G., LEYSEN, M., LEE, J. & SCHULER, D. Radar reflectivity estimation using multiple SAR scenes of the same target: technique and applications. *IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing-A Scientific Vision for Sustainable Development, 1997. IEEE*, 1047-1050.
- DEMIRKAYA, O. & H. ASYALI, M. 2004. Determination of image bimodality thresholds for different intensity distributions. *Signal Processing: Image Communication*, 19, 507-516.

- DEVILLE, P., LINARD, C., MARTIN, S., GILBERT, M., STEVENS, F. R., GAUGHAN, A. E., BLONDEL, V. D. & TATEM, A. J. 2014. Dynamic population mapping using mobile phone data. *Proc Natl Acad Sci U S A*, 111, 15888-93.
- DU, Y., ZHANG, Y., LING, F., WANG, Q., LI, W. & LI, X. 2016. Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sensing*, 8, 354.
- EC - JRC. 2020. *Global Surface Water - Data Access* [Online]. Available: <https://global-surface-water.appspot.com/download> [Accessed 26.01 2022].
- ESA. 1998. *The Copernicus programme* [Online]. Available: https://www.esa.int/Applications/Observing_the_Earth/Copernicus [Accessed 14.01 2022].
- ESA. 2014. *The Sentinel-1 Toolbox*, [Online]. Available: <https://sentinel.esa.int/web/sentinel/toolboxes/sentinel-1> [Accessed 06.01 2022].
- ESRI. 2020. *ESRI 10-Meter Land Cover* [Online]. Available: <https://livingatlas.arcgis.com/landcover/> [Accessed 04.01 2022].
- EUROPEAN SPACE AGENCY (ESA). 2014. *Sentinel online* [Online]. Available: <https://sentinels.copernicus.eu/web/sentinel/home> [Accessed 03.01 2022].
- FAN, X., LIU, Y., WU, G. & ZHAO, X. 2020. Compositing the Minimum NDVI for Daily Water Surface Mapping. *Remote Sensing*, 12.
- FARR, T. G., ROSEN, P. A., CARO, E., CRIPPEN, R., DUREN, R., HENSLEY, S., KOBRICK, M., PALLER, M., RODRIGUEZ, E., ROTH, L., SEAL, D., SHAFFER, S., SHIMADA, J., UMLAND, J., WERNER, M., OSKIN, M., BURBANK, D. & ALSDORF, D. 2007. The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45.
- FEYISA, G. L., MEILBY, H., FENSHOLT, R. & PROUD, S. R. 2014. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, 140, 23-35.
- FROST, V. S., STILES, J. A., SHANMUGAN, K. S. & HOLTZMAN, J. C. 1982. A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on pattern analysis and machine intelligence*, 157-166.
- FUNK, C., PETERSON, P., LANDSFELD, M., PEDREROS, D., VERDIN, J., SHUKLA, S., HUSAK, G., ROWLAND, J., HARRISON, L., HOELL, A. & MICHAELSEN, J. 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2, 150066.
- FÜREDER, P., LANG, S., HAGENLOCHER, M., TIEDE, D., WENDT, L. & ROGENHOFER, E. Earth observation and GIS to support humanitarian operations in refugee/IDP camps. ISCRAM, 2015.
- GAO, B.-C. 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58, 257-266.
- GAUGHAN, A. E., STEVENS, F. R., LINARD, C., JIA, P. & TATEM, A. J. 2013. High resolution population distribution maps for Southeast Asia in 2010 and 2015. *PLoS One*, 8, e55882.
- GEOFABRIK GMBH. 2007. *Geofabrik download server* [Online]. Available: <https://www.geofabrik.de/data/download.html> [Accessed 04.01 2022].

- GEUDTNER, D., TORRES, R., SNOEIJ, P., DAVIDSON, M. & ROMMEN, B. Sentinel-1 system capabilities and applications. 2014 IEEE Geoscience and Remote Sensing Symposium, 2014. IEEE, 1457-1460.
- GOBEYN, S., VAN WESEMAEL, A., NEAL, J., LIEVENS, H., EERDENBRUGH, K. V., DE VLEESCHOUWER, N., VERNIEUWE, H., SCHUMANN, G. J. P., DI BALDASSARRE, G., BAETS, B. D., BATES, P. D. & VERHOEST, N. E. C. 2017. Impact of the timing of a SAR image acquisition on the calibration of a flood inundation model. *Advances in Water Resources*, 100, 126-138.
- GOOGLE LLC. 2010. *Google Earth Engine* [Online]. Available: <https://earthengine.google.com> [Accessed 04.01 2022].
- GOOGLE LLC. 2020. *WorldPop Global Project Population Data: Constrained Estimated Age and Sex Structures of Residential Population per 100x100m Grid Square* [Online]. Available: https://developers.google.com/earth-engine/datasets/catalog/WorldPop_GP_100m_pop_age_sex_cons_unadj [Accessed 25.01 2022].
- GOOGLE LLC. 2021. *Resampling and Reducing Resolution* [Online]. Available: <https://developers.google.com/earth-engine/guides/resample?hl=en> [Accessed 08.01 2022].
- GORELICK, N., HANCHER, M., DIXON, M., ILYUSHCHENKO, S., THAU, D. & MOORE, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27.
- GROUP ON EARTH OBSERVATION. 2005. *GEO - who we are* [Online]. Available: https://earthobservations.org/geo_community.php [Accessed 14.01 2022].
- GULÁCSI, A. & KOVÁCS, F. 2020. Sentinel-1-Imagery-Based High-Resolution Water Cover Detection on Wetlands, Aided by Google Earth Engine. *Remote Sensing*, 12.
- HOEKMAN, D. H. & REICHE, J. 2015. Multi-model radiometric slope correction of SAR images of complex terrain using a two-stage semi-empirical approach. *Remote Sensing of Environment*, 156, 1-10.
- HONG, S., JANG, H., KIM, N. & SOHN, H. G. 2015. Water area extraction using RADARSAT SAR imagery combined with Landsat imagery and terrain information. *Sensors (Basel)*, 15, 6652-67.
- HUANG, W., DEVRIES, B., HUANG, C., LANG, M., JONES, J., CREED, I. & CARROLL, M. 2018. Automated Extraction of Surface Water Extent from Sentinel-1 Data. *Remote Sensing*, 10.
- HUMDATA. 2022. *South Sudan population estimates* [Online]. Available: <https://data.humdata.org/dataset/south-sudan-administrative-levels-0-2-2020-population-estimates> [Accessed].
- ICEYE. 2021. *Flood Monitoring* [Online]. Available: <https://www.iceye.com/solutions/flood-monitoring> [Accessed 24.01 2022].
- IGNATENKO, V., LAURILA, P., RADIUS, A., LAMENTOWSKI, L., ANTROPOV, O. & MUFF, D. 2020. ICEYE Microsatellite SAR Constellation Status Update: Evaluation of First Commercial Imaging Modes. *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*.
- JUTZ, S. L. G. 2020. *The Copernicus Space Component: Today's Status and the Future* [Online]. European parliament, Brussels: Copernicus4Regions Info Session. Available: http://www.nereus-regions.eu/wp-content/uploads/2020/01/The-Copernicus-Space-Component_Simon-L.-G.-Jutz.pdf [Accessed 06.01 2022].

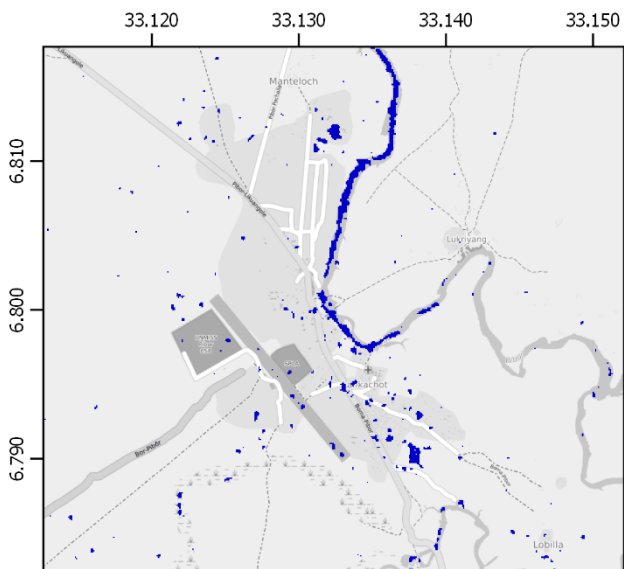
- KARRA, K., KONTGIS, C., STATMAN-WEIL, Z., MAZZARIELLO, J. C., MATHIS, M. & BRUMBY, S. P. 2021. Global land use / land cover with Sentinel 2 and deep learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*.
- KUAN, D. T., SAWCHUK, A. A., STRAND, T. C. & CHAVEL, P. 1985. Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE transactions on pattern analysis and machine intelligence*, 165-177.
- LANG, S., FÜREDER, P., KRANZ, O., CARD, B., ROBERTS, S. & PAPP, A. 2015. Humanitarian emergencies: causes, traits and impacts as observed by remote sensing. *Remote sensing of water resources, disasters, and urban studies*. CRC Press.
- LANG, S., FÜREDER, P., RIEDLER, B., WENDT, L., BRAUN, A., TIEDE, D., SCHOEPFER, E., ZEIL, P., SPRÖHNLE, K., KULESSA, K., ROGENHOFER, E., BÄUERL, M., ÖZE, A., SCHWENDEMANN, G. & HOCHSCHILD, V. 2019. Earth observation tools and services to increase the effectiveness of humanitarian assistance. *European Journal of Remote Sensing*, 53, 67-85.
- LANG, S., FÜREDER, P. & ROGENHOFER, E. 2018. Yearbook on Space Policy 2016: Space for Sustainable Development. In: AL-EKABI, C. & FERRETTI, S. (eds.). Springer.
- LANG, S., SCHOEPFER, E., ZEIL, P. & RIEDLER, B. 2017. Earth Observation for Humanitarian Assistance. *GI_Forum*, 1, 157-165.
- LEE, J.-S. 1980. Digital image enhancement and noise filtering by use of local statistics. *IEEE transactions on pattern analysis and machine intelligence*, 165-168.
- LEE, J. S., JURKEVICH, L., DEWAELE, P., WAMBACQ, P. & OOSTERLINCK, A. 1994. Speckle filtering of synthetic aperture radar images: A review. *Remote Sensing Reviews*, 8, 313-340.
- LI, D. 2010. An overview of earth observation and geospatial information service. *Geospatial technology for earth observation*, 1-25.
- LI, Y., MARTINIS, S. & WIELAND, M. 2019. Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 178-191.
- LIANG, J. & LIU, D. 2020. A local thresholding approach to flood water delineation using Sentinel-1 SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159, 53-62.
- LINARD, C., GILBERT, M., SNOW, R. W., NOOR, A. M. & TATEM, A. J. 2012. Population distribution, settlement patterns and accessibility across Africa in 2010. *PLoS One*, 7, e31743.
- LONG, S., FATOYINBO, T. E. & POLICELLI, F. 2014. Flood extent mapping for Namibia using change detection and thresholding with SAR. *Environmental Research Letters*, 9, 035002.
- LOPES, A., NEZRY, E., TOUZI, R. & LAUR, H. 1993. Structure detection and statistical adaptive speckle filtering in SAR images. *International Journal of Remote Sensing*, 14, 1735-1758.
- MAIN-KNORN, M., PFLUG, B., LOUIS, J., DEBAECKER, V., MÜLLER-WILM, U. & GASCON, F. Sen2Cor for sentinel-2. Image and Signal Processing for Remote Sensing XXIII, 2017. International Society for Optics and Photonics, 1042704.

- MAÎTRE, H., LANG, S., SUN, H., LIN, C. Y., LIU, J., LEI, B., FENG, J., WONG, N. & SO, H. K.-H. 2009. A comparison of SAR image speckle filters. *MIPPR 2009: Remote Sensing and GIS Data Processing and Other Applications*.
- MARKERT, K. N., MARKERT, A. M., MAYER, T., NAUMAN, C., HAAG, A., POORTINGA, A., BHANDARI, B., THWAL, N. S., KUNLAMAI, T., CHISHTIE, F., KWANT, M., PHONGSAPAN, K., CLINTON, N., TOWASHIRAPORN, P. & SAAH, D. 2020. Comparing Sentinel-1 Surface Water Mapping Algorithms and Radiometric Terrain Correction Processing in Southeast Asia Utilizing Google Earth Engine. *Remote Sensing*, 12.
- META. 2019. *Data for Good at Meta (previously Facebook)* [Online]. Available: <https://dataforgood.facebook.com/dfg/tools/high-resolution-population-density-maps> [Accessed 25.01 2022].
- MSF. 2019. *A state of emergency declared as flooding in South Sudan ensues* [Online]. Available: <https://msf.or.ke/en/magazine/state-emergency-declared-flooding-south-sudan-ensues-0> [Accessed 14.01 2022].
- MSF. 2020. Available: <https://www.msf.org/heavy-floods-threaten-lives-thousands-people-eastern-south-sudan> [Accessed 14.01 2022].
- MSF. 2021. *Third year of severe floods leaves nearly 800,000 people struggling* [Online]. Available: <https://www.msf.org/third-year-floods-leaves-people-struggling-south-sudan> [Accessed 14.01 2022].
- MSF. 2022. *South Sudan: Responding to the worst floods in decades* [Online]. Available: <https://www.doctorswithoutborders.org/what-we-do/news-stories/story/south-sudan-responding-worst-floods-decades> [Accessed].
- MULLISSA, A., VOLLRATH, A., ODONGO-BRAUN, C., SLAGTER, B., BALLING, J., GOU, Y., GORELICK, N. & REICHE, J. 2021. Sentinel-1 SAR Backscatter Analysis Ready Data Preparation in Google Earth Engine. *Remote Sensing*, 13.
- OLOFSSON, P., FOODY, G. M., HEROLD, M., STEHMAN, S. V., WOODCOCK, C. E. & WULDER, M. A. 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148, 42-57.
- OPENSTREETMAP FOUNDATION. 2004. *OpenStreetMap* [Online]. Available: <https://www.openstreetmap.org/> [Accessed 04.01 2022].
- OTSU, N. 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9, 62-66.
- PEARSON, R. L. & MILLER, L. D. 1972. Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie. *Remote sensing of environment*, VIII, 1355.
- PEEL, M. C., FINLAYSON, B. L. & MCMAHON, T. A. 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.*, 11, 1633-1644.
- 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.
- PERONA, P. & MALIK, J. 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, 12, 629-639.
- PLATE, E. J. 2002. Flood risk and flood management. *Journal of Hydrology*, 267, 2-11.
- RAMM, F., TOPF, J. & CHILTON, S. 2011. *OpenStreetMap: using and enhancing the free map of the world*, UIT Cambridge Cambridge.

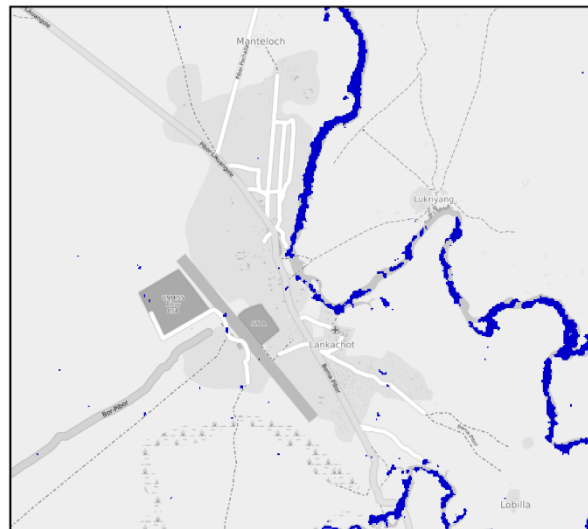
- SIRKO, W., KASHUBIN, S., RITTER, M., ANNKAH, A., BOUCHARB, Y. S. E., DAUPHIN, Y., KEYSERS, D., NEUMANN, M., CISSE, M. & QUINN, J. 2021. Continental-scale building detection from high resolution satellite imagery. *arXiv preprint arXiv:2107.12283*.
- SPRÖHNLE, K., TIEDE, D., SCHOEPFER, E., FÜREDER, P., SVANBERG, A. & ROST, T. 2014. Earth observation-based dwelling detection approaches in a highly complex refugee camp environment—a comparative study. *Remote Sensing*, 6, 9277-9297.
- STEVENS, F. R., GAUGHAN, A. E., LINARD, C. & TATEM, A. J. 2015. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PLoS One*, 10, e0107042.
- TATEM, A. J. 2017. WorldPop, open data for spatial demography. *Sci Data*, 4, 170004.
- TATEM, A. J., NOOR, A. M., VON HAGEN, C., DI GREGORIO, A. & HAY, S. I. 2007. High resolution population maps for low income nations: combining land cover and census in East Africa. *PLoS One*, 2, e1298.
- THORBURN, D. H. 1922. The Pibor River. *The Geographical Journal*, 60, 210-217.
- UDDIN, MATIN & MEYER 2019. Operational Flood Mapping Using Multi-Temporal Sentinel-1 SAR Images: A Case Study from Bangladesh. *Remote Sensing*, 11.
- UNITED NATIONS. 2020. *Step-by-Step: Recommended Practice: Flood Mapping and Damage Assessment Using Sentinel-1 SAR Data in Google Earth Engine* [Online]. Available: <https://www.un-spider.org/advisory-support/recommended-practices/recommended-practice-google-earth-engine-flood-mapping/step-by-step> [Accessed 08.01 2022].
- UNIVERSITY OF SALZBURG - Z_GIS. 2021. *GEOHUM* [Online]. Available: <https://geohum.zgis.at/> [Accessed 14.01 2022].
- UNIVERSITY OF SOUTHAMPTON. 2013. *WorldPop* [Online]. Available: <https://www.worldpop.org/> [Accessed 03.01 2022].
- VERPOORTER, C., KUTSER, T. & TRANVIK, L. 2012. Automated mapping of water bodies using Landsat multispectral data. *Limnology and Oceanography: Methods*, 10, 1037-1050.
- VOIGT, S., SCHOEPFER, E., FOURIE, C. & MAGER, A. Towards semi-automated satellite mapping for humanitarian situational awareness. IEEE Global Humanitarian Technology Conference (GHTC 2014), 2014. IEEE, 412-416.
- VOLLRATH, A., MULLISSA, A. & REICHE, J. 2020. Angular-Based Radiometric Slope Correction for Sentinel-1 on Google Earth Engine. *Remote Sensing*, 12.
- WENDT, L., HILBERG, S., ROBL, J., BRAUN, A., ROGENHOFER, E., DIRNBERGER, D., STRASSER, T., FÜREDER, P. & LANG, S. 2015. Using Remote Sensing and GIS to Support Drinking Water Supply in Refugee/IDP Camps. *GI_Forum*, 1, 449-458.
- WENDT, L., LANG, S. & ROGENHOFER, E. 2017. Monitoring of Refugee and Camps for Internally Displaced Persons Using Sentinel-2 Imagery – A Feasibility Study. *GI_Forum*, 1, 172-182.
- WORLDPOP. 2020a. *Top-down estimation modelling: constrained vs unconstrained* [Online]. Available: https://www.worldpop.org/methods/top_down_constrained_vs_unconstrained#:~:text=WorldPop%20top-down%20modelling%20methods,to%20grid%20cell-based%20counts. [Accessed 25.01 2022].

WORLDPOP. 2020b. *WorldPop methods* [Online]. Available:
<https://www.worldpop.org/methods> [Accessed 25.01 2022].

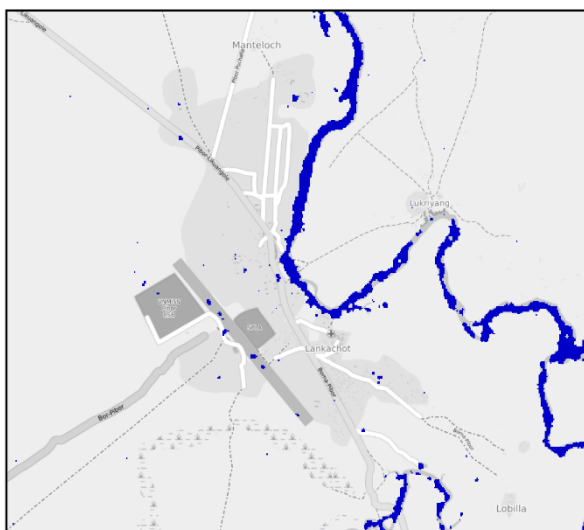
Annex 1



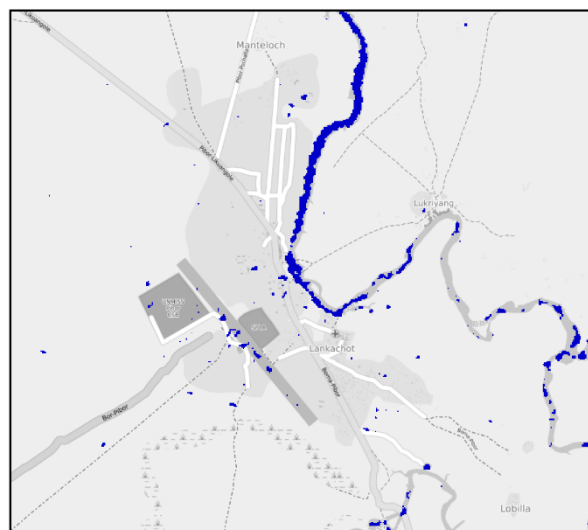
06.06.2017



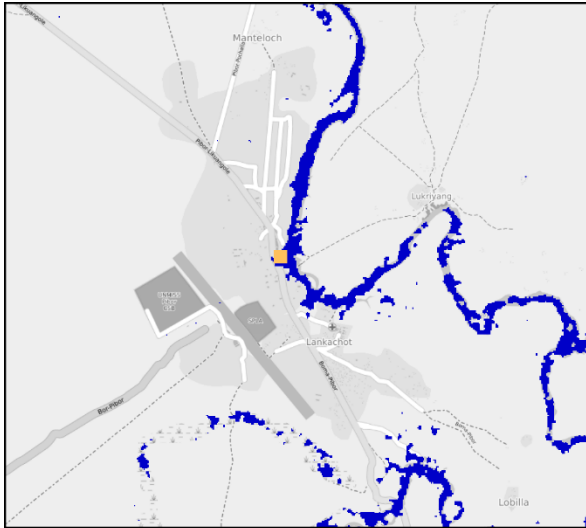
05.08.2017



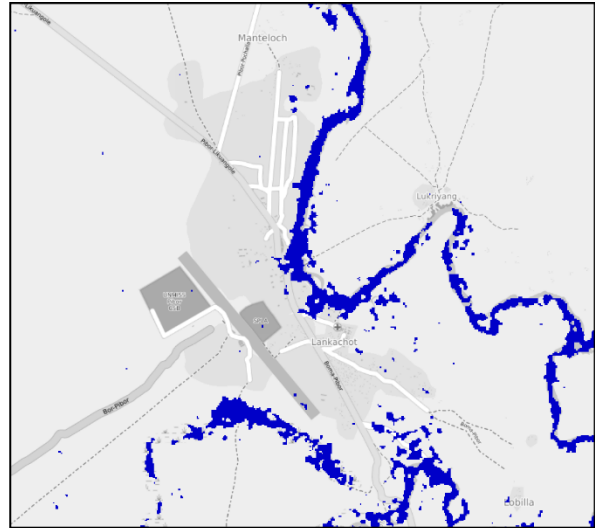
11.08.2017



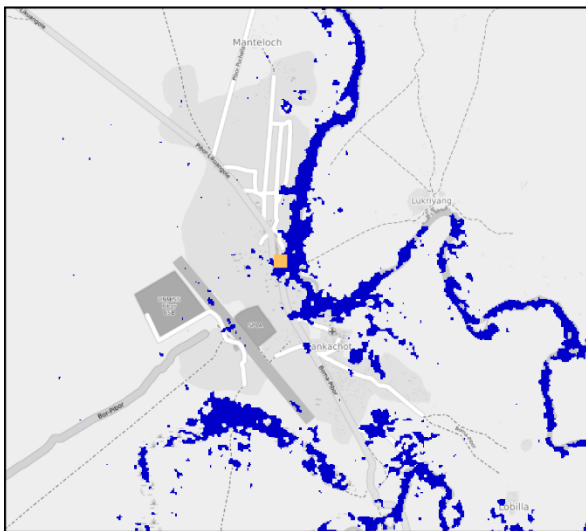
17.08.2017



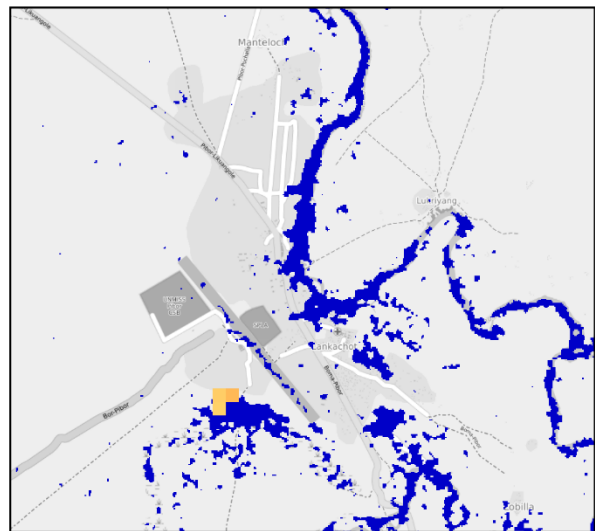
22.09.2017



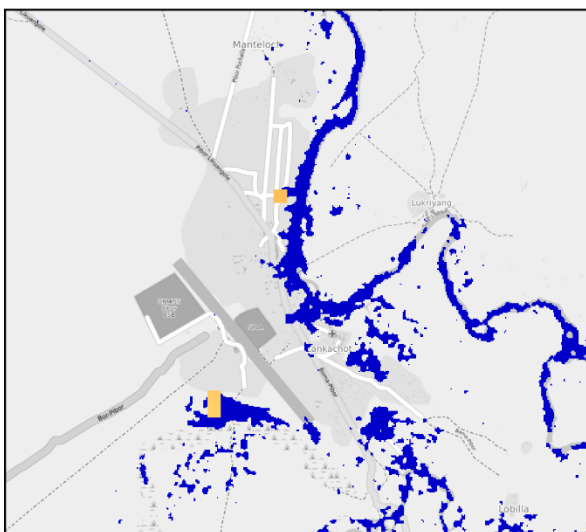
28.09.2017



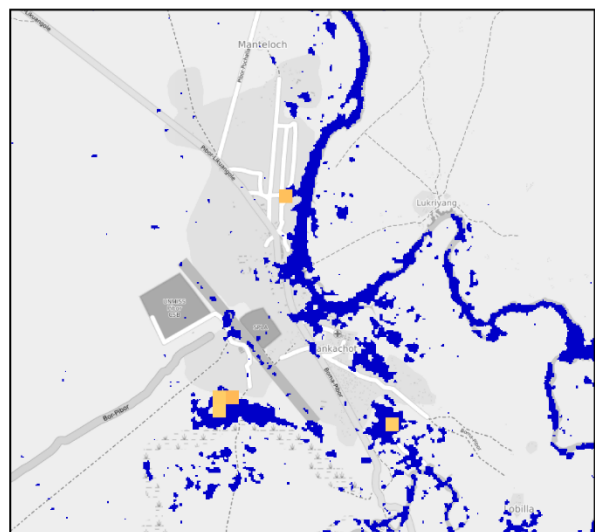
04.10.2017



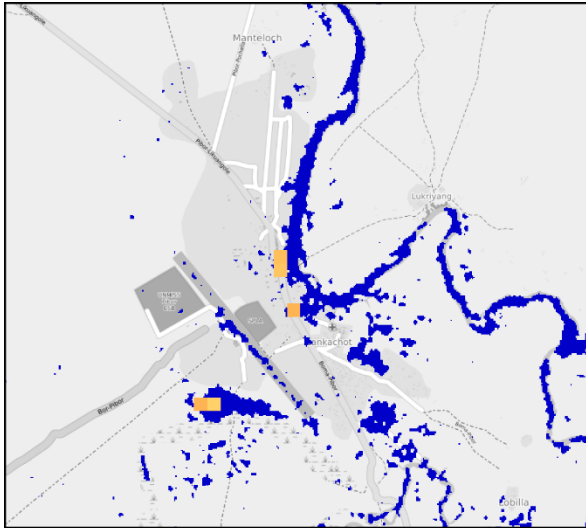
10.10.2017



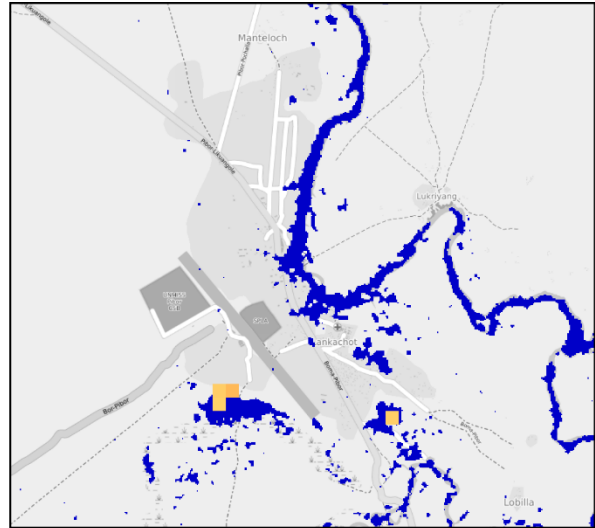
22.10.2017



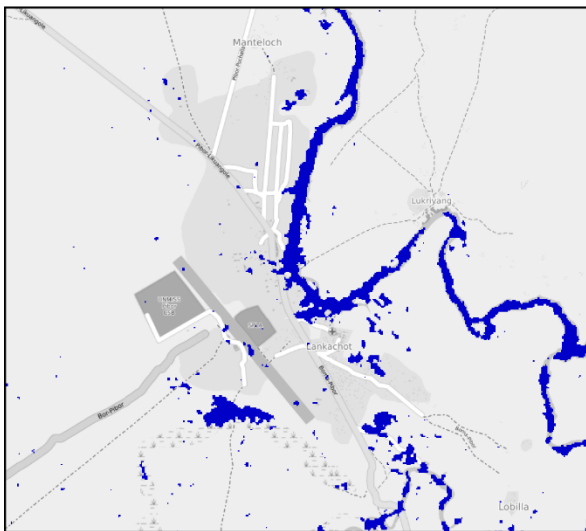
03.11.2017



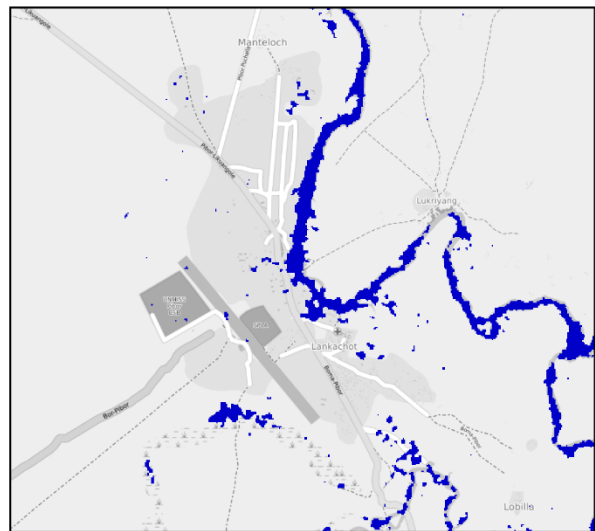
09.11.2017



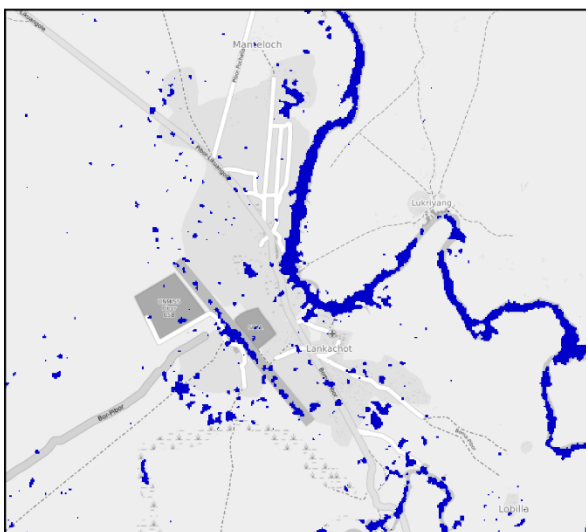
15.11.2017



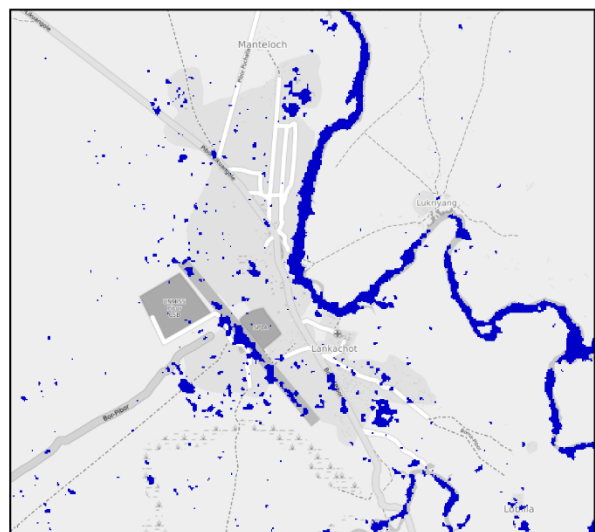
21.11.2017



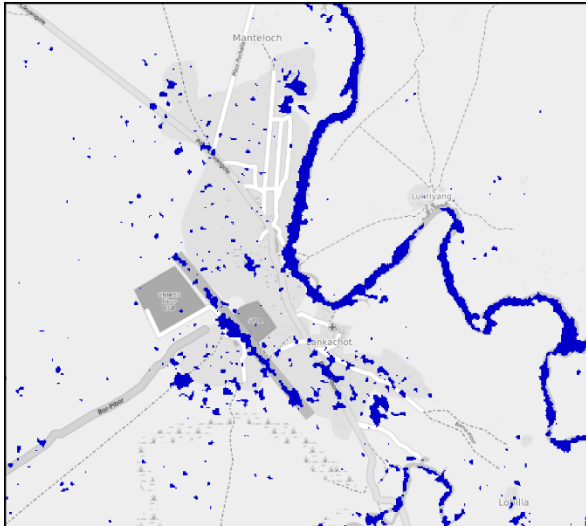
27.11.2017



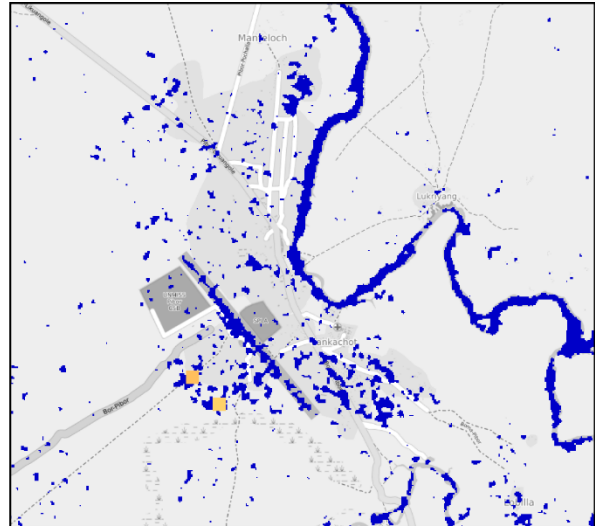
03.12.2017



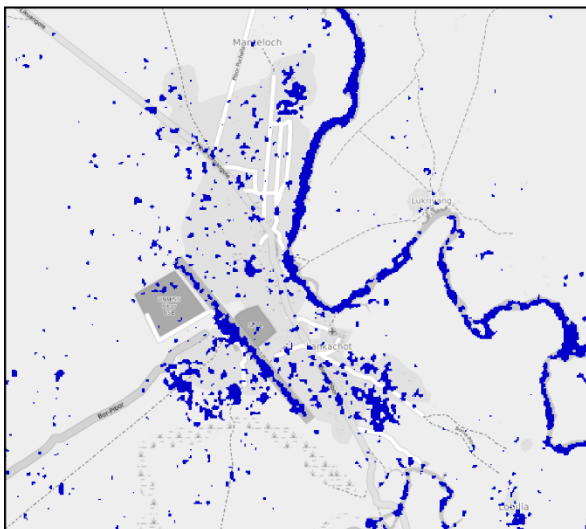
09.12.2017



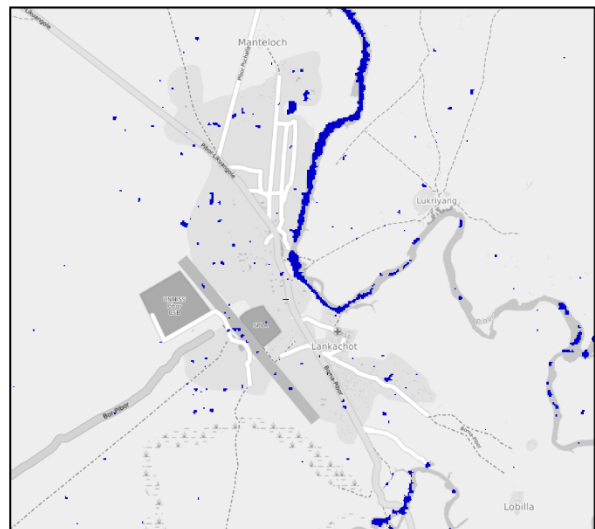
15.12.2017



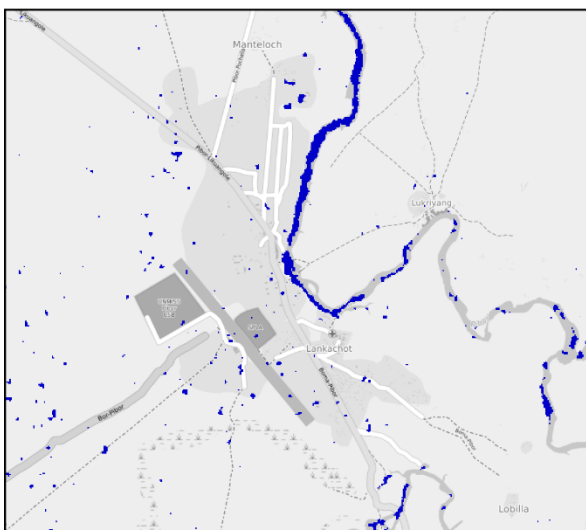
21.12.2017



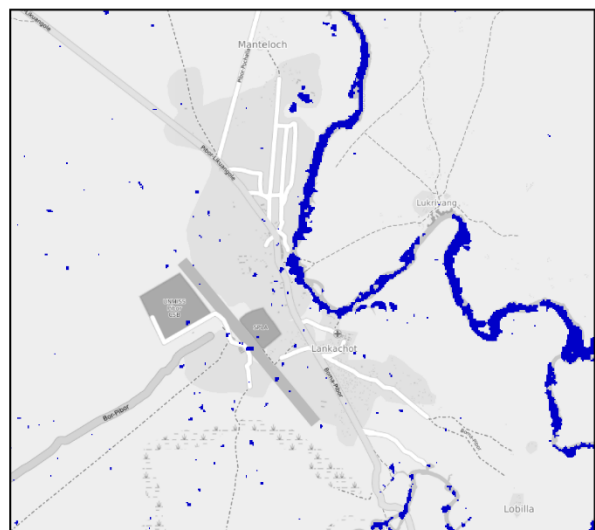
27.12.2017



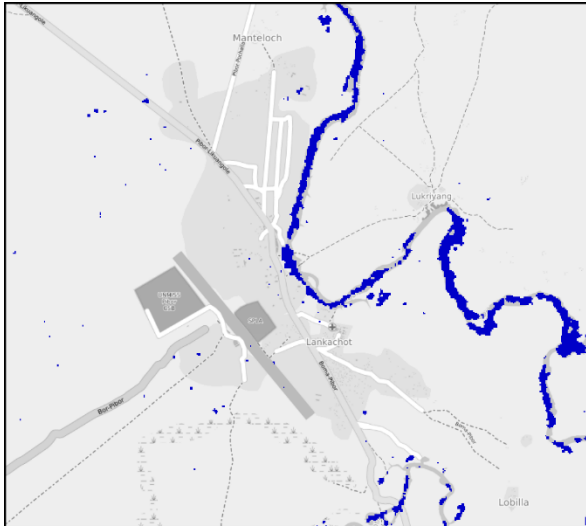
01.06.2018



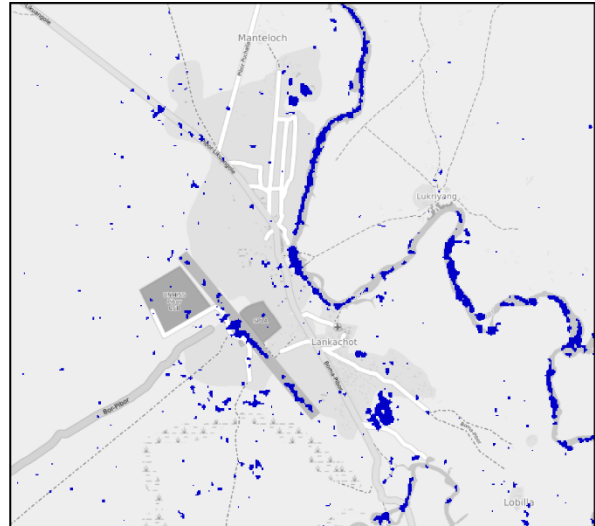
25.06.2018



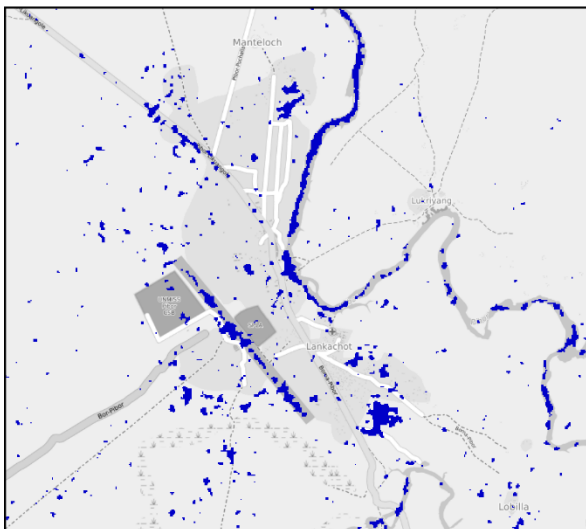
13.07.2018



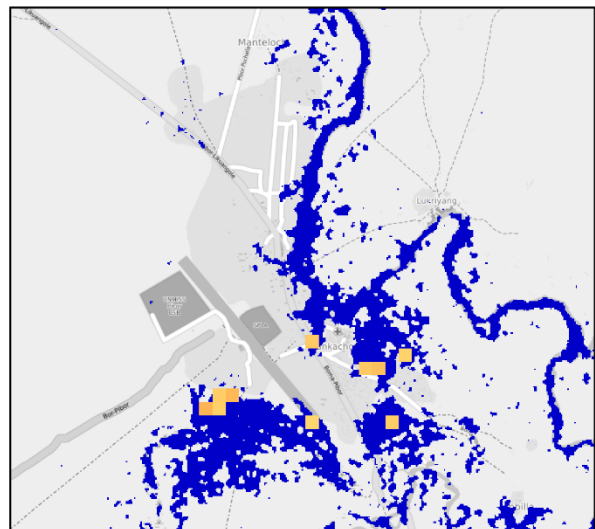
06.08.2018



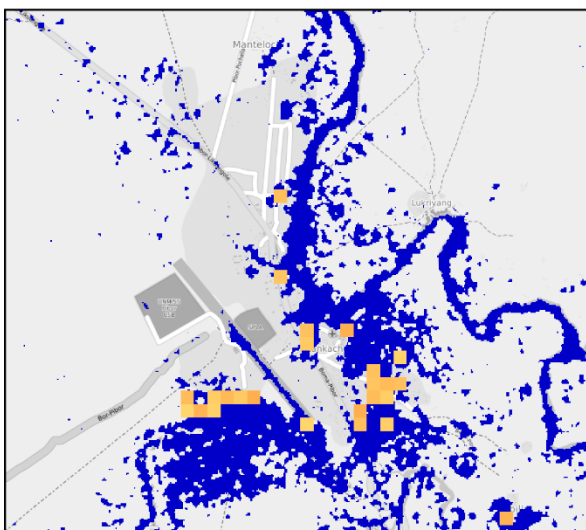
05.10.2018



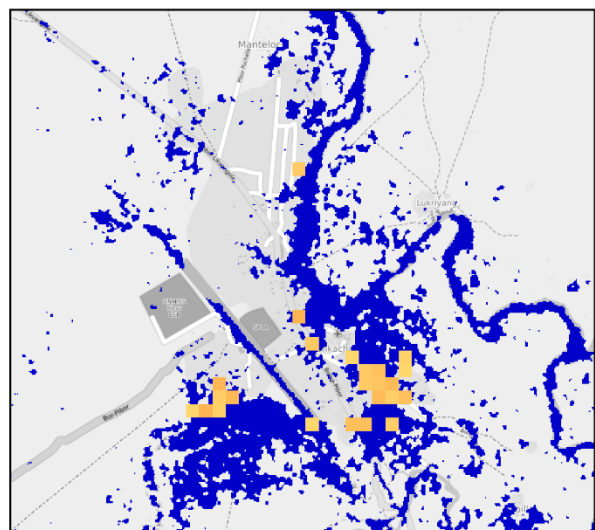
10.11.2018



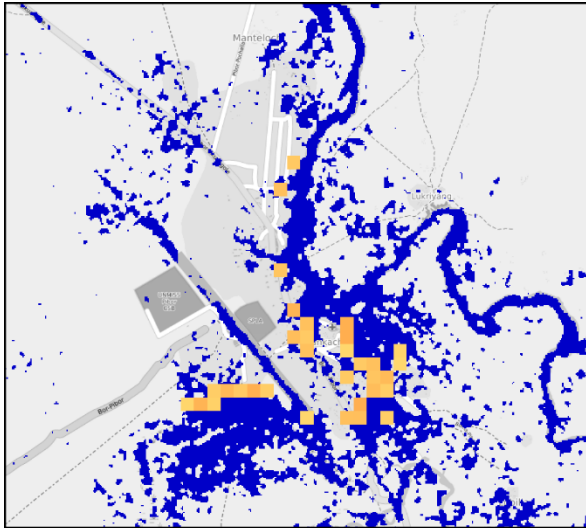
18.10.2019



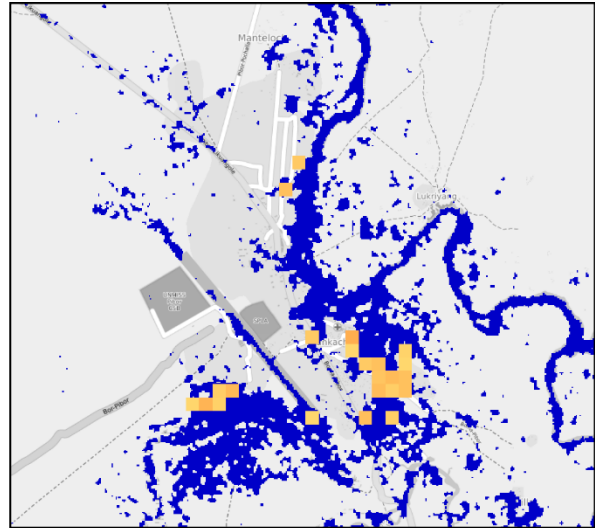
24.10.2019



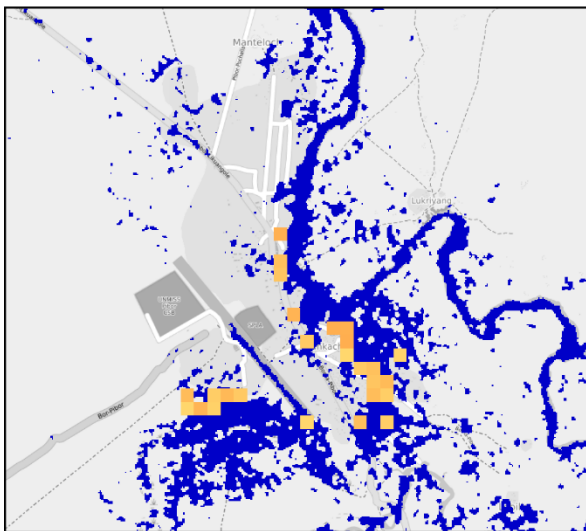
30.10.2019



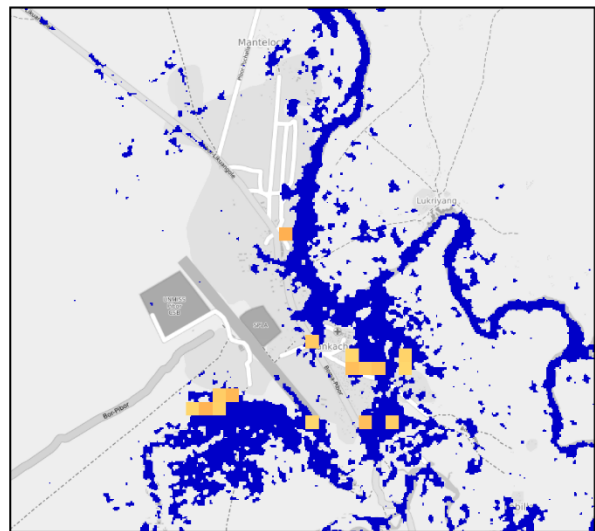
05.11.2019



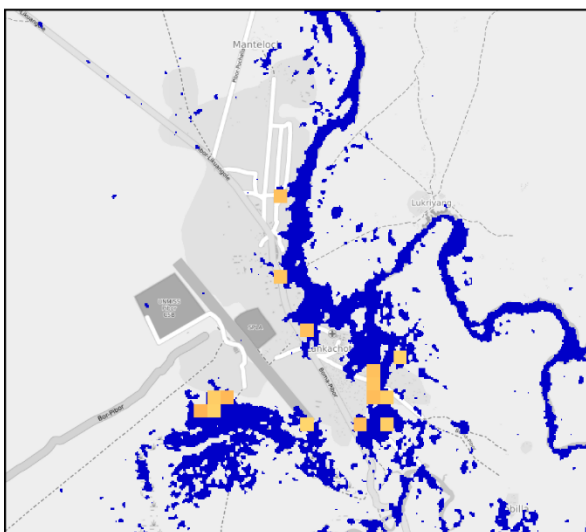
11.11.2019



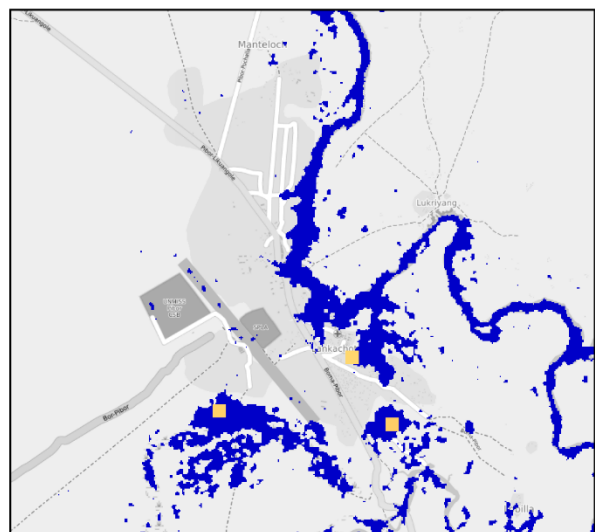
17.11.2019



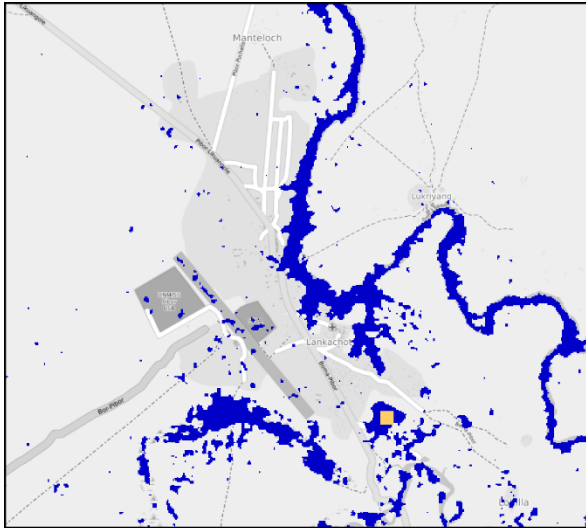
23.11.2019



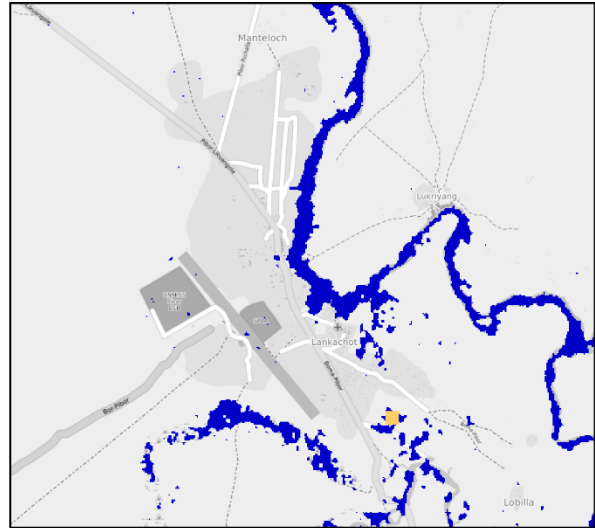
29.11.2019



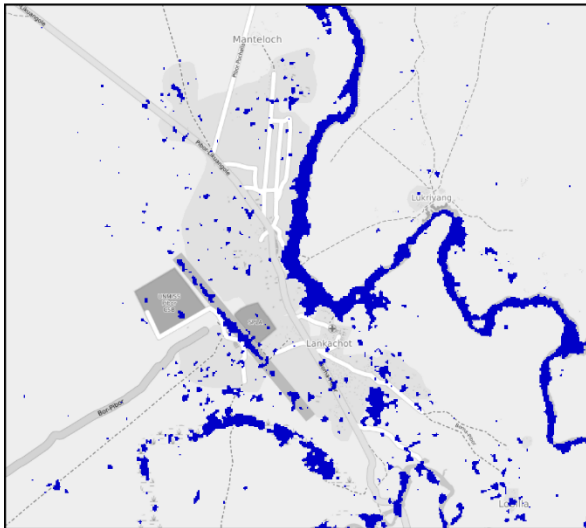
05.12.2019



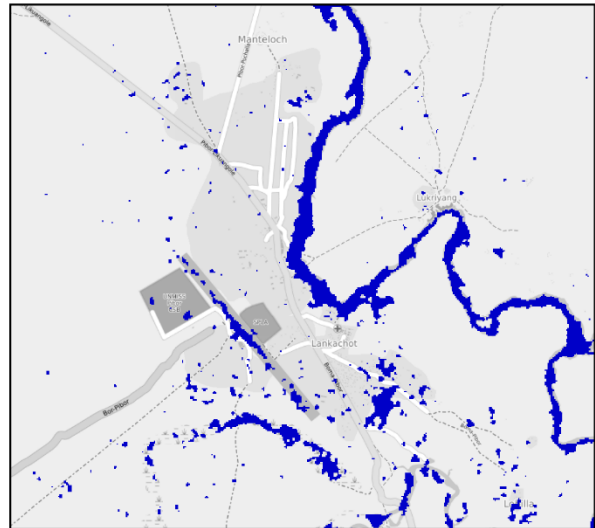
11.12.2019



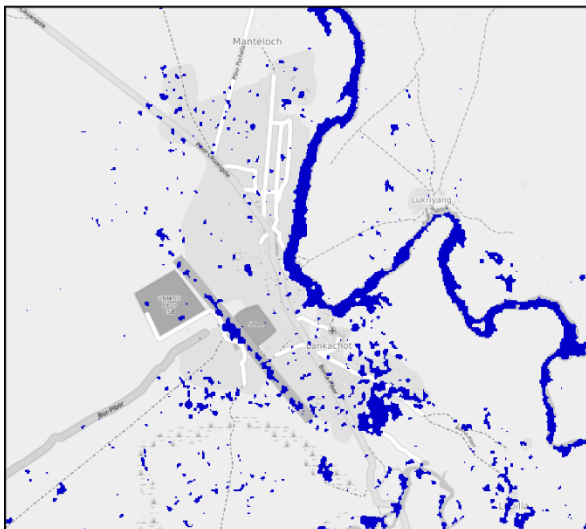
17.12.2019



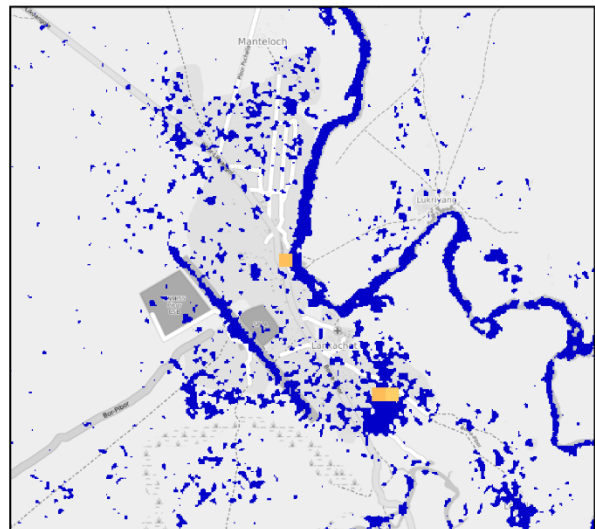
23.12.2019



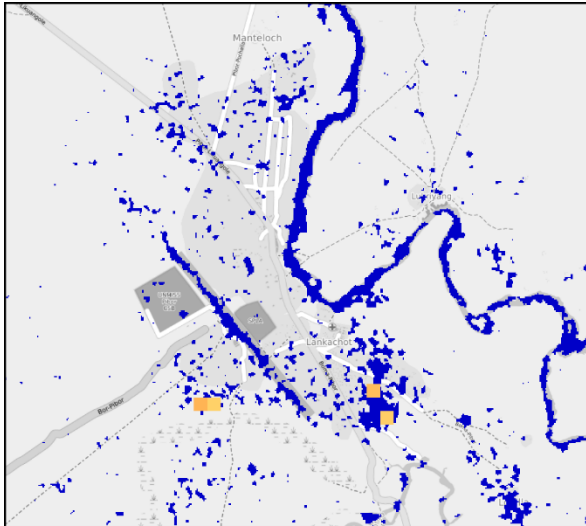
29.12.2019



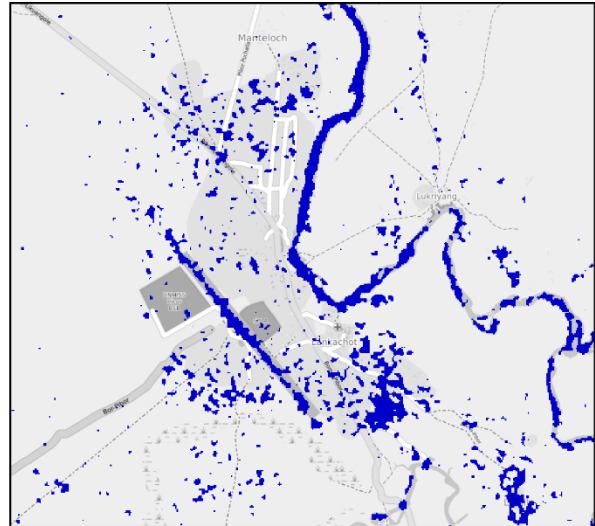
10.01.2020



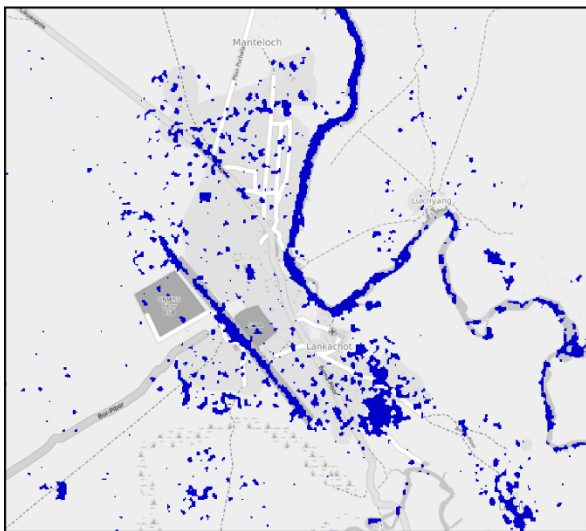
22.01.2020



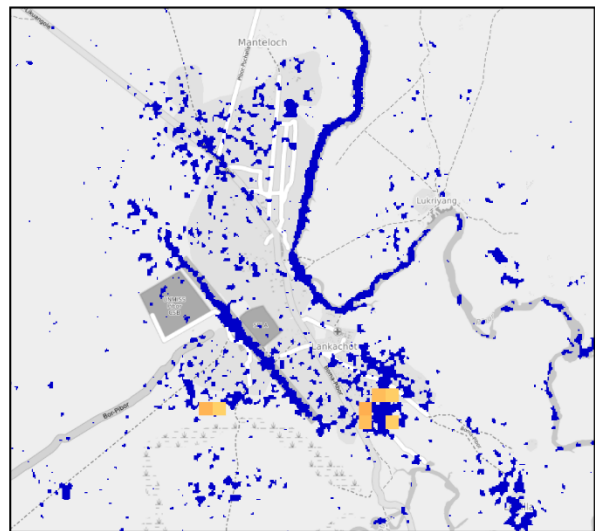
28.01.2020



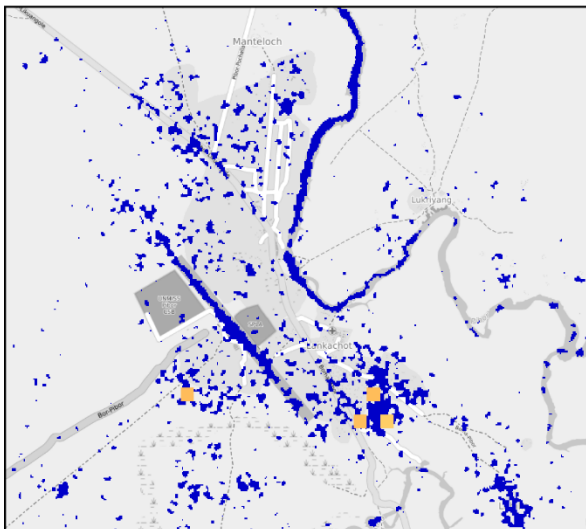
09.02.2020



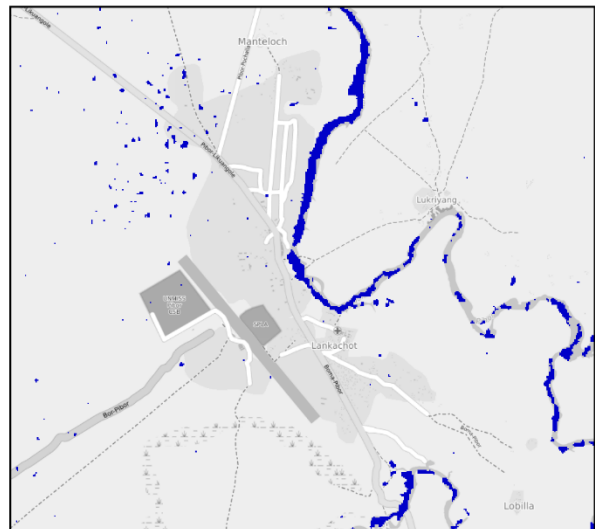
15.02.2020



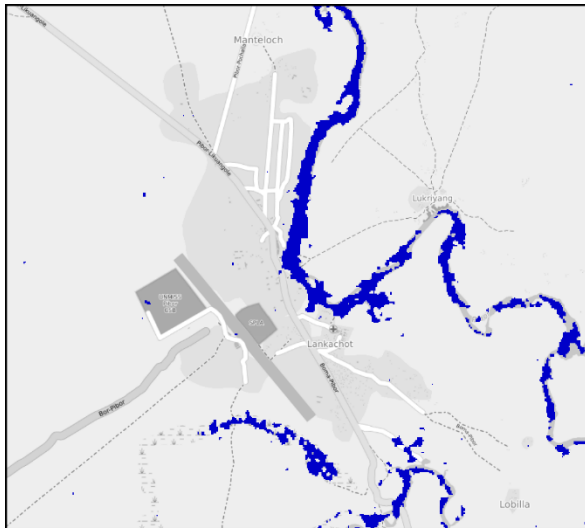
21.02.2020



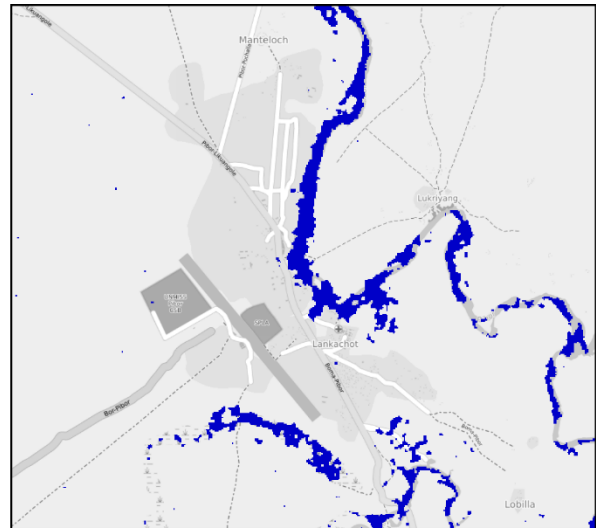
04.03.2020



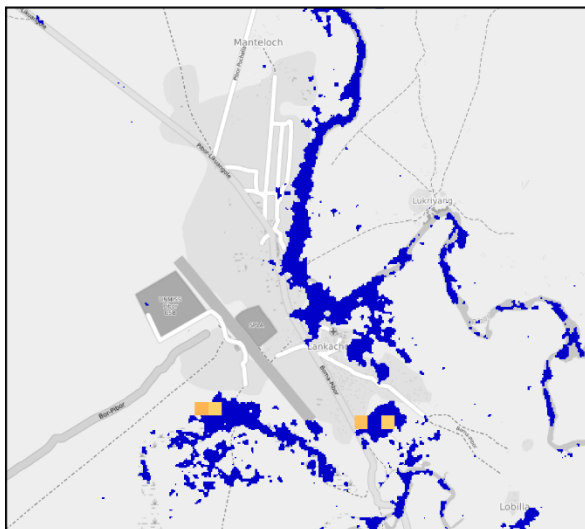
08.06.2020



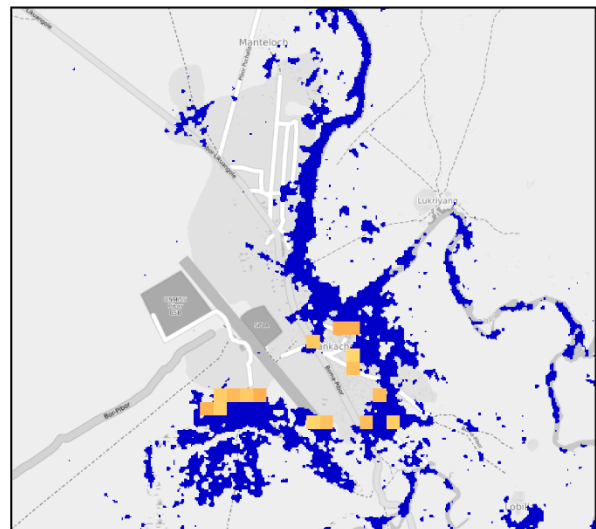
07.08.2020



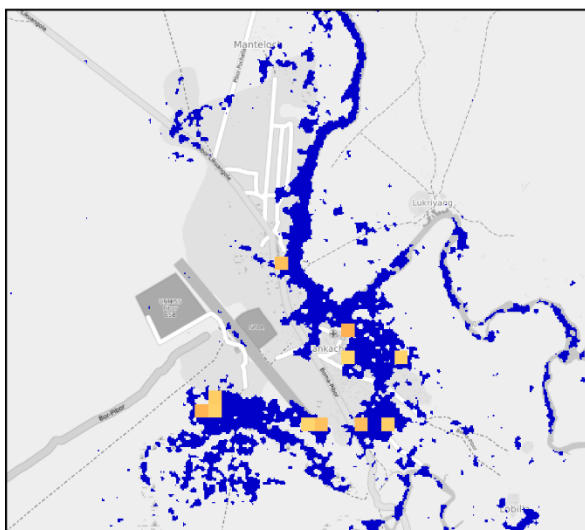
19.08.2020



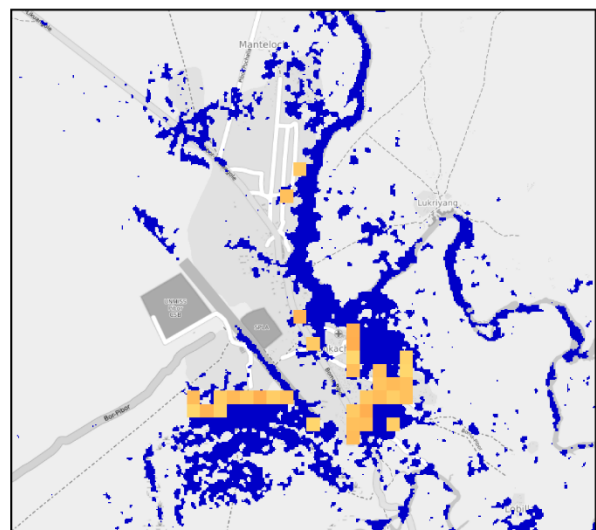
31.08.2020



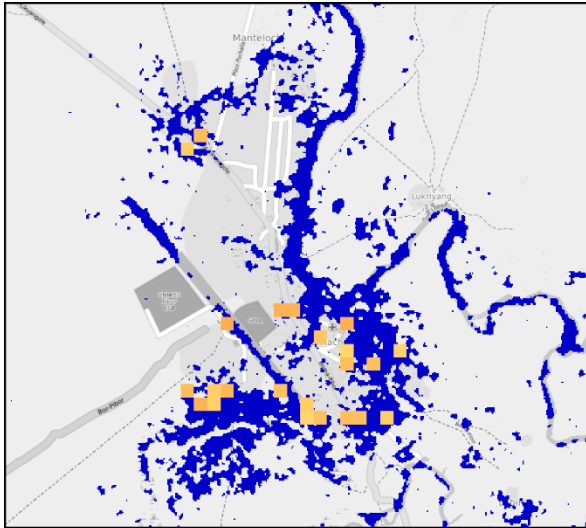
12.09.2020



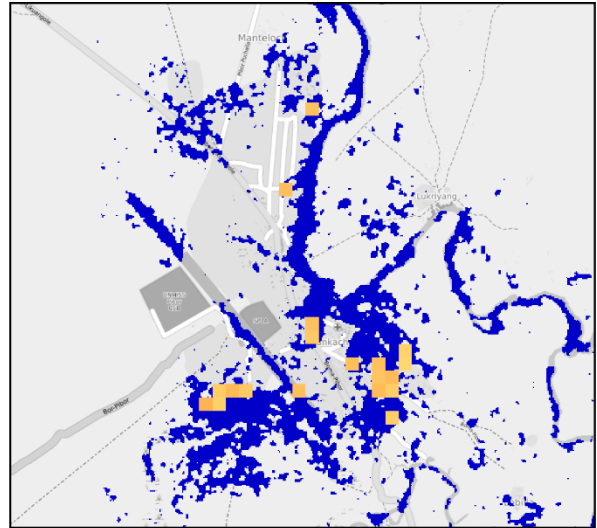
24.09.2020



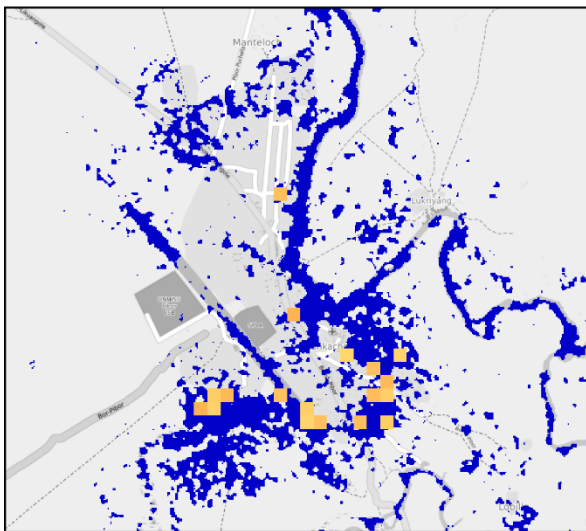
30.09.2020



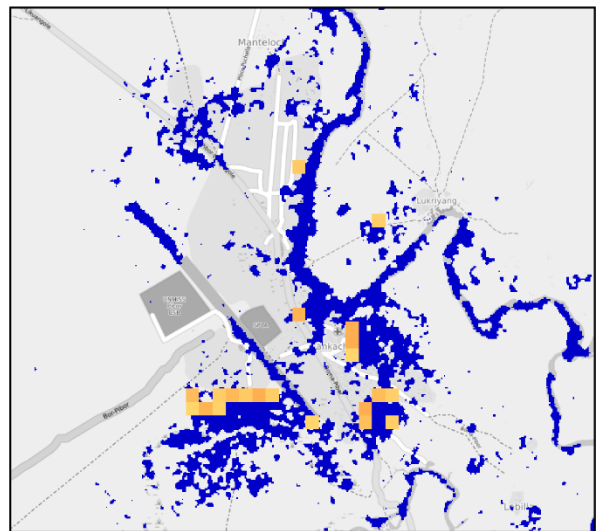
06.10.2020



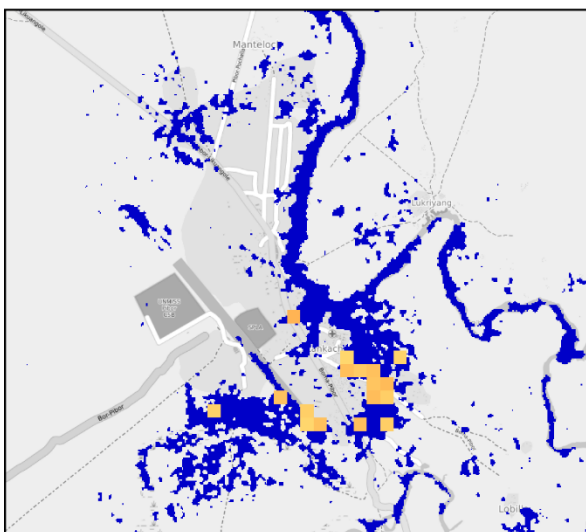
24.10.2020



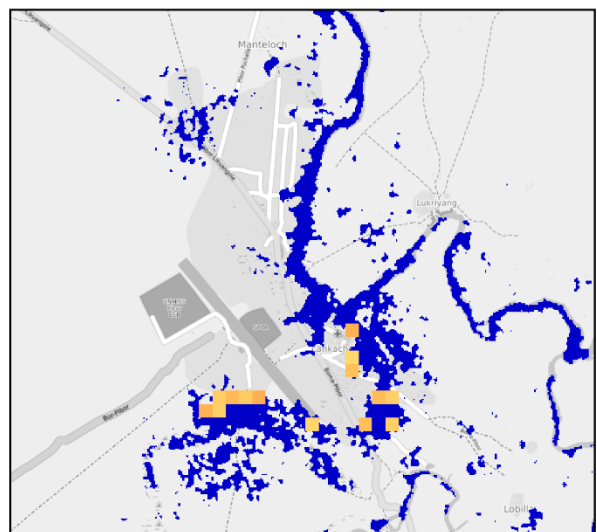
30.10.2020



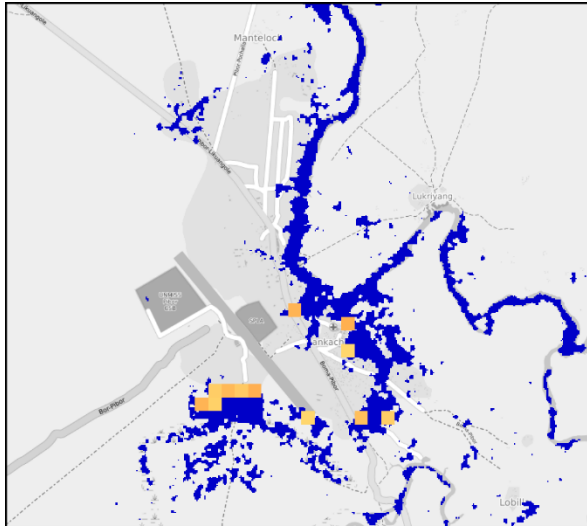
05.11.2020



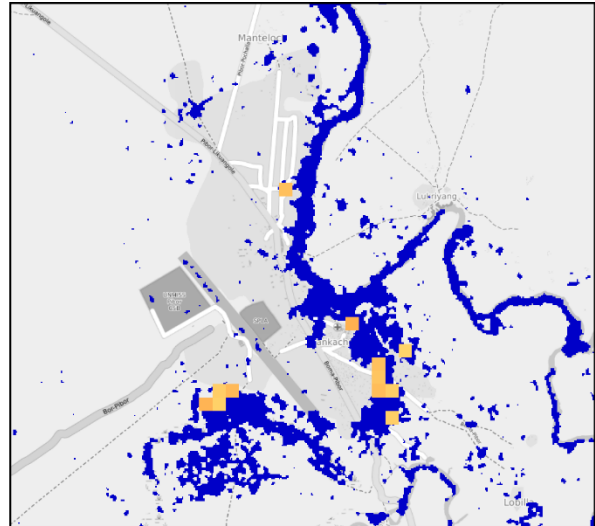
11.11.2020



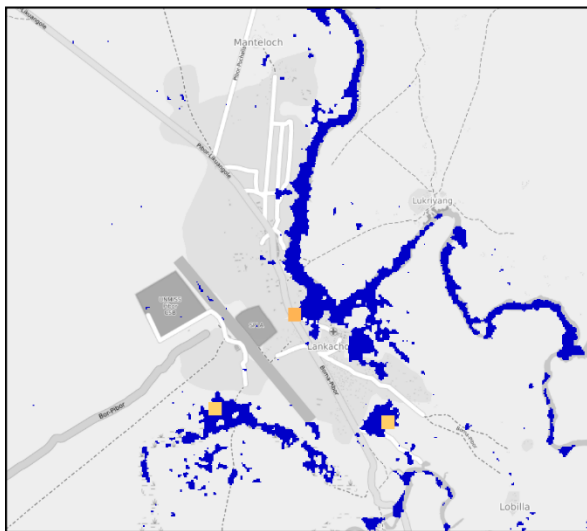
17.11.2020



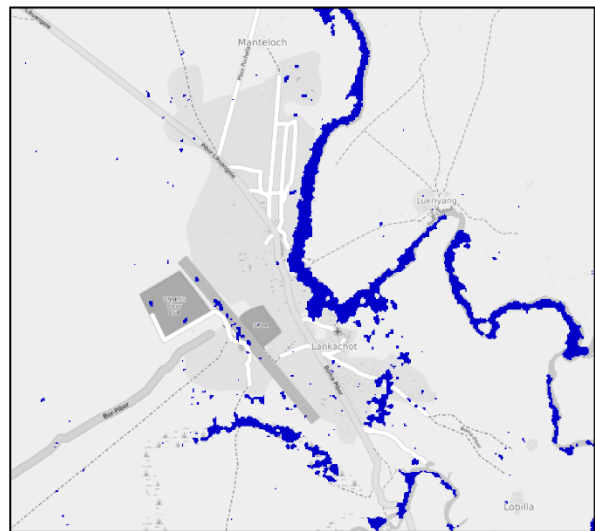
23.11.2020



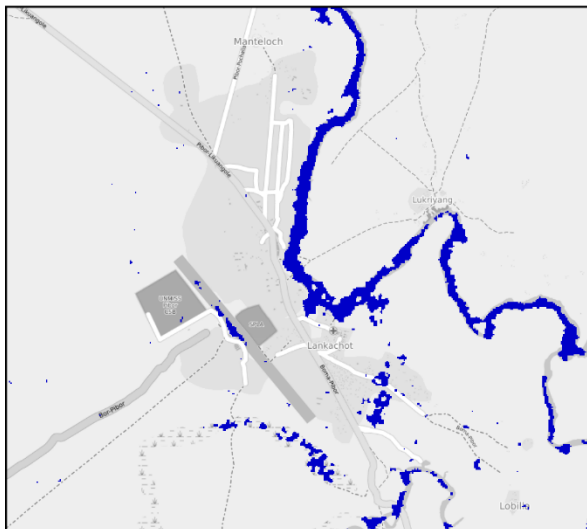
29.11.2020



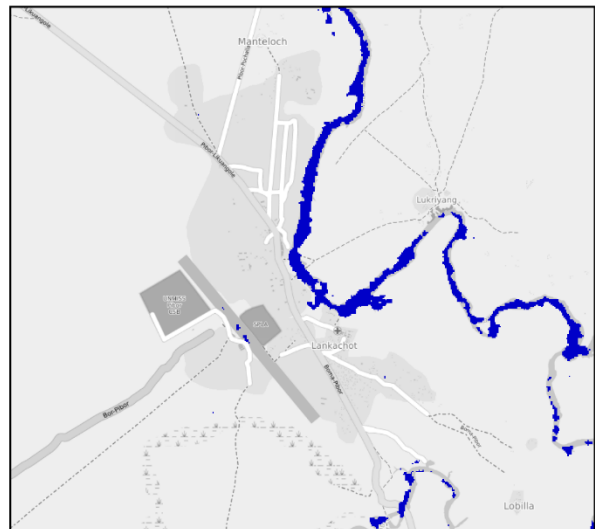
05.12.2020



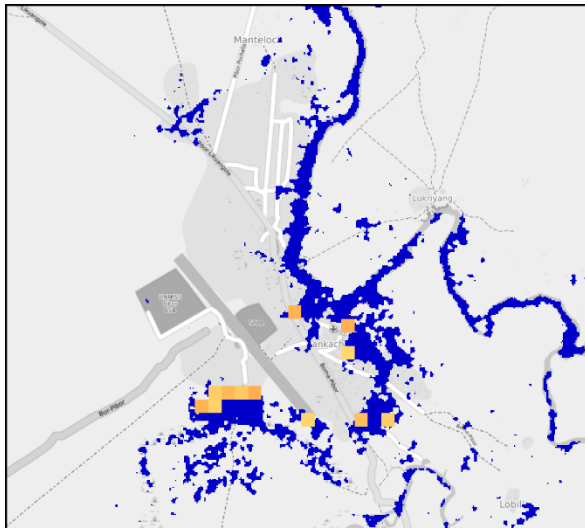
17.12.2020



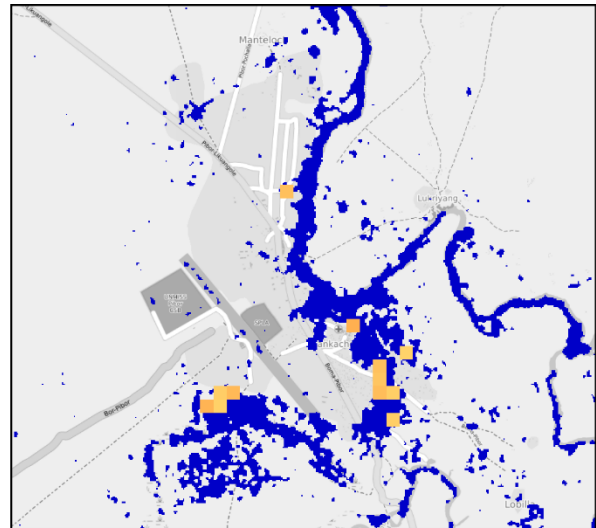
23.12.2020



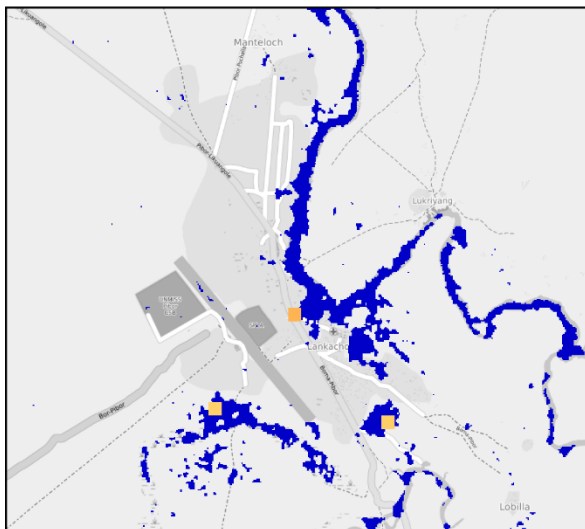
04.01.2021



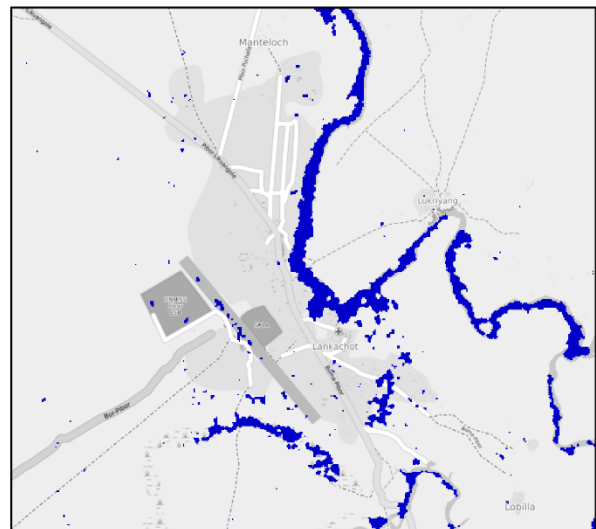
23.11.2020



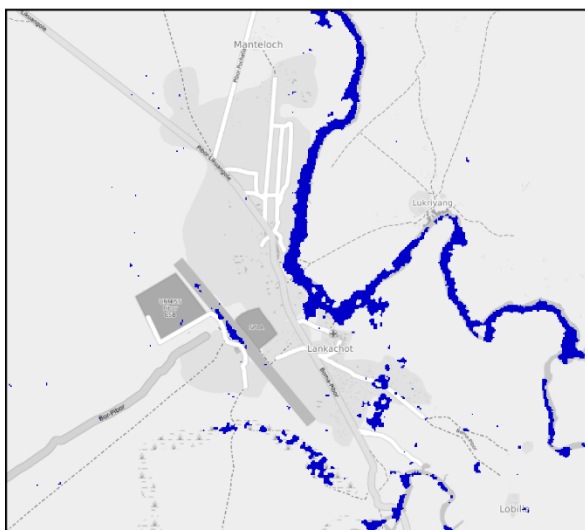
29.11.2020



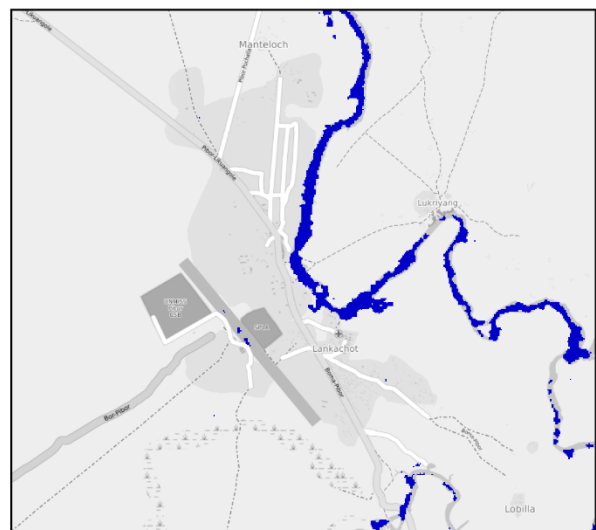
05.12.2020



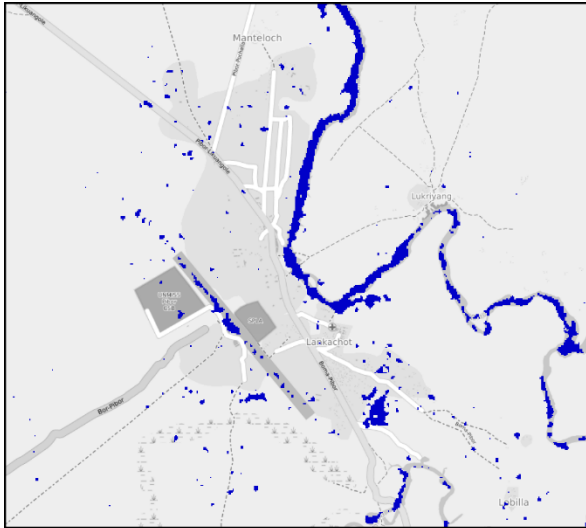
17.12.2020



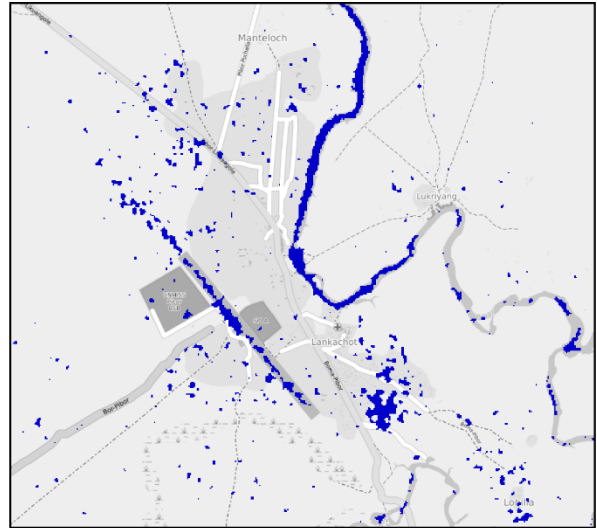
23.12.2020



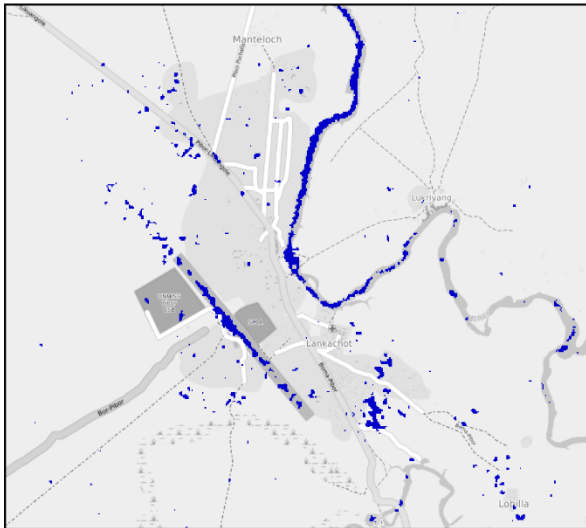
04.01.2021



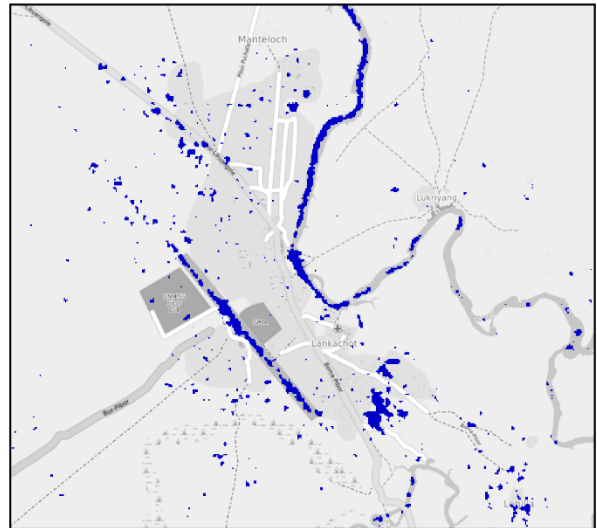
10.01.2021



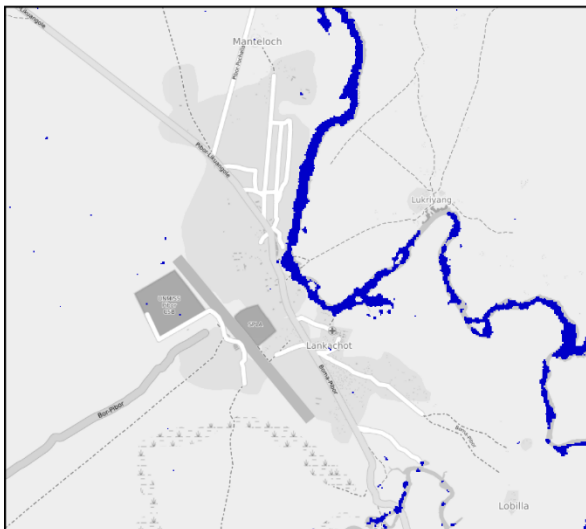
22.01.2021



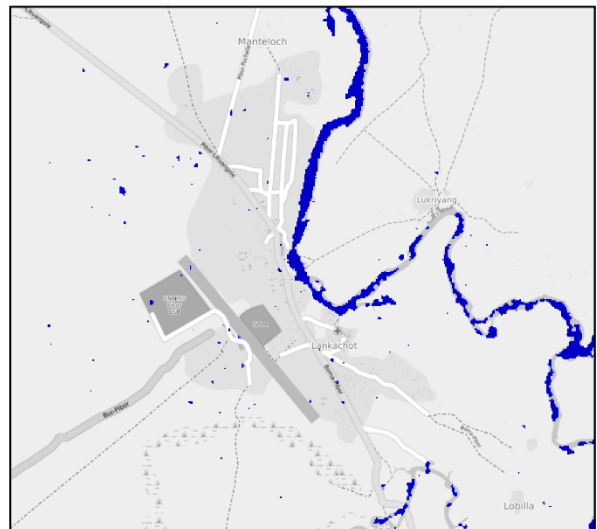
28.01.2021



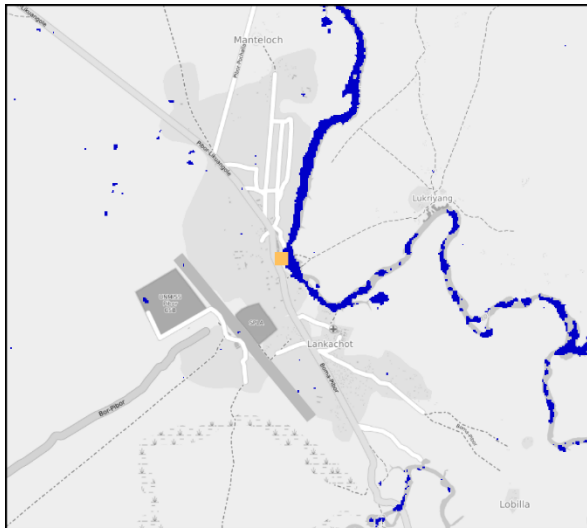
03.02.2021



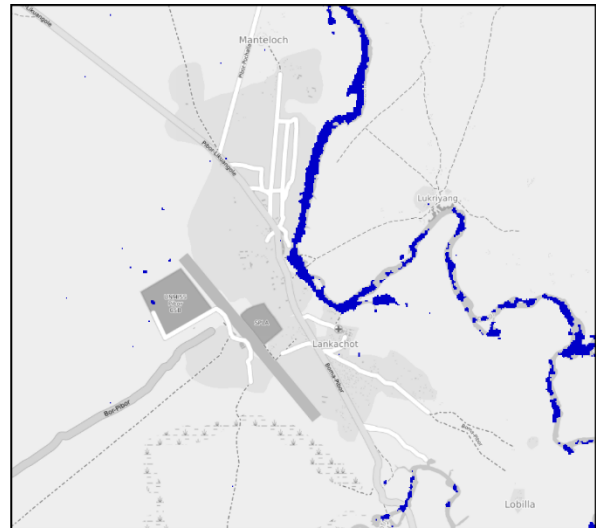
21.07.2021



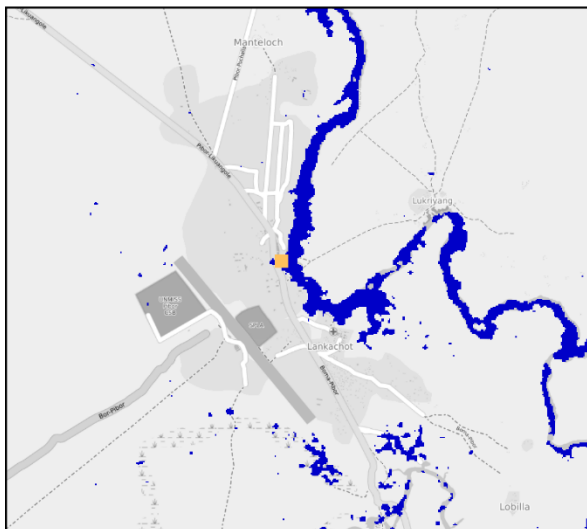
02.08.2021



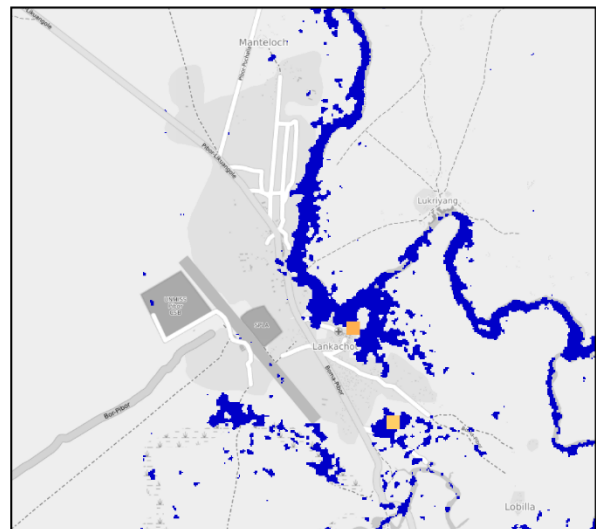
08.08.2021



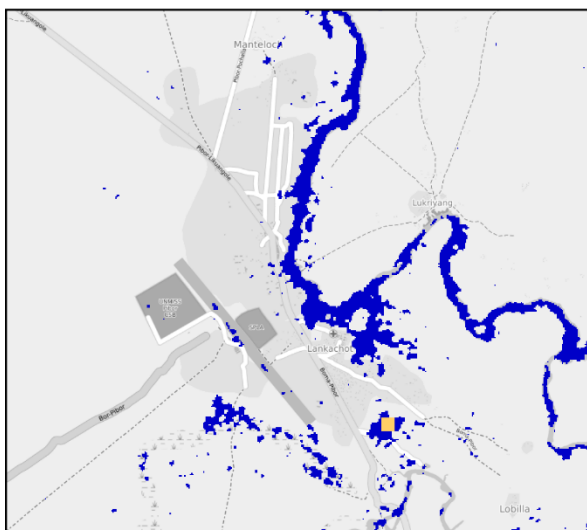
07.09.2021



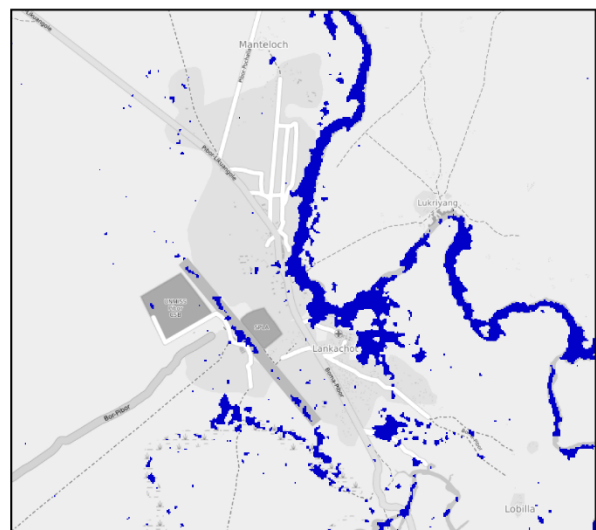
19.09.2021



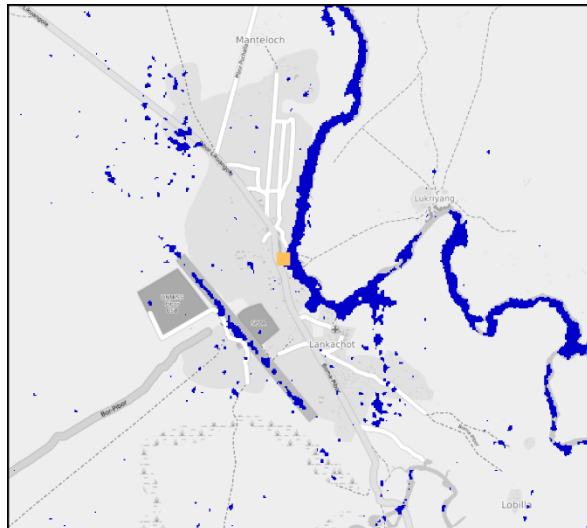
07.10.2021



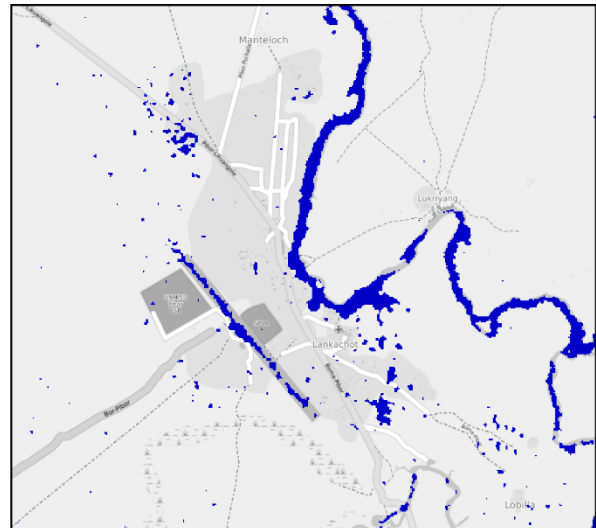
19.10.2021



25.10.2021



18.11.2021



24.11.2021

Figure 58. Pibor Post area: whole series of map products.

Table 16. Threshold and population: whole series statistics.

| Date | Threshold | Population affected |
|------------|-----------|---------------------|
| 06/06/2017 | -12.7169 | 117 |
| 05/08/2017 | -12.8763 | 0 |
| 11/08/2017 | -12.5597 | 0 |
| 17/08/2017 | -12.8771 | 0 |
| 22/09/2017 | -12.7137 | 15 |
| 28/09/2017 | -12.5572 | 0 |
| 04/10/2017 | -13.1967 | 15 |
| 10/10/2017 | -13.2042 | 40 |
| 22/10/2017 | -13.3563 | 38 |
| 03/11/2017 | -13.3554 | 66 |
| 09/11/2017 | -13.5228 | 74 |
| 15/11/2017 | -13.5248 | 52 |
| 21/11/2017 | -13.6859 | 0 |
| 27/11/2017 | -13.3681 | 0 |
| 03/12/2017 | -13.2078 | 10 |
| 09/12/2017 | -13.3624 | 133 |
| 15/12/2017 | -13.5204 | 305 |
| 21/12/2017 | -12.8805 | 662 |
| 27/12/2017 | -13.04 | 701 |
| 01/06/2018 | -12.56 | 0 |
| 25/06/2018 | -12.5588 | 32 |
| 13/07/2018 | -12.7186 | 0 |
| 06/08/2018 | -12.5573 | 0 |
| 05/10/2018 | -12.8736 | 55 |
| 10/11/2018 | -12.5656 | 265 |
| 18/10/2019 | -13.0288 | 130 |
| 24/10/2019 | -12.5647 | 491 |
| 30/10/2019 | -14.0068 | 358 |
| 05/11/2019 | -13.2072 | 644 |
| 11/11/2019 | -14.0014 | 344 |
| 17/11/2019 | -12.8776 | 450 |
| 23/11/2019 | -12.7259 | 208 |
| 29/11/2019 | -13.2037 | 209 |
| 05/12/2019 | -13.5226 | 34 |
| 11/12/2019 | -12.7154 | 12 |
| 17/12/2019 | -13.5255 | 12 |
| 23/12/2019 | -13.1983 | 19 |
| 29/12/2019 | -14.0012 | 0 |
| 10/01/2020 | -13.8497 | 19 |

| Date | Threshold | Population affected |
|------------|-----------|---------------------|
| 22/01/2020 | -13.2062 | 399 |
| 28/01/2020 | -13.5218 | 234 |
| 09/02/2020 | -13.3623 | 252 |
| 15/02/2020 | -13.6789 | 91 |
| 21/02/2020 | -13.363 | 379 |
| 04/03/2020 | -13.5214 | 535 |
| 08/06/2020 | -12.5536 | 0 |
| 07/08/2020 | -13.0348 | 0 |
| 19/08/2020 | -13.0406 | 0 |
| 31/08/2020 | -14.1622 | 55 |
| 12/09/2020 | -14.3246 | 230 |
| 24/09/2020 | -14.9594 | 157 |
| 30/09/2020 | -14.3236 | 509 |
| 06/10/2020 | -15.123 | 375 |
| 24/10/2020 | -15.118 | 354 |
| 30/10/2020 | -14.313 | 345 |
| 05/11/2020 | -14.7931 | 400 |
| 11/11/2020 | -14.6408 | 272 |
| 17/11/2020 | -14.9556 | 196 |
| 23/11/2020 | -14.8075 | 171 |
| 29/11/2020 | -12.7283 | 168 |
| 05/12/2020 | -14.8033 | 41 |
| 17/12/2020 | -14.642 | 0 |
| 23/12/2020 | -14.9568 | 0 |
| 04/01/2021 | -15.1229 | 0 |
| 10/01/2021 | -14.6389 | 0 |
| 22/01/2021 | -13.9996 | 0 |
| 28/01/2021 | -14.4839 | 0 |
| 03/02/2021 | -14.1579 | 9 |
| 21/07/2021 | -13.0407 | 0 |
| 02/08/2021 | -13.3608 | 0 |
| 08/08/2021 | -13.6765 | 15 |
| 07/09/2021 | -12.7152 | 0 |
| 19/09/2021 | -12.5542 | 15 |
| 07/10/2021 | -12.8819 | 30 |
| 19/10/2021 | -13.2008 | 12 |
| 25/10/2021 | -12.888 | 0 |
| 18/11/2021 | -13.3575 | 15 |
| 24/11/2021 | -13.8488 | 0 |

Annex 2

Single date application

```

/*Copyright (c) 2022 Francesco Collivignarelli
Permission is hereby granted, free of charge, to any person
obtaining a copy of this software and associated documentation
files (the "Software"), to deal in the Software without
restriction, including without limitation the rights to use,
copy, modify, merge, publish, distribute, sublicense, and/or sell
copies of the Software, and to permit persons to whom the
Software is furnished to do so, subject to the following
conditions:
The above copyright notice and this permission notice shall be
included in all copies or substantial portions of the Software.
THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND,
EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES
OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND
NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT
HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY,
WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING
FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR
OTHER DEALINGS IN THE SOFTWARE.*/

//----- OPTIONS -----
var datestart = '2020-09-30';
var dateend = '2020-10-01';
var relorb = 123;

var exporttodrive = true;

//Area of interest. Based on subnational administrative boundaries (NAME3_EN field)
//https://data.humdata.org/dataset/south-sudan-administrative-boundaries:
var unit_level = 'ADM3_EN';
var unit_name = 'Pibor';

var useaoi = true;
var ao = ee.FeatureCollection('users/fcolli/ssd_admbnda_adm3_imwg_nbs_2021')
    .filter(ee.Filter.eq(unit_level, unit_name));
//var ao = ee.FeatureCollection('users/fcolli/Extent_Edit');

//Db threshold to be considered as surface water (VV channel)
var otsu_method = true; //Change to false to use a 'Crisp' method
var otsu_type = 'constrained'; //Change to 'unconstrained' to use the ordinary otsu method

//Use external landcover
//ESA: https://esa-worldcover.org/en
//ESRI: https://www.arcgis.com/home/item.html?id=d6642f8a4f6d4685a24ae2dc0c73d4ac
//if false use OSM instead
var use_ext_landcover = true;

//Post-classification mode filter option
var cleaning = false;

//Population source
var population_source = 'WorldPOP'; //change to 'GHSL' if JRC layer has to be used instead (250m)
//change to 'WorldPOP' if WorldPOP constrained has to be used (100m)
//Display Options
var displayOriginal = false;
var displayTerrainCorrected = false;
var drawroads = true;
var drawplaces = true;
var addlabels = false;

//Load assets:
var table = ee.FeatureCollection('users/fcolli/roads_SS_20211011');
var table2 = ee.FeatureCollection('users/fcolli/places_points_SS_20211011_english');

//----- FLOOD MAPPING -----
//Prepare borders and convert to geometry, in the case ao = false, cut over SS country borders
var countries = ee.FeatureCollection('USDOS/LSIB_SIMPLE/2017');
var p = countries.filter(ee.Filter.eq('country_na', 'South Sudan'));

if (useaoi) {
    var pGeom = ao.geometry();
    //Map.centerObject(pGeom);
} else {
    var pGeom = p.geometry();
}

//Zoom to...:
//Map.setCenter(33.0700, 8.54, 12); //Nasir
//Map.setCenter(33.1300, 6.8000, 14); //Pibor

//Load the Sentinel-1 ImageCollection, filter accordingly.
var sentinel1 = ee.ImageCollection('COPERNICUS/S1_GRD')
    .filter(ee.Filter.listContains('transmitterReceiverPolarisation', 'VH'))

```

```

.filter(ee.Filter.listContains('transmitterReceiverPolarisation', 'VV'))
.filter(ee.Filter.eq('orbitProperties_pass', 'DESCENDING'))
.filter(ee.Filter.eq('instrumentMode', 'IW'))
.filterDate(datestart, dateend)
.filterMetadata('relativeOrbitNumber_start', 'equals', relorb);

//Apply terrain correction to avoid false positives
var noTerrainCorrection = ee.Image(sentinell.mosaic().clip(pGeom));

//Terrain correction
sentinell = sentinell.map(terrainCorrection);
var yesTerrainCorrection = ee.Image(sentinell.mosaic().clip(pGeom));

//Gamma-MAP
var s1_gammaMap = sentinell.map(gammaMap);
var s1_gammaMap = ee.Image(s1_gammaMap.mosaic().clip(pGeom));

//THRESHOLDING
var threshold = -15.0; //Crisp threshold, not used in the case of Otsu method

if (otsu_method) {
  if (use_ext_landcover) {
    //ESA
    //var landcover = ee.ImageCollection('ESA/WorldCover/v100')
    //  .first().clip(pGeom);
    //var mask = landcover.eq(80);
    //ESRI
    var landcover = ee.ImageCollection("projects/sat-io/open-datasets/landcover/ESRI_Global-LULC_10m")
      .mosaic().clip(pGeom);
    var mask = landcover.eq(1);
  } else {
    var landcover = ee.Image('users/fcolli/Waterway_4326_buff_reduced_fix_zip')
      .clip(pGeom);
    var mask = landcover.eq(1);
  }

  if (otsu_type == 'constrained') {
    Map.addLayer(mask, {
      min: 0,
      max: 1
    }, "mask", 0);
    var s1_gammaMap_msk = s1_gammaMap.mask(mask);
    threshold = applyOtsu(s1_gammaMap_msk);
  } else {
    threshold = applyOtsu(s1_gammaMap);
  }
}

var fld = yesTerrainCorrection.select('VV').lte(threshold);
var fldgamma = s1_gammaMap.select('VV').lte(threshold);

var mask = fld.eq(1);
fld = fld.updateMask(mask);

var maskgamma = fldgamma.eq(1);
fldgamma = fldgamma.updateMask(maskgamma);

//Cleaning filter
if (cleaning) {
  var fld_filtered = majorityFilter(fld);
  var fld_filtered_gamma = majorityFilter(fldgamma);
}

//----- IMPACT ON POPULATION -----
// Load JRC Global Human Settlement Population Density layer
// Resolution: 250. Number of people per cell is given.
// Alternatively, WorldPop WorldPop/GP/100m/pop at 100x100m.
if (population_source == 'GHSL') {
  var dictionary_pop = 'population_count';
  var population_count = ee.Image('JRC/GHSL/P2016/POP_GPW_GLOBE_V1/2015').clip(pGeom);
  var pop_resolution = 250;
} else {
  var dictionary_pop = 'population';
  var population_count = ee.Image(ee.ImageCollection('WorldPop/GP/100m/pop_age_sex_cons_unadj')
    .filter(ee.Filter.eq('country', 'SSD')).mosaic()).select('population').clip(pGeom);

  //var population_count =
ee.Image(ee.ImageCollection('WorldPop/GP/100m/pop').filterDate('2019').mosaic()).clip(pGeom);
  var pop_resolution = 100;
}

// Calculate the amount of exposed population
// get population projection
var popprojection = population_count.projection();

// Reproject flood layer to population layer scale
var flooded_res1 = fldgamma
  .reproject({

```

```

    crs: popprojection,
    scale: pop_resolution
  });

var description = 'surface_water_' + unit_name + '_' + datestart;
if (exporttodrive) {
  Export.image.toDrive({
    image: fldgamma,
    description: description,
    folder: 'Results',
    scale: 10,
    region: pGeom,
    maxPixels: 1e9
  });
}

// Create a raster showing exposed population only using the resampled flood layer
var population_exposed = population_count
  .updateMask(flooded_res1)
  .updateMask(population_count);

description = 'population_exposed_' + unit_name + '_' + datestart;
if (exporttodrive) {
  Export.image.toDrive({
    image: population_exposed,
    description: description,
    folder: 'Results',
    scale: pop_resolution,
    region: pGeom,
    maxPixels: 1e9
  });
}

//Sum pixel values of exposed population raster
var stats = population_exposed.reduceRegion({
  reducer: ee.Reducer.sum(),
  geometry: pGeom,
  scale: pop_resolution,
  maxPixels: 1e9
});

description = 'SAR_gamma_' + unit_name + '_' + datestart;
if (exporttodrive) {
  Export.image.toDrive({
    image: s1_gammaMap,
    description: description,
    folder: 'Results',
    scale: 10,
    region: pGeom,
    maxPixels: 1e9
  });
}

// get number of exposed people as integer
var number_pp_exposed = stats.getNumber(dictionary_pop).round();
print("Affected population:");
print(number_pp_exposed);

//Get map boundry
var cur_c = Map.getBounds();
var cur_rec = ee.Geometry.Rectangle(cur_c[0], cur_c[1], cur_c[2], cur_c[3]);

//Count only in the displayed area. Sum pixel values of exposed population raster
var stats2 = population_exposed.reduceRegion({
  reducer: ee.Reducer.sum(),
  geometry: cur_rec,
  scale: pop_resolution,
  maxPixels: 1e9
});

// get number of exposed people as integer
var number_pp_exposed2 = stats2.getNumber(dictionary_pop).round();
print("Affected population (display):");
print(number_pp_exposed2);

//Statistics
var maxpop = population_exposed.reduceRegion({
  reducer: ee.Reducer.max(),
  geometry: cur_rec,
  scale: pop_resolution
}).getNumber(dictionary_pop);
var minpop = population_exposed.reduceRegion({
  reducer: ee.Reducer.min(),
  geometry: cur_rec,
  scale: pop_resolution
}).getNumber(dictionary_pop);
var meapop = population_exposed.reduceRegion({
  reducer: ee.Reducer.mean(),
  geometry: cur_rec,

```

```

    scale: pop_resolution
  }).getNumber(dictionary_pop);
var stdpop = population_exposed.reduceRegion({
  reducer: ee.Reducer.stdDev(),
  geometry: cur_rec,
  scale: pop_resolution
}).getNumber(dictionary_pop);

//Client side stretch calculation
var stretchpopmin = 5; //minpop.getInfo() - 2.0*stdpop.getInfo();
var stretchpopmax = 50; //maxpop.getInfo() + 2.0*stdpop.getInfo();

//----- IMPACT ON ROADS -----
//Convert flood map to vectors. This layer is then used to filter the roads in the display section
var fld_converted = fld.reduceToVectors({
  reducer: ee.Reducer.countEvery(),
  geometry: cur_rec,
  scale: 100,
  tileSize: 16,
  maxPixels: 1e8
});

//----- DISPLAY -----
//SAR Data
//Some display
if (displayOriginal)
  Map.addLayer(noTerrainCorrection.select('VV'), {
    min: -17,
    max: -5
  }, "VV original", 0);
if (displayTerrainCorrected)
  Map.addLayer(terrainCorrection.select('VV'), {
    min: -17,
    max: -5
  }, "VV terrain correction", 0);
Map.addLayer(s1_gammaMap.select('VV'), {
  min: -17,
  max: -5
}, "VV Gamma-MAP");

if (drawroads) {
  //Roads All:
  var roads = ee.FeatureCollection(table).filterBounds(pGeom);
  Map.addLayer(roads.draw({
    color: 'green',
    strokeWidth: 1
  }), {}, 'Roads');

  //Roads hampered:
  var roads_h = roads.filterBounds(fld_converted);
  Map.addLayer(roads_h.draw({
    color: '#dd1c77',
    strokeWidth: 2
  }), {}, 'Roads affected');
}
//Flood maps
Map.addLayer(fldgamma, {
  'min': 0,
  'max': 1,
  'palette': ['white', 'blue']
}, 'Water gamma');
if (cleaning) {
  Map.addLayer(fld_filtered, {
    'min': 0,
    'max': 1,
    'palette': ['white', 'red']
  }, 'Water filtered');
}
//Places:
if (drawplaces) {
  var places = ee.FeatureCollection(table2).filterBounds(pGeom);
  Map.addLayer(places.filter(ee.Filter.eq('fclass', 'hamlet')), {
    color: '#b3e2cd',
    opacity: 0.2
  }, 'Hamlet');
  Map.addLayer(places.filter(ee.Filter.eq('fclass', 'village')), {
    color: '#fcd0ac',
    opacity: 0.2
  }, 'Village');
  Map.addLayer(places.filter(ee.Filter.eq('fclass', 'town')), {
    color: '#cbd5e8',
    opacity: 0.2
  }, 'Town');
  Map.addLayer(places.filter(ee.Filter.eq('fclass', 'suburb')), {
    color: '#f4cae4',
    opacity: 0.2
  }, 'Suburb');
  Map.addLayer(places.filter(ee.Filter.eq('fclass', 'city')), {

```

```

        color: '#e6f5c9',
        opacity: 0.2
    }, 'City');
}
//Population
var visParam = {
    min: stretchpopmin,
    max: stretchpopmax,
    palette: ['yellow', 'orange', 'red'],
};
Map.addLayer(population_exposed, visParam, 'Population Exposed (per cell)');

print(population_exposed);

//Add labels
if (addlabels) {
    //Add labels
    var text = require('users/gena/packages:text');

    // scale text font relative to the current map scale
    var scale = Map.getScale() * 1;

    var labels = places.map(function (feat) {
        feat = ee.Feature(feat);
        var name = ee.String(feat.get("name"));
        var centroid = feat.geometry().centroid();
        var t = text.draw(name, centroid, scale, {
            fontSize: 18,
            textColor: 'red',
            outlineWidth: 5,
            outlineColor: 'white'
        });
        return t;
    });

    labels = ee.ImageCollection(labels);
    Map.addLayer(labels, {}, 'Place labels');
}

//----- LEGEND -----
makelegendgraduated('Population', visParam);
makeLegend(drawroads);

```

Multitemporal application

```

/*Copyright (c) 2022 Francesco Collivignarelli
Permission is hereby granted, free of charge, to any person
obtaining a copy of this software and associated documentation
files (the "Software"), to deal in the Software without
restriction, including without limitation the rights to use,
copy, modify, merge, publish, distribute, sublicense, and/or sell
copies of the Software, and to permit persons to whom the
Software is furnished to do so, subject to the following
conditions:
The above copyright notice and this permission notice shall be
included in all copies or substantial portions of the Software.
THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND,
EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES
OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND
NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT
HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY,
WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING
FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR
OTHER DEALINGS IN THE SOFTWARE.*/
//----- OPTIONS -----
-----
var datestart = ee.Date('2021-08-01');
var dateend = ee.Date('2021-12-31');
var relorb = 123;

var datetoshow = 'none';
var showsun = false;

var exporttodrive = true;
var storeexportbatch = false;

var bmaxtresh = 0.75;

//Area of interest. Based on subnational administrative boundaries (NAME3_EN field)
//https://data.humdata.org/dataset/south-sudan-administrative-boundaries:
var unit_level = 'ADM3_EN';
var unit_name = 'Pibor';

var useaoi = true;
var ao = ee.FeatureCollection('users/fcolli/ssd_admbnda_adm3_imwg_nbs_2021')
    .filter(ee.Filter.eq(unit_level, unit_name));

```

```

//var aoi = ee.FeatureCollection('users/fcolli/Extent_Edith');

//----- FLOOD MAPPING -----
//Prepare borders and convert to geometry, in the case aoi = false, cut over SS country borders
var countries = ee.FeatureCollection('USDOS/LSIB_SIMPLE/2017');
var p = countries.filter(ee.Filter.eq('country_na', 'South Sudan'));

if (useaoi) {
  var pGeom = aoi.geometry();
  //Map.centerObject(pGeom);
} else {
  var pGeom = p.geometry();
}

var sentinell_tc = sentinell.map(terrainCorrection);
var sentinell_tc_mosaic = mosaicByDate(sentinell_tc.sort('system:time_start', true));
var sentinell_tc_mosaic_gamma = sentinell_tc_mosaic.map(gammaMap);
var resultOtsu = sentinell_tc_mosaic_gamma.map(applyOtsu);
var sentinell_tc_mosaic_gamma_bmax = resultOtsu.map(calculateBmax);
var resultOtsuGt = sentinell_tc_mosaic_gamma_bmax
  .filter(ee.Filter.gt('bmax', bmaxtresh))
  .filter(ee.Filter.lte('threshold', -12.5));
var changebands = sentinell_tc_mosaic_gamma_bmax.select(['Water'], ['Water']);
var changebandsGt = resultOtsuGt.select(['Water'], ['Water']);
var changebandsGtPop = changebandsGt.map(population);
var changebandsPop = changebands.map(population);
var changebandsPopSpatial = changebandsGt.map(population2);
var changebandsPopSpatialSelected = changebandsPopSpatial.filter(ee.Filter.gt('population_exposed', 0));

if (datetoshow != 'none') {
  var filtered = changebandsGt.filter(ee.Filter.eq('system:id', datetoshow));

  var cut = filtered.map(function (image) {
    return image.clip(pGeom);
  });

  Map.addLayer(cut.select('Water'), {
    'min': 0,
    'max': 1,
    'palette': ['white', 'blue']
  }, 'Water gamma');
}

var count = changebandsGt.size();

var sum = changebandsGt.reduce(ee.Reducer.sum()).clip(pGeom).divide(count).toDouble();

if (showsum) {
  Map.addLayer(sum, {
    'min': 0,
    'max': 500,
  }, 'Water occurrence');
}

var description1 = unit_name + '_' +
  'table_final';

var description2 = unit_name + '_' +
  'sum_final';

var description3 = unit_name + '_' +
  'table_gt_final';

if (exporttodrive) {
  Export.table.toDrive({
    collection: changebandsPop,
    description: description1,
    folder: 'Results',
    fileFormat: 'csv',
    selectors: ['system:id', 'threshold', 'population_exposed', 'bmax', 'm1', 'm2', 'p1', 'p2',
'sigmab', 'sigmat']
  });
  Export.table.toDrive({
    collection: changebandsGtPop,
    description: description3,
    folder: 'Results',
    fileFormat: 'csv',
    selectors: ['system:id', 'threshold', 'population_exposed', 'bmax', 'm1', 'm2', 'p1', 'p2',
'sigmab', 'sigmat']
  });
  Export.image.toDrive({
    image: sum,
    description: description2,
    folder: 'Results',
    scale: 10,
    region: pGeom,
    maxPixels: 1e9
  });
}

```



```

}

if (storeexportbatch) {
  var batch = require('users/fitoprincipe/geetools:batch');
  batch.Download.ImageCollection.toDrive(changebandsGt, 'ResultsWater', {
    name: '{date}_water',
    scale: 10,
    region: aoi,
    type: 'byte'
  });

  var batch2 = require('users/fitoprincipe/geetools:batch');
  batch2.Download.ImageCollection.toDrive(changebandsPopSpatialSelected, 'ResultsPop', {
    name: '{date}_pop',
    scale: 100,
    region: aoi,
    type: 'float'
  });
}

```

Utility and processing functions

```

/*Copyright (c) 2022 Francesco Collivignarelli
Permission is hereby granted, free of charge, to any person
obtaining a copy of this software and associated documentation
files (the "Software"), to deal in the Software without
restriction, including without limitation the rights to use,
copy, modify, merge, publish, distribute, sublicense, and/or sell
copies of the Software, and to permit persons to whom the
Software is furnished to do so, subject to the following
conditions:
The above copyright notice and this permission notice shall be
included in all copies or substantial portions of the Software.
THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND,
EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES
OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND
NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT
HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY,
WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING
FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR
OTHER DEALINGS IN THE SOFTWARE.*/
//----- TERRAIN CORRECTION AND SPECKLE FILTERING -----
// Originally implemented by Andreas Vollrath (ESA), inspired by Johannes Reiche (Wageningen)
function terrainCorrection(image) {
  var imgGeom = image.geometry();
  var srtm = ee.Image('USGS/SRTMGL1_003').clip(imgGeom); // 30m srtm
  var sigma0Pow = ee.Image.constant(10).pow(image.divide(10.0));

  // Article ( numbers relate to chapters)
  // 2.1.1 Radar geometry
  var theta_i = image.select('angle');
  var phi_i = ee.Terrain.aspect(theta_i)
    .reduceRegion(ee.Reducer.mean(), theta_i.get('system:footprint'), 1000)
    .get('aspect');

  // 2.1.2 Terrain geometry
  var alpha_s = ee.Terrain.slope(srtm).select('slope');
  var phi_s = ee.Terrain.aspect(srtm).select('aspect');

  // 2.1.3 Model geometry
  // reduce to 3 angle
  var phi_r = ee.Image.constant(phi_i).subtract(phi_s);

  // convert all to radians
  var phi_rRad = phi_r.multiply(Math.PI / 180);
  var alpha_sRad = alpha_s.multiply(Math.PI / 180);
  var theta_iRad = theta_i.multiply(Math.PI / 180);
  var ninetyRad = ee.Image.constant(90).multiply(Math.PI / 180);

  // slope steepness in range (eq. 2)
  var alpha_r = (alpha_sRad.tan().multiply(phi_rRad.cos())).atan();

  // slope steepness in azimuth (eq 3)
  var alpha_az = (alpha_sRad.tan().multiply(phi_rRad.sin())).atan();

  // local incidence angle (eq. 4)
  var theta_lia = (alpha_az.cos().multiply((theta_iRad.subtract(alpha_r)).cos())).acos();
  var theta_liaDeg = theta_lia.multiply(180 / Math.PI);
  // 2.2
  // Gamma_nought_flat
  var gamma0 = sigma0Pow.divide(theta_iRad.cos());
  var gamma0dB = ee.Image.constant(10).multiply(gamma0.log10());
  var ratio_l = gamma0dB.select('VV').subtract(gamma0dB.select('VH'));

  // Volumetric Model
  var nominator = (ninetyRad.subtract(theta_iRad).add(alpha_r)).tan();

```

```

var denominator = (ninetyRad.subtract(theta_iRad)).tan();
var volModel = (nominator.divide(denominator)).abs();

// apply model
var gamma0_Volume = gamma0.divide(volModel);
var gamma0_VolumeDB = ee.Image.constant(10).multiply(gamma0_Volume.log10());

// we add a layover/shadow mask to the original implementation
// layover, where slope > radar viewing angle
var alpha_rDeg = alpha_r.multiply(180 / Math.PI);
var layover = alpha_rDeg.lt(theta_i);

// shadow where LIA > 90
var shadow = theta_liaDeg.lt(85);

// calculate the ratio for RGB vis
var ratio = gamma0_VolumeDB.select('VV').subtract(gamma0_VolumeDB.select('VH'));

var output = gamma0_VolumeDB.addBands(ratio).addBands(alpha_r).addBands(phi_s).addBands(theta_iRad)
    .addBands(layover).addBands(shadow).addBands(gamma0dB).addBands(ratio_1);

return image.addBands(
    output.select(['VV', 'VH'], ['VV', 'VH']),
    null,
    true);
}

function powerToDb(img) {
    return ee.Image(10).multiply(img.log10());
}

function dbToPower(img) {
    return ee.Image(10).pow(img.divide(10));
}

function gammaMap(img) {
    var ksize = 3;
    var enl = 5;
    var bandNames = img.bandNames();

    // Convert image from dB to natural values
    var nat_img = dbToPower(img);

    // Square kernel, ksize should be odd (typically 3, 5 or 7)
    var weights = ee.List.repeat(ee.List.repeat(1, ksize), ksize);

    // ~~(ksize/2) does integer division in JavaScript
    var kernel = ee.Kernel.fixed(ksize, ksize, weights, ~~(ksize / 2), ~~(ksize / 2), false);

    // Get mean and variance
    var mean = nat_img.reduceNeighborhood(ee.Reducer.mean(), kernel);
    var variance = nat_img.reduceNeighborhood(ee.Reducer.variance(), kernel);

    // "Pure speckle" threshold
    var ci = variance.sqrt().divide(mean); // square root of inverse of enl

    // If ci <= cu, the kernel lies in a "pure speckle" area -> return simple mean
    var cu = 1.0 / Math.sqrt(enl);

    // If cu < ci < cmax the kernel lies in the low textured speckle area -> return the filtered value
    var cmax = Math.sqrt(2.0) * cu;

    var alpha = ee.Image(1.0 + cu * cu).divide(ci.multiply(ci).subtract(cu * cu));
    var b = alpha.subtract(enl + 1.0);
    var d =
mean.multiply(mean).multiply(b).multiply(b).add(alpha.multiply(mean).multiply(nat_img).multiply(4.0 *
enl));
    var f = b.multiply(mean).add(d.sqrt()).divide(alpha.multiply(2.0));

    var caster = ee.Dictionary.fromLists(bandNames, ee.List.repeat('float', 3));
    var img1 = powerToDb(mean.updateMask(ci.lte(cu))).rename(bandNames).cast(caster);
    var img2 = powerToDb(f.updateMask(ci.gt(cu)).updateMask(ci.lt(cmax))).rename(bandNames).cast(caster);
    var img3 = img.updateMask(ci.gte(cmax)).rename(bandNames).cast(caster);

    // If ci > cmax do not filter at all (i.e. we don't do anything, other than masking)
    var result = ee.ImageCollection([img1, img2, img3])
        .reduce(ee.Reducer.firstNonNull()).rename(bandNames);

    // Compose a 3 band image with the mean filtered "pure speckle", the "low textured" filtered and the
    unfiltered portions
    return result;
}

function makeLegend(drawroads) {
    //LEGEND
    var legend = ui.Panel({
        style: {
            position: 'bottom-left',

```

```

        padding: '8px 15px'
    }
});

// Create legend title
var legendTitle = ui.Label({
    value: 'Road condition',
    style: {
        fontWeight: 'bold',
        fontSize: '14px',
        margin: '0 0 4px 0',
        padding: '0'
    }
});

// Add the title to the panel
if (drawroads === true) {
    legend.add(legendTitle);
}

//
// Creates and styles 1 row of the legend.
var makeRow = function (color, name) {

    // Create the label that is actually the colored box.
    var colorBox = ui.Label({
        style: {
            backgroundColor: color,
            // Use padding to give the box height and width.
            padding: '8px',
            margin: '0 0 4px 0'
        }
    });

    // Create the label filled with the description text.
    var description = ui.Label({
        value: name,
        style: {
            margin: '0 0 4px 6px'
        }
    });

    // return the panel
    return ui.Panel({
        widgets: [colorBox, description],
        layout: ui.Panel.Layout.Flow('horizontal')
    });
};

// Palette with the colors
var palette = ['green', '#dd1c77'];

// name of the legend
var names = ['Good', 'Flooded'];

// Add color and and names
for (var i = 0; i < 2; i++) {
    legend.add(makeRow(palette[i], names[i]));
}

// add legend to map (alternatively you can also print the legend to the console)
if (drawroads === true) {
    Map.add(legend);
}

//LEGEND
// set position of panel
legend = ui.Panel({
    style: {
        position: 'bottom-left',
        padding: '8px 15px'
    }
});

// Create legend title
legendTitle = ui.Label({
    value: 'Water extent',
    style: {
        fontWeight: 'bold',
        fontSize: '14px',
        margin: '0 0 4px 0',
        padding: '0'
    }
});

// Add the title to the panel
legend.add(legendTitle);

```

```

// Creates and styles 1 row of the legend.
makeRow = function (color, name) {

    // Create the label that is actually the colored box.
    var colorBox = ui.Label({
        style: {
            backgroundColor: color,
            // Use padding to give the box height and width.
            padding: '8px',
            margin: '0 0 4px 0'
        }
    });

    // Create the label filled with the description text.
    var description = ui.Label({
        value: name,
        style: {
            margin: '0 0 4px 6px'
        }
    });

    // return the panel
    return ui.Panel({
        widgets: [colorBox, description],
        layout: ui.Panel.Layout.Flow('horizontal')
    });
};

// Palette with the colors
palette = ['blue'];

// name of the legend
names = ['Surface water'];

// Add color and names
for (i = 0; i < 1; i++) {
    legend.add(makeRow(palette[i], names[i]));
}

// add legend to map (alternatively you can also print the legend to the console)
Map.add(legend);

return;
}

function makelegendgraduated(title, viz) {
    function makeColorBarParams(palette) {
        return {
            bbox: [0, 0, 1, 0.1],
            dimensions: '100x10',
            format: 'png',
            min: 0,
            max: 1,
            palette: palette,
        };
    }

    // Create the color bar for the legend.
    var colorBar = ui.Thumbnail({
        image: ee.Image.pixelLonLat().select(0),
        params: makeColorBarParams(viz.palette),
        style: {
            stretch: 'horizontal',
            margin: '0px 8px',
            maxHeight: '24px'
        }
    });

    // Create a panel with three numbers for the legend.
    var legendLabels = ui.Panel({
        widgets: [
            ui.Label(viz.min, {
                margin: '4px 8px'
            }),
            ui.Label(
                ((viz.max - viz.min) / 2 + viz.min), {
                margin: '4px 8px',
                textAlign: 'center',
                stretch: 'horizontal'
            }),
            ui.Label(viz.max, {
                margin: '4px 8px'
            })
        ],
        layout: ui.Panel.Layout.flow('horizontal')
    });

    var legendTitle = ui.Label({

```

```

        value: title,
        style: {
            fontWeight: 'bold'
        }
    });

    // Add the legendPanel to the map.
    var legendPanel = ui.Panel([legendTitle, colorBar, legendLabels]);
    //move to bottom-left
    legendPanel.style().set({
        position: 'bottom-left'
    });
    Map.add(legendPanel);
    return;
}

//----- MAJORITY FILTER -----
function majorityFilter(classification) {
    //Scale to set for filtering purpose
    var SCALE = 10;
    //calculate mode
    var mode = classification.focal_mode().reproject('EPSG:4326', null, SCALE);

    //Weighted smoothing
    //using a 3x3 window
    //euclidean distance weighting from corners
    var weights = [[1, 2, 1],
        [2, 3, 2],
        [1, 2, 1]];

    //create a 3x3 kernel with the weights
    //kernel W and H must equal weight H and H
    var kernel = ee.Kernel.fixed(3, 3, weights);

    //apply mode on neighborhood using weights
    //and force operations to be done at native scale.
    //cleans only the flood map generated from the not filtered data.
    var filtered = classification.reduceNeighborhood({
        reducer: ee.Reducer.mode(),
        kernel: kernel,
    }).reproject('EPSG:4326', null, SCALE);

    var maskfil = filtered.eq(1);
    filtered = filtered.updateMask(maskfil);

    return filtered;
}

function constructHistChart(histogram, threshold) {
    var counts = ee.List(ee.Dictionary(histogram).get('histogram'));
    var buckets = ee.List(ee.Dictionary(histogram).get('bucketMeans'));
    // construct array for visualization of threshold in chart
    var segment = ee.List.repeat(0, counts.size());
    var maxFrequency = ee.Number(counts.reduce(ee.Reducer.max()));
    var threshIndex = buckets.indexOf(threshold);
    segment = segment.set(threshIndex, maxFrequency);
    var histChart = ui.Chart.array.values(ee.Array.cat([counts, segment], 1), 0, buckets)
        .setSeriesNames(['Values', 'Threshold'])
        .setChartType('ColumnChart');

    return histChart;
}

function applyOtsu(img) {
    // Compute the histogram of the VV band. The mean and variance are only FYI.
    var histogram = img.select('VV').reduceRegion({
        reducer: ee.Reducer.histogram(255, 0.005, 1e6)
            .combine('mean', null, true)
            .combine('variance', null, true),
        geometry: pGeom,
        scale: 10,
        bestEffort: true,
    });

    // Chart the histogram
    //print(Chart.image.histogram(img.select('VV'), geometry, 10));

    // Return the DN that maximizes interclass variance in VV (in the region).
    var otsu = function (histogram) {
        var counts = ee.Array(ee.Dictionary(histogram).get('histogram'));
        var means = ee.Array(ee.Dictionary(histogram).get('bucketMeans'));
        var size = means.length().get([0]);
        var total = counts.reduce(ee.Reducer.sum(), [0]).get([0]);
        var sum = means.multiply(counts).reduce(ee.Reducer.sum(), [0]).get([0]);
        var mean = sum.divide(total);

        var indices = ee.List.sequence(1, size);

```

```

// Compute between sum of squares, where each mean partitions the data.
var bss = indices.map(function (i) {
  var aCounts = counts.slice(0, 0, i);
  var aCount = aCounts.reduce(ee.Reducer.sum(), [0]).get([0]);
  var aMeans = means.slice(0, 0, i);
  var aMean = aMeans.multiply(aCounts)
    .reduce(ee.Reducer.sum(), [0]).get([0])
    .divide(aCount);
  var bCount = total.subtract(aCount);
  var bMean = sum.subtract(aCount.multiply(aMean)).divide(bCount);
  return aCount.multiply(aMean.subtract(mean).pow(2)).add(
    bCount.multiply(bMean.subtract(mean).pow(2)));
});

//print(ui.Chart.array.values(ee.Array(bss), 0, means));
// Return the mean value corresponding to the maximum BSS.
return means.sort(bss).get([-1]);
};

print(histogram.get('VV_histogram'));
var threshold = otsu(histogram.get('VV_histogram'));

description = unit_name + ' ' + datestart + ' Sentinel-1 VV';
var chart = constructHistChart(histogram.get('VV_histogram'), threshold)
  .setOptions({
    title: description,
    hAxis: {
      title: 'Values',
    },
    vAxis: {
      title: 'Count'
    }
  });

print("Calculated threshold:", threshold);
print('Thresholding histogram:', chart);

return threshold;
}

/*-----MOSAIC BY DATE-----*/
function mosaicByDate(imcol) {
  // imcol: An image collection
  // returns: An image collection
  var imlist = imcol.toList(imcol.size());

  var unique_dates = imlist.map(function (im) {
    return ee.Image(im).date().format("YYYY-MM-dd");
  }).distinct();

  var mosaic_imlist = unique_dates.map(function (d) {
    d = ee.Date(d);

    var im = imcol
      .filterDate(d, d.advance(1, "day"))
      .mosaic();

    return im.set(
      "system:time_start", d.millis(),
      "system:id", d.format("YYYY-MM-dd"));
  });

  return ee.ImageCollection(mosaic_imlist);
}

/*-----BIMODALITY TEST-----*/
//BMAX STARTS HERE
function calculateBmax(img) {
  var landcover = ee.ImageCollection("projects/sat-io/open-datasets/landcover/ESRI_Global-LULC_10m")
    .mosaic();
  var mask = landcover.eq(1).clip(pGeom);
  var img_msk = img.mask(mask);

  var initialTresh = img_msk.get('threshold');
  var tempImg = ee.Image.constant(initialTresh);

  var water = img_msk.select('VV').lte(tempImg);
  var p1 = water.reduceRegion({
    reducer: ee.Reducer.mean(),
    geometry: pGeom,
    bestEffort: true,
    scale: 10,
  }).get('VV');

  p1 = ee.Number(ee.Algorithms.If(p1, p1, 0.99));
  var p2 = ee.Number(1).subtract(p1);

```

```

var m = img_msk.select('VV').updateMask(water).rename('m1').addBands(
  img_msk.select('VV').updateMask(water.not()).rename('m2'));

var mReduced = m.reduceRegion({
  reducer: ee.Reducer.mean(),
  geometry: pGeom,
  bestEffort: true,
  scale: 10,
});

var m1 = ee.Number(mReduced.get('m1'));
var m2 = ee.Number(mReduced.get('m2'));

m1 = ee.Number(ee.Algorithms.If(m1, m1, -20));
m2 = ee.Number(ee.Algorithms.If(m2, m2, -5));

var sigmab = p1.multiply(p2.multiply(m1.subtract(m2).pow(2)));
var sigmat = ee.Number(img_msk.reduceRegion({
  reducer: ee.Reducer.variance(),
  geometry: pGeom,
  bestEffort: true,
  scale: 10,
}).get('VV'));

sigmat = ee.Number(ee.Algorithms.If(sigmat, sigmat, 2));
var bmax = sigmab.divide(sigmat);

var id = img_msk.get('system:id');

return img.set({
  'date': id,
  'threshold': initialTresh,
  'p1': p1,
  'p2': p2,
  'm1': m1,
  'm2': m2,
  'sigmab': sigmab,
  'sigmat': sigmat,
  'bmax': bmax
});
}

function population(img) {
  var dictionary_pop = 'population';
  var population_count = ee.Image(ee.ImageCollection('WorldPop/GP/100m/pop_age_sex_cons_unadj')
    .filter(ee.Filter.eq('country', 'SSD')).mosaic()).select('population').clip(pGeom);

  //var population_count =
  ee.Image(ee.ImageCollection('WorldPop/GP/100m/pop').filterDate('2019').mosaic()).clip(pGeom);
  var pop_resolution = 100;

  var popprojection = population_count.projection();

  var reprojected = img.select('Water')
    .reproject({
      crs: popprojection,
      scale: pop_resolution
    });

  var population_exposed = population_count
    .updateMask(reprojected)
    .updateMask(population_count);

  var stats2 = population_exposed.reduceRegion({
    reducer: ee.Reducer.sum(),
    geometry: pGeom,
    scale: pop_resolution,
    maxPixels: 1e9
  });

  var numberExp = stats2.getNumber(dictionary_pop).round();

  img.addBands(
    population_exposed);

  return img.set({
    'population_exposed': numberExp
  });
}

function population2(img) {
  var dictionary_pop = 'population';
  var population_count = ee.Image(ee.ImageCollection('WorldPop/GP/100m/pop_age_sex_cons_unadj')
    .filter(ee.Filter.eq('country', 'SSD')).mosaic()).select('population').clip(pGeom);

  //var population_count =
  ee.Image(ee.ImageCollection('WorldPop/GP/100m/pop').filterDate('2019').mosaic()).clip(pGeom);
  var pop_resolution = 100;

```

```
var popprojection = population_count.projection();

var reprojected = img.select('Water')
  .reproject({
    crs: popprojection,
    scale: pop_resolution
  });

var population_exposed = population_count
  .updateMask(reprojected)
  .updateMask(population_count);

var stats2 = population_exposed.reduceRegion({
  reducer: ee.Reducer.sum(),
  geometry: pGeom,
  scale: pop_resolution,
  maxPixels: 1e9
});

var numberExp = stats2.getNumber(dictionary_pop).round();

var id = img.get('system:id');

return population_exposed.set({
  'date': id,
  'population_exposed': numberExp,
});
}
```