Automating Global Geospatial Data Set Analysis

Visualizing flood disasters in the cities of the Global South

Ohto Nygren

NYGREN OHTO: Automating Global Geospatial Dataset Analysis – Visualizing flood disasters in the cities of the Global South

Flooding is the most devastating natural hazard affecting tens of millions of people yearly and causing billions of USD dollars in damages globally. The people most affected by flooding globally are those with a high level of everyday vulnerability and limited resources for flood protection and recovery. Geospatial data from the Global South is severely lacking, and geospatial proficiency needs to be improved at a local level so that geospatial data and data analysis can be efficiently utilized in disaster risk reduction schemes and urban planning in the Global South. This thesis focuses on the use of automated global geospatial dataset analysis in disaster risk reduction in the Global South by using the Python programming language to produce an automated flood analysis and visualization model. In this study, the automated model was developed and tested in two, highly relevant cases: in the city of Bangkok, Thailand, and in the urban area of Tula de Allende, Mexico.

The results of the thesis show that with minimal user interaction, the automated flood model ingests flood extent and depth data produced by ICEYE, a global population estimation raster produced by the German Aerospace Agency (DLR) and OpenStreetMap (OSM) data, performs multiple relevant analyses of these data, and produces an interactive map highlighting the severity and effects of a flooding event. The automated flood model performs consistently and accurately while producing key statistics and standardized visualizations of flooding events which offers first responders a very fast first estimation of the scale of a flooding event and helps plan an appropriate response anywhere around the globe.

Global geospatial data sets are often created to examine large scale geographical phenomena; however, the results of this thesis show that they can also be used to analyze detailed local-level phenomena when paired together with supporting data. The advantage of using global geospatial data sets is that when sufficiently accurate and precise, they remove the most time-consuming part of geospatial analysis: finding suitable data. Fast reaction is of utmost importance in the first hours of a natural hazard like flooding, thus, automated analysis produced on a global scale could significantly help international humanitarian aid and first responders. Using an automated model also standardizes the results removing human errors and interpretation from the results enabling the accurate comparison of historical flood data in due time.

Tulvat ovat luonnonilmiöihin liittyvistä riskeistä tuhoisimpia, ja ne vaikuttavat kymmeniin miljooniin ihmisiin vuosittain sekä aiheuttavat miljardien dollarien vahingot maailmanlaajuisesti. Tulvista kärsivät usein maailmanlaajuisesti ne ihmiset, jotka ovat jo ennestään haavoittuvia ja joilla on suhteellisesti heikoimmat keinot suojautua tulvilta ja selviytyä tulvan aiheuttamista tuhoista. Monissa globaalin etelän maissa on niukasti paikkatietoaineistoa ja paikkatieto-osaamista on syytä lisätä erityisesti paikallisella tasolla, jotta paikkatietoaineistoa ja analyysin hyödynnettävyyttä voidaan parantaa katastrofiriskien vähentämissuunnitelmissa sekä kaupunkisuunnittelussa globaalissa etelässä. Tämä opinnäytetyö keskittyy automatisoidun globaalin paikkatietoaineiston analyysin hyödyntämiseen katastrofiriskien vähentämisessä globaalissa etelässä käyttämällä Python-ohjelmointikieltä automatisoidun tulva-analyysi- ja visualisointimallin tuottamiseen. Tässä tutkimuksessa automatisoitua mallia kehitettiin ja testattiin kahdessa tulvariskien kannalta erittäin relevantissa tapauksessa: Bangkokissa, Thaimaassa ja Tula de Allende:n kaupunkialueella, Meksikossa.

Tämän tutkielman tulokset osoittavat, että automatisoitu tulvamalli osaa lukea ICEYE:n tuottaman tulvan laajuus- ja syvyysaineiston, Saksan ilmailu- ja avaruuskeskuksen (DLR) tuottaman maailmanlaajuisen väestönarviorasterin, sekä OpenStreetMap (OSM) -aineiston, suorittaa aineistolle tulvan tuhojen tulkinnan kannalta olennaisia analyyseja, ja tuottaa lopputuloksena interaktiivisen kartan, joka korostaa tulvatapahtuman laajuutta ja vaikutuksia. Automatisoitu tulvamalli toimii johdonmukaisesti ja tuottaa tilastoja sekä standardoituja visualisointeja tulvatapahtumista, mikä tarjoaa ensivastehenkilöille erittäin nopean ensimmäisen arvion tulvatapahtuman laajuudesta. Tämä auttaa kohdentamaan pelastustoimenpiteitä riskitilanteessa vaihtelevissa ympäristöissä eri puolilla maailmaa.

Globaalit paikkatietoaineistot luodaan usein laajojen maantieteellisten ilmiöiden tutkimiseen, mutta tämän tutkielman tulokset osoittavat kuitenkin, että niillä voidaan analysoida myös hyvin paikallistason ilmiöitä, kun ne yhdistetään muihin relevantteihin tietolähteisiin. Globaalien paikkatietoaineistojen käytön etuna on, että ollessaan riittävän tarkkoja ne poistavat paikkatietoanalyysin aikaa vievimmän osan: sopivan tiedon löytämisen. Nopea reagointi on äärimmäisen tärkeää luonnonuhkien, kuten tulvien, ensimmäisinä tunteina ja kansainvälisen humanitaarisen avun ja ensivastetoimijoiden tulisi hyödyntää maailmanlaajuisia automatisoituja analyysejä. Automaattinen malli myös standardoi tulokset poistaen tuloksista inhimilliset virheet ja tulkinnat, mikä mahdollistaa historiallisten tulvatietojen tarkan vertailun.

Table of contents

1   Introduction

" The flood came out. No one could see anyone else; they could not be recognized in the catastrophe. The flood roared like a bull, like a wild ass screaming in the winds. The darkness was total, there was no sun."

Excerpt of the epic of Atra-Hasis dated to the 17th century BC

Floods have always existed and will always exist as long as there is water on Earth. They are a part of the ecosystems of our planet. Even though floods are often spoken of in a very negative light, floods are a natural phenomenon and they do not become natural disasters until they affect the physical and social vulnerabilities of people and societies (Jackson et al. 2017). A disaster is often defined as an unplanned event in which the needs of the affected communities outweigh the available resources (Furin 2018). The interaction between floods and humans has always captured the imagination of the human population and as a sign of this interest, floods are mentioned even in some of the oldest written texts found on Earth, including the epic of Atra-Hasis, the epic of Gilgamesh, and the Bible. In ancient texts, floods are often depicted as devastating events brought on by the gods that wipe out all the life to punish the humankind and restart life on Earth. Although floods have never actually reached such biblical proportions that they would have covered the entire planet, floods are the most devastating natural hazards and the leading cause of death by natural hazards (Doocy et al. 2013; Mosavi et al. 2018).

Understanding different perspectives to explain flooding as a phenomenon and to develop strategies for flood disaster reduction is important because flooding causes loss of human lives and extensive damage to infrastructure every year around the globe. The impacts of floods can also lead to physical and mental health problems such as infectious and parasitic diseases, nutritional diseases, diseases of the circulatory system, depression, post-traumatic stress disorder, and increased monthly health expenditure (Nomura et al. 2016; Puteh et al. 2019). Floodwaters carry around trash and sewage water and contaminate drinking water, causing cholera outbreaks (Penrose et al. 2010). Due to climate change floods have become more frequent, intensive, and more devastating than before and this trend is predicted to only get worse in the near future, affecting especially the poorest and most vulnerable areas (Depietri et al. 2011; Alfieri et al. 2016; Krellenberg et al. 2016; Hyndman & Hyndman 2017). Population

growth, increasing urbanization, and extensive changes in forest cover and land use are also contributing to the problem (Doocey et al. 2013).

According to the Swiss Re Institute, floods caused economic losses amounting to USD 82 billion globally in the year 2021, which means a 205 % increase from the USD 40 billion estimation in 2016 by the OECD (Financial Management…2016, Extreme flood events…2022). It can take years for affected cities and communities to recover from a flood disaster in developed countries, let alone in developing countries, or the countries in the so-called Global South, especially since poor people cannot afford insurance against flood damage. Humans have always tried to mitigate, prevent, and predict flood disasters and other natural hazards. The strategies of flood disaster risk reduction have evolved; and include today intensified global cooperation, different kinds of civil-society and grassroots initiatives, sophisticated software systems, using new kinds of digital data, and geographic information systems (GIS) (Al-Tahir & Mahabir 2014). One global problem is that geospatial information is relatively scarcely available in many rapidly developing countries. This concerns especially information related to vulnerable informal settlements (Falco et al. 2019; Young et al. 2020).

Mapping flood risk areas quickly and reliably is essential to reducing damages and loss of life caused by flooding. If maps of flooding extents are produced quickly enough, they can be used by first responders and humanitarian aid organizations, which help in efforts of recovery from large-scale disasters (e.g., the Red Cross). During times of crisis, accurate information is as vital as clean water and food (Meier 2011). If these flooding extent maps are produced with a similar template consistently, historical flood risk areas can be mapped for risk mitigation planners and city officials to be used in risk reduction strategies and urban planning. Simultaneously, the need and utility of more open-source geospatial data can be tracked. The critical risk areas can be identified for future improvements and essential services for these areas can be planned to withstand flooding either by reinforcing the buildings, constructing flood preventative measures, increasing people's awareness of the risks, or by relocating the settlements just outside the critical risk area.

The aim of this Master's thesis is to utilize Python code and open-source geospatial web tools (e.g., Kepler.gl, Mapbox GL, and Leaflet) to create and deploy a dynamic web visualization model that ingests terrain data, population data, ground data, and flood analyses/predictions to assess impacted areas for adequate First Response efforts and timely action. This aim is planned

to be achieved by creating an automated, standardized, and rapid disaster mapping processes flow using digital global data and open-source tools. If implemented and used in real-life cases, this process would support the United Nation's Sustainable Development Goal (SDG) #11: *Make cities and human settlements inclusive, safe, resilient, and sustainable* (Make cities…2021). The solution developed in this thesis uses the World Settlement Footprint (WSF) population data provided by the German Aerospace Center (DLR) to get estimations of how many people are affected by the flooding events, SAR satellite-based data of the space company ICEYE to obtain flooding extent and flood depths, and OpenStreetMap (OSM) data to link the flood and population information to essential services and the number of inundated buildings.

The detailed research questions of this study are the following:

1. How to combine multi-source geospatial data to produce automated flood risk assessments and visualizations in cities of the Global South?

2. What is the accuracy and reliability of these spatio-temporal estimates?

3. What is the practical value and challenge of the developed automatization approach in operative disaster risk reduction?

The geovisualization to be created in this thesis will utilize three globally scalable data sources. The combination of using SAR satellite data from ICEYE, the WSF dataset from the DLR, and OMS data will make the process repeatable anywhere in the world. The final accuracy will depend mostly on how accurate and complete the OSM data from the region in question is. ICEYE can monitor the required conditions anywhere in the world with millimeter-level precision and the WSF data has been validated by crowdsourcing and using 900 000 validation samples. The accuracy and precision of OSM data are highly dependent on how actively the geographical location has been mapped. The automated geovisualization will be created by using the Python coding language to create a script that produces an interactive, locally hosted flood map of the flooding event. The process performs consistently, and it is repeatable anywhere in the world without any user input apart from inserting the population raster data and the flood raster data.

2 Theoretical framework

2.1 Disaster Risk Reduction

Disaster risk reduction (DRR) is a systematic approach that aims to identify, prevent, and reduce damage caused by natural hazards such as floods, earthquakes, droughts, and tropical cyclones (What is Disaster…2021; Disaster risk…2021). Although natural hazards are often considered as natural disasters, it is important to understand that "natural disasters" are not totally natural; rather, they are shaped by a dynamic interplay of biophysical and sociopolitical processes and have multiple economic, social, and psychological effects, especially on vulnerable urban neighborhoods and rural communities in the Global South (Goh, 2019). Understanding the different dimensions of flood disasters enables the identification of why these natural processes are so harmful especially to already vulnerable populations and how to mitigate or prevent their harmful effects (Silva & Mena 2020).

In 2015 all the United Nations (UN) member states accepted the Sendai framework for disaster risk reduction for 2015–2030 (Sendai Framework…2015). The agreement defined four main priorities: 1) understanding disaster risk, 2) strengthening disaster risk governance to manage disaster risk, 3) investing in disaster risk reduction for resilience, and 4) enhancing disaster preparedness for effective response, and for "building back better" in recovery, rehabilitation, and reconstruction.

In the academic literature, there has been considerable discussion and debate on how to best approach DRR. Currently, there are two main, somewhat different, and contradicting paradigms to explain it, including 1) the hazard paradigm and 2) the vulnerability paradigm (Gaillard & Mercer 2012). In the following, I will present the main similarities and the main differences between these approaches. Natural hazards are extremely complex, and thus it is important to understand different perspectives to explain them and different methods to develop disaster risk reduction strategies (Tierney 2007). The focus of this analysis is mainly on the ways different approaches explain and understand the drivers of flood disasters and the role of disaster risk reduction in mitigating the effects of flooding events.

2.1.1 Hazard paradigm

The hazard paradigm mainly considers that disasters are a result of insufficient preparation and people exposed to them do not take extreme natural hazards seriously enough (Gaillard & Mercer 2012). The hazard paradigm focuses mostly on the hazards themselves and how to prevent them by building better infrastructure and developing early warning mechanisms. The dominant DRR initiatives, grounded on the hazard paradigm, rely frequently on a fairly top-down-style approach, with a heavy emphasis on technology, scientific research, and national government intervention.

Examples of the flood management projects, grounded on the hazard paradigm, include huge projects of flood infrastructure, implemented in various parts of the Global South since the 1950s, such as Cambodia, Laos, Mexico, and Bolivia (Molle 2009; Molle et al. 2009). Many of these projects were based on the idea of mastering natural hazards through environmental engineering and hard infrastructure, consisting of the construction of floodwalls, embankments, levees, and borders. Since the 1990s, such projects have been increasingly criticized for trying to control nature by technological means, without sufficiently taking into account the involved ecological and social aspects (Nygren 2016; Coates 2022).

More recent examples of flood management methods developed based on the hazard paradigm are recent room for rivers initiatives and plans for sponge cities. Initially, many room for rivers projects were developed in the Netherlands and later on, they have been suggested to be used in the Global South situations, like in Dar es Salaam, Tanzania Transforming Tanzania's…2019). The room for river flood prevention strategies aim to give rivers in urban areas space to extend more freely during the period of intensive flooding, by creating city parks and other kinds of green infrastructure on the flood plains (Klijn et al. 2018). This would reduce the devastating consequences of flooding for humans and infrastructure and reduce the probability of embankment failures (Figure 1).

Figure 1. Proposed changes for the Msimbazi River Basin, Tanzania (Source: The Msimbazi opportunity 2019)

However, in Dar es Salaam as well as in many other cities of the Global South, there is a risk that these kinds of strategies would indicate removing the densely populated informal settlements from the critical risk areas, which could then promote social exclusion and increase the everyday vulnerability of the poor population. It could also stimulate the building of new housing complexes with high-quality services and flood preventative infrastructure in such areas, pricing out even the middle-income residents from the area, while increasing already existing social segregation.

The sponge cities approaches aim to reduce urban flooding by mimicking nature in urban planning and reducing the amount of gray infrastructure, such as roads, dams, pipes, and concrete walls, by replacing them with green infrastructures, like retention ponds, parks, wetlands, and natural waterways (Gies 2018). The concept is known as sponge cities in China, green infrastructure in Europe, and low-impact urban development in the United States. The idea is to remove the impermeable surfaces in urban environments and to increase natural surfaces, which allows water to soak into the ground while filtering pollutants from the water. This approach soaks water and holds it roughly in place like a sponge helping to reduce urban runoff which is a major reason behind urban flooding and water pollution. An example of this approach is a proposed park in Bangkok, Thailand which would collect and store excess water during the rainy season in a city where many of the canals have been paved over and flooding is a reoccurring problem (Gray 2018).

Flood prevention measures and strategies developed and utilized by leaning on the hazard paradigm tend to focus heavily on protecting the economically most important areas and areas where the most influential and rich people live in. Flood protection projects such as dredging

and expanding canals or constructing flood barriers are often implemented on the lands that are inhabited by the poor, while plans to build such infrastructure on lands owned by the rich and the powerful are rejected (Win 2017). Flood prevention strategies such as blocking and diverting floodwaters to other areas often most heavily affect the local population that is the most vulnerable and like in the case of Bangkok, the inner city has been protected at the expense of the marginalized population living in the peripheries and in informal settlements (Marks 2015). The focus on protecting the urban elites by water diverging channels often makes the floods even worse in the low-lying rural areas and near the water-transfer canals as water is diverted to these areas either intentionally or as an unintended and overlooked consequence in order to save the cities.

Approaching DRR from the perspective, based on the hazard paradigm, requires considerable funding and expertise, and often excludes the most vulnerable populations. Local people and communities are often seen as helpless in facing natural hazards and there is a persisting attitude that they have nothing of value to give to DRR as they do not have sufficient scientific knowledge and qualifications and they lack the needed technical expertise and economic resources to be engaged in the planning and building of large-scale flood prevention structures. An important aspect to consider is also the fact that according to the recent evaluation by the Swiss Re Institute, in developing countries only 5 percent of severe flood losses were insured in 2021, compared to 34 percent in advanced economies (Extreme flood events…2022). Insurance is expensive and a luxury that poor people cannot afford, and, in some cases, they probably would not even be able to get it, due to the fact that they live in informal settlements in areas prone to flooding which are considered areas of high-risk liability and thus inhabitants' applications for flood insurance are often rejected by insurance companies.

2.1.2   Vulnerability paradigm

According to the vulnerability paradigm, natural hazards often affect the population groups who are already the most vulnerable and least equipped to manage or mitigate natural hazards. These people are often poor and do not have access to resources, such as adequate flood prevention infrastructure, water and sanitation services, or early detection applications (Greater Impact…2017). In addition, these people often live in the areas most probable to be affected by floods because they cannot afford to live anywhere else, and the wealthier population does not want to live in these risk-prone areas (Uitto & Shaw 2016).

The vulnerability paradigm has emphasized the unequal distribution of effects of floods and other natural hazards in many parts of the Global South. This especially because due to rapid urbanization and growing population rates many cities have expanded into ecologically and hydrologically sensitive floodplains, low-lying areas, riverbanks, and coastal zones. This kind of urban expansion often occurs in the form of informal settlements, for example in many African cities, informal construction makes up most of the housing being built (Fekade 2000; Watson 2020). Cities in the Global South also often lack active city planning and enforced building and environmental regulations. In addition, many residents living in risk-prone areas usually cannot afford engineers or other kinds of experts and qualified professionals to help them carry out flood protection projects, but rather hire self-taught amateurs, who might not have adequate resources, expertise, or materials to build proper flood protection structures (Sakijege et al. 2014). In some cases, the structures built by the hired amateurs or by the occupants themselves, unfortunately, cause the damage to be even more severe or last longer by trapping the water inside the floodwalls for example. Furthermore, due to a lack of large-scale coordination and supervision, "do-it-yourself"-flood protections built by individual households can make some areas overflow even more or cause the overall flooding to be much worse due to the cumulative effect.

The vulnerability paradigm puts much emphasis on the active involvement of local residents in the DRR planning and implementation. It points out that in areas where flooding is a regular occurrence, the local people are not just passive bystanders nor completely powerless to prevent or mitigate the floods despite their lack of economic resources (Sakijege et al. 2012). Local residents often possess considerable knowledge concerning how to prevent, cope with, and recover from natural hazards, as the risk of such events is a significant part of their everyday lives. This aspect is even stated in the Sendai framework which recommends the inclusion of people from different social and economic backgrounds in the planning of DRR strategies, different people can then bring their own experiences and unique perspectives to the table (Sendai Framework…2015). The vulnerability paradigm highlights that including the local populations in the planning and design of DRR schemes is extremely important in order to avoid feelings of resentment and abandonment among the local population, who often see the outside experts as intimidating and condescending and lacking the detailed understanding of local contextually grounded problems, concerns, desires, and preferred ways of life (Marchezini 2019). The supporters of the vulnerability paradigm also point out that many of the innovations

designed by Northern experts and high-ranked planners have, in fact, gotten many influences from local traditions. For example, the sponge city concept, promoted by those supporting the hazard paradigm, was, in fact, partly inspired by thousands of years-old water management practices used by Chinese farmers, and observed by the creator of the sponge city concept Yu Kongjian during his childhood (Gies 2018).

Based on these issues, supporters of the vulnerability paradigm emphasize the importance of creating bridges and interconnection between scientific knowledge and local knowledge, in order to learn from one another. Otherwise, important local-based knowledge on disasters, and practices of disaster prevention and disaster recovery, as well as of remediation strategies might be lost. Furthermore, without relevant co-sharing and co-creation of knowledge, it might be difficult to get local residents committed to long-term disaster prevention. It has also proven impossible to simply relocate people living in critical flood risk areas as these people have nowhere else to go and they are prepared to live with the risk of flooding for a chance to survive.

Approaching DRR from the perspective of the hazard paradigm also often includes scenarios, models, and predictions of the changes in physical geography and in the natural environment, in addition to active engagement of local residents and local traditions. Changes in climatic conditions are modeled and assessed for the future in 10-year intervals for example and these changes are then compared with data about socio-economic and demographic conditions in the present (Birkmann 2020). The demographic situation in many parts of the developed world especially is going through a big change in the population pyramid with a rapidly aging population, which means that elderly people will in the near future be a significant portion of the population. An extreme example of this is the looming population crisis in Japan (Sadafumi 2017). Climate change also affects demographics and people with different socio-economic conditions in different ways. An example of this are heat waves, which will become more prevalent in several parts of the world due to the effects of climate change. Older people are much more susceptible to heat stress and people from minorities or with a low socioeconomic status have much less green infrastructure which cools local climates in their neighborhoods and offers shelter from the sun (Norton et al. 2013; Silva et al. 2018; Rowland-Shea et al. 2020; Venter et al. 2020; Schwartz 2021). Predictions and estimates of the change in demographics and socioeconomic development should also be included in the models which aim to predict the future of disaster susceptibility in any given geographic location.

According to the vulnerability paradigm, scientists and other professionals working in the field of DRR should also pay more attention to just disaster governance. Governance means a set of structured arrangements and institutional processes between all of the different actors involved in the decision-making and implementation of strategies planned (Tierney 2012). In disaster governance, this means cooperation between different actors working on disaster risk reduction or disaster response, including local government officials, NGOs, intergovernmental organizations, subcontractors, private companies, and local residents, as a single authority cannot command compliance from different groups of stakeholders

According to the vulnerability approach, it is also important to note that disasters also do not respect administrative borders created by humans and management of DRR often requires cooperation between neighboring countries, international development cooperation agencies, and different kinds of non-governmental sectors. Especially in the Global South, disaster response efforts are often run by NGOs or international development cooperation agencies, and without good governance, a vast amount of the aid money might be directed to bureaucracies or issues that are not the most relevant ones concerning disaster remediation. Thus, there is a need to follow the guidelines of transparency and accountability. The problem with such international aid agencies and NGOs which mobilize only shortly after a disaster is sometimes also that they might lack sufficient in-country experience and they usually do not provide long-term projects of disaster recovery. Thus, much aid money is only going to immediate help, while affected people need to cope with long-term economic and social vulnerabilities on their own, as the first responders and organizations leave to other parts of the world to help in another catastrophe.

The ideas involved in the hazard paradigm, especially those related to recent room for rivers and sponge cities initiatives provide better consideration of ecological aspects of natural hazards, and how to make technologies of risk prevention and mitigation more adaptable to them. The vulnerability paradigm, in turn, offers a more thorough understanding of how natural hazards often affect the groups of the population who are already the most vulnerable and least equipped to manage or mitigate natural hazards. It also provides guidelines to carefully consider local concerns and needs in disaster risk reduction and to seek ways to better involve local residents and local decision-makers in such projects.

2.2   Geospatial data and methods in disaster risk reduction

2.2.1   Geospatial data and GIS in disaster risk reduction

In this chapter, I will explain the geospatial data and the main methods used in most of the disaster risk reduction schemes throughout the world, and with specific relevance especially to the cases of the Global South. Geospatial data and geographic information systems (GIS) play a major role in DRR and are essential to all aspects of disaster management (Goodchild & Glennon 2010; Al-Tahir & Mahabir 2014; Young et al. 2020). GIS data can quickly pinpoint the areas most affected or probable to be affected by different disasters and is essential for navigating the areas and for getting a detailed understanding of different socio-spatial spaces in the affected areas. The intricacies of socio-spatial spaces are formed by the interaction between the different people and societies occupying the urban space and are defined by the social constructs created by the inhabitants living in these spaces (Zieleniec 2018). Satellite images give a birds-eye view of the scope of the disaster and can be processed relatively quickly to provide invaluable information to first responders. Action in the first hours after a disaster caused by a natural hazard is critical for saving lives and first responders need as much relevant information as possible as quickly as possible (Sudden Onsets 2022).

Data used in disaster response efforts must be acquired and/or produced as quickly as possible, while also being as accurate and up-to-date as possible and examples of such services created to cater to this need are the Copernicus Emergency Management Service (CEMS) which is part of the European Union's Earth observation program and the United Nations Institute for Training and Research - Operational Satellite Applications Programme (UNITAR/UNOSAT) Rapid Mapping Service.

CMES provide rapid mapping services free of charge during natural hazard conditions, human-made emergency situations, and humanitarian crises all over the globe by using satellite imagery and other geospatial data (The Emergency…2022). According to their website information: "The EMS process can only be triggered by or through an Authorized User (AU), which include National Focal Points in the EU Member States and countries participating in the Copernicus programme, as well as European Commission services and the European External Action Service (EEAS)." The service they provide depends on the urgency of the emergency response and at the fastest, they promise to produce their First Estimate Fast impact assessment ready-to-print-maps to be ready within three hours and vector data to be ready

within two hours from receiving the first usable images of the natural hazard-affected disaster in question (Rapid Mapping 2022). The First Estimate Products (FEP) are derived from the first suitable post-event images, and they are intended to be used to highlight possibly affected areas with more detailed products that include the extent of the impacted area and the damage grade being promised within 10-12 hours of the request. The Rapid Mapping service is available 24/7 every day of the year and it has been activated 128 times for First Estimate products and 1,538 times for Grading Products between 01.04.2012 and 24.04.2022 (Rapid Mapping 2022).

UNOSAT has been operational since 2003 and it provides satellite image analysis during humanitarian emergencies, including both disasters caused by natural hazards as well as human conflict-situations (UNOSAT Rapid…2022). The Rapid Mapping service is operational 24/7 every day of the year free of charge to UN sister agencies and humanitarian entities operating in line with UN policies. Deliverable products include maps, GIS-ready data, statistics, and reports. Phase 1 emergency response deliverables are promised within 24 to 72 hours of a disaster and include preliminary definition of scope, scale, severity of crisis & rapid assessment design of a disaster (UNOSAT Geospatial Catalogue…2022). This includes mapping the extent of a hazard, affected population estimates, humanitarian access, and raping damage estimation.

A visual aid to the extent of a disaster is a powerful tool of communication and it helps make decisions informed by facts and data (The power of maps 2022). For example, through information on an Excel spreadsheet containing a list of addresses and a description of the damage done it is much more difficult to understand the magnitude of the disaster and its exact damages than through a map of the area that highlights all of the damaged buildings with different colors depending on the scope of damage to each particular building. Even graphs and charts cannot show the scope of a disaster as effectively as a map. It is extremely important to clean up and visualize data into easy-to-understand and informative packages, especially in times of crisis (Meier 2011). Humanitarian aid agencies and emergency management personnel work around the clock during times of crisis to rapidly collect, verify, process, and analyze data in order to generate up-to-date situational awareness information to be used by first responders and long-term humanitarian aid operations (Aman et al. 2014, Schröter et al. 2020). Geospatial data can also be used to analyze past trends for future risk areas and to develop early warning systems for imminent floods or other disasters.

2.2.2 Earth observation and remote sensing

Earth observation (EO) consists of a myriad of methods for gathering various types of information about our planet Earth (Li 2021). The techniques may vary from different remote-sensing methods to in-situ measurements on the ground, but they all share the same common goal of gathering data and mapping the planet's physical, chemical, and biological systems. In-situ observations have been around for centuries as humans have studied our planet and the longest-lasting continuous climate data has been recorded since 1767 and daily since 1815 by the Radcliffe Observatory of Oxford University (Met Office award…2015). In-situ observations are valuable and when spread through a large network of observation posts it is possible to monitor the weather and other phenomena in very large geographical areas.

The research field of EO was really revolutionized by the invention and utilization of airplanes and satellites. Now when people refer to EO they almost exclusively mean remote-sensing methods that include various sensors, like optical and multispectral cameras, synthetic-aperture radar (SAR), or LiDAR attached to satellites, airplanes, or drones for fast and efficient monitoring of the Earth and gathering of data (Tomás & Li 2017). The main reasons for remote-sensing to be a revolutionary advancement for EO are that it enables the monitoring of the planet with higher frequency over long periods of time over large spatial extents, almost real-time collection of data over these large spatial extents, the production of data is much more cost-effective, and the data the sensors provide is objective and independent (Palacios-Lopez et al. 2021). SAR and satellite images can cover hundreds of kilometers in one image and the entire planet can be scanned in a few days. Even areas previously physically inaccessible like the ocean floor or the bottom of lakes can be mapped with sonar systems and mountain regions are easily mapped by satellites. What people choose to do and represent with EO data is subject to different sorts of interpretation and even manipulation, yet the raw data collected by the sensors represents what is happening at the observed location and is not affected by human biases which makes it a relatively powerful source of information. In-situ observations on the ground are still necessary for the validation of remote-sensing data and combining in-situ observations with remote-sensing often provides context and validation of both types of EO data.

The satellite industry was the playing field of large, government-funded projects for a long time, but in the 21st century, private companies have started to take an increasing position in

this field. For example, Maxar technologies produce over 3.8 million square kilometers of high-resolution satellite imagery every day, and the Finnish microsatellite company ICEYE is capable of monitoring any place on Earth in any weather night and day with its microsatellite X-band synthetic aperture radar sensor constellation (Maxar is a leading…2022, Persistent monitoring…2022). ICEYE has recently grown its satellite constellation large enough to enable persistent monitoring of natural hazards worldwide 24/7 (Persistent monitoring by ICEYE…2021). There are still limitations on capacity and downlinking, so monitoring a place's exact time and location is not always possible, however, ICEYE's core focus is flood capture and they have optimized their flood monitoring systems to capture flooding events and have meteorologists monitoring and predicting future flooding events across the globe (Kitajima 2021; Rosen 2021).

Synthetic aperture radar (SAR) data differs from optical images as it is an active data collection method (Moreira et al. 2013; What is Synthetic…2020). SAR technology is based on remotely mapping the reflectivity of objects and environments with high spatial accuracy by sending electromagnetic waves toward the Earth and measuring the backscatter with their sensors (Ferro-Famil & Pottier 2016). This can be compared to bats using sound waves to navigate in dark caves. Different objects and surfaces reflect electromagnetic waves in different ways, and this makes it possible to differentiate objects and surfaces from the SAR data. Water and paved surfaces, such as roads or airports, produce a specular reflection making it appear very dark in raw SAR images, urban structures produce double-bounce scattering and appear bright white, while the majority of it consists of diffuse scattering which is produced by rough surfaces (Figure 2, Learn Synthetic…2021). Figure 2 shows the SAR imagery of the city center of Bangkok, Thailand and the region around Tula de Allende, Mexico derived from the ICEYE archives. The imagery looks very different from traditional optical satellite imagery; in these SAR images it is relatively easy to distinguish water (black color), paved roads, buildings, agricultural fields, and mountains.

Figure 2. SAR imagery of the city center of Bangkok, Thailand (left) and the region around Tula de Allende, Mexico (right) from the ICEYE archives (Source: ICEYE Imagery Archive 2022).

The ICEYE SAR satellites operate in three different imaging modes, Strip Mode, Spot Mode, and Scan Mode (Completing the picture…2021). The resolution and area covered differ considerably, depending on the imaging mode, because the radar beam can be steered, and the illumination time can be altered. The longer the beams are fixed on one position, the higher the image resolution. Standard Strip mode images have a ground resolution of three meters and cover an area of 30km by 50km, standard Spot mode has a resolution of one meter but only covers an area of 5km by 5km, and Scan mode can cover an area of 100km by 100km, but with a reduced resolution (that is still at least better than 15 meters). SAR has many capabilities that are not possible with optical imagery, for example, clouds, smog, volcanic ash, or sandstorms do not affect X-band radars, and this enables SAR data to be collected regardless of the weather conditions on Earth, which is a huge advantage (Completing the picture…2021). Humanitarian aid organizations can follow the effects of natural hazards like hurricanes, forest fires, or erupting volcanoes without sunlight and without having to wait for the clouds and smoke to disappear which can take a relatively long time. By using SAR satellites, a flood caused by a hurricane can be immediately mapped and lava flows from volcanoes can be monitored before the volcanic ash dissipates. The accuracy of SAR sensors is also not affected by the distance to the ground, which makes it possible to get very high-resolution data even from space. This characteristic, paired with Coherent Change Detection (CCD) makes it possible to track changes on the ground in extremely high detail, even higher than a person on the ground would be able to detect. With ICEYE's SAR satellites it is possible to see if a vehicle has been driven over a field if products have been moved around in a shipyard, or if a forest has been cut down.

CCD means taking high-resolution satellite images from precisely the same geographical area over time to detect the change (What is Coherent…2022). Another advantage of SAR imagery is the possibility to control the resolution and coverage of the imagery because the SAR illumination is controlled electronically. This also makes it possible to update the algorithms as they develop. SAR measurements are also extremely precise and when properly calibrated the geolocation is precise to the scale of single pixels.

2.2.3   Global geospatial data sets

The availability of geospatial data, like the one produced by Google Maps, ICEYE, and Maxar, or used by emergency professionals is often taken for granted in the Global North. However, this kind of geospatial data is often lacking in the Global South or at the very least lacking in detail and accuracy. Controlling information is and has always been a very significant source of authority and power during the human history and many political regimes in the world cling to every possible source of data, reluctant to make their data open-source or deliberately manipulating data given to the public (Fuentes 2019; Carlitz & McLellan 2020; Evans 2020). Many kinds of data are also managed and controlled by international organizations and private companies, such as Google, ICEYE, and Maxar, with legal copyrights to restrict the use of data without permission or payment. In addition, Google is highly protective of its data as it makes significant investments yearly to upkeep, gather, and systemize this data.

For these reasons, open data is essential, especially for many parts of the Global South, where there is a considerable lack of GIS expertise, and a lack of highly sophisticated and expensive equipment for the collection of geospatial data, as well as a lack of systematically collected and analyzed data. An independent evaluation of the UNOSAT Raping Mapping service included an online survey which revealed that 78 percent of responders who rely on the service would use another open-source service if the UNOSAT Rapid Mapping service did not exist (Independent Evaluation…2018). Therefore, creating, maintaining, updating, and improving open-source data is essential for the efficiency and equality of global humanitarian aid, the development of DRR schemes, and improvement of geospatial preparedness (Hernaiz 2019). Many academics, international organizations, and NGOs have started to turn to volunteered geographic information (VGI) to help countries in the Global South to solve this problem. This especially because VGI data is free, large amounts are created continuously and it is often the only GIS data available in the Global South (Goodchild & Li 2012; Yilma 2019).

OpenStreetMap (OSM) is the most famous example of VGI data. OSM is an interactive web map and at the same time it is the most extensive open-source geospatial database in the world, thus, it is considered a valuable source of information that is used by thousands of websites, mobile apps, and hardware devices, as well as by NGOs and the United Nations in humanitarian aid and disaster response efforts (Anderson et al. 2019; Hernaiz 2019; OpenStreetMap: About 2021). OSM basemaps are used by many big companies and raw OSM data is utilized for analysis and programming by companies like Apple, Lyft, and TeleNav and by Pokémon Go, a highly popular mobile game based on geospatial technology (Major OpenStreetMap Consumers 2022). OpenStreetMap can be described as the Wikipedia of maps as anyone can create an account and edit the map to contribute GIS data about roads, buildings, services, public transportation routes, waterways, natural parks, addresses, and many more aspects in any place in the world (About OpenStreetMap 2021). The map and its data are created and managed by a community of mappers who are all volunteers. By being based entirely on volunteer work OSM forms a certain kind of global community mapping effort that is based on co-creating and co-sharing of knowledge. OSM was created to offer an alternative and more importantly a free source of geographical data (FAQ 2021). OSM data can be freely downloaded, copied, distributed, transmitted, and built upon by anyone under the Open Data Commons Open Database License (ODbL) as long as OpenStreetMap and its contributors are credited. The documentation is licensed under the Creative Commons Attribution-ShareAlike 2.0 license (CC BY-SA 2.0).

Recent years have seen major growth in the utilization of professional, paid editors of OSM by large corporations (Anderson et al. 2019). Companies like Apple, Microsoft, Amazon, Uber, TeleNav, and Mapbox have mobilized to map areas that are of special interest to their companies' strategies (Figure 3). Apple and Mapbox are active all around the globe as they are focused on updating maps for their projects, while companies like TeleNav and Amazon focus solely on areas most important to their company's success. Microsoft has invested significantly in deep learning, computer vision, and artificial intelligence to generate high-quality building footprints with the goal of increasing global building footprint coverage and it has released hundreds of millions of these building footprints as open data to be used by OpenStreetMap and humanitarian efforts (Building Footprints 2022). Amazon.com, Inc. employs over 400 paid editors, who edit road names, turn restrictions, directionality, and road connectivity to OSM in

several countries to improve the maps for their operations and for everyone else using OSM
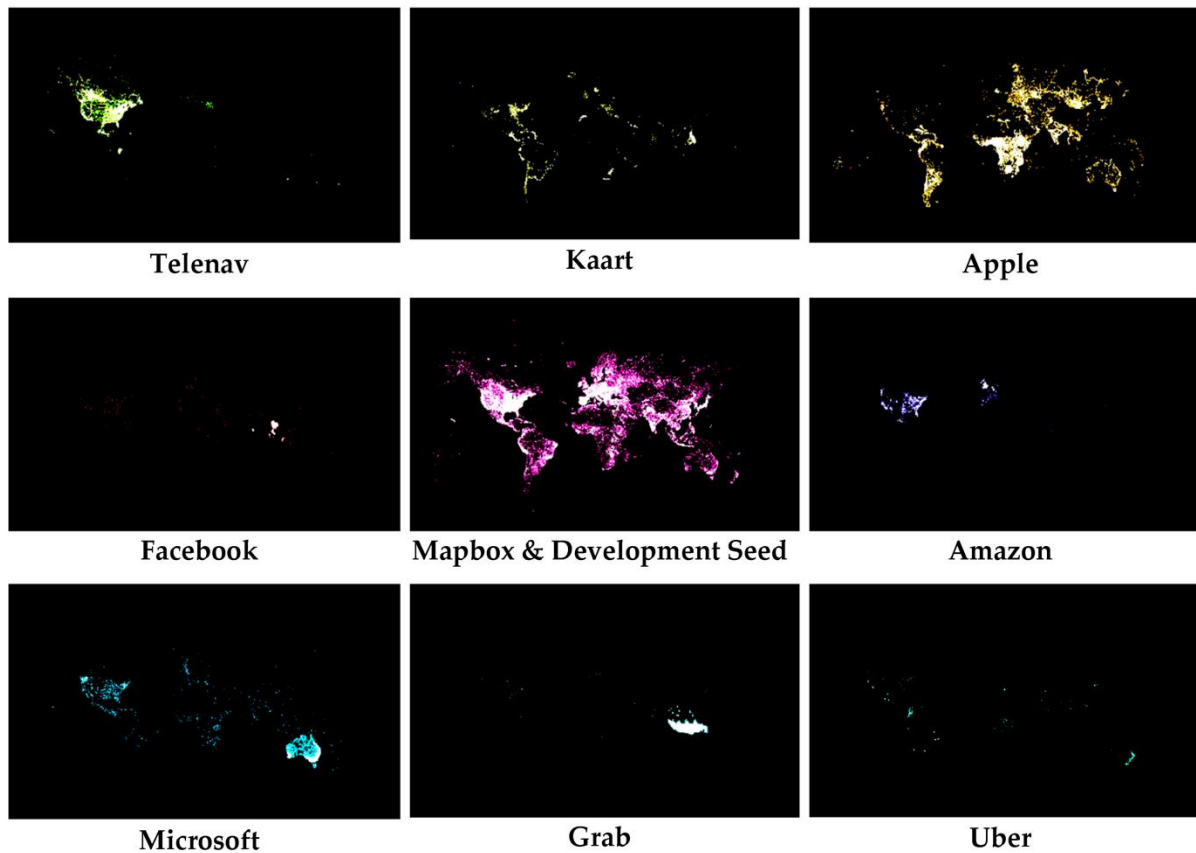(Navigating OpenStreetMap…2020).

Figure 3. World maps show where different companies are editing OSM (Source: Anderson et al. 2019).

Another solution to the lack of geospatial data is organizing Mapathon events. Mapathons are
coordinated mapping events by projects like Missing Maps, or the Humanitarian
OpenStreetMap Team where particular groups of people comes together to map previously
unmapped areas around the world using satellite imagery. This data is then used with disaster
risk assessment and to update OpenStreetMap. These events can also be organized by local high
schools, as has been done in Turku, Finland for example (Mapathon 2022). In this way, the high
school students get experience in GIS, and they actually help map unmapped areas, often
located in the Global South. The Resilience Academy has carried out similar work in Tanzania
with university students. The Resilience Academy is a partnership between several Tanzanian
universities, the University of Turku, the Government of Tanzania, the World Bank, and the
Foreign, Commonwealth and Development Office (FCDO, UK) (About Resilience Academy
2022). It aims to "equip young people with the tools, knowledge, and skills to address the

world's most pressing urban challenges and ensure resilient urban development." (Resilience Academy 2022).

Despite all the positive sides of a crowdsourcing projects like OSM, they are not free of criticism either. Although OSM is an initiative to democratize data, it also contains certain risks of being used as a form or in service of data colonialism (Young et al. 2021). Data colonialism is defined as a process in which data that is produced by users and citizens is later appropriated by private corporations, governments, and/or NGOs for their own benefit and for profit-making, without the user necessarily even realizing how the profit is made and how thoroughly their lives are being surveilled (Thatcher et al. 2016). OSM was created in London, England, to map the world by using crowdsourcing and its success has led to the normalization of what a crowdsourcing project in the Global South should look like, based mostly on the views of what the developers in the Global North consider to be important (Young et al. 2021). Thus, there is a risk that without good collaboration with local people in the Global South as is strongly advocated in the organized editing guidelines, OSM mapping projects might focus solely on the topics that are of high interest to the Global North developers, corporations, and users (Anderson et al. 2019). Besides, most of the OSM mapping by users is carried out using satellite imagery produced by governments of the developed world or by high-tech companies based in the Global North or in the countries of rapidly emerging economies. Using satellites as an all-seeing eye in the sky might also end up as being a certain form of a top-down view of the world, and thus a certain form of data colonialism. Data produced by these satellites are not owned by developing countries and the local populations have little possibility to decide what is being mapped and how it is mapped. In addition, if the mapping is done just by using satellite imagery, without any detailed understanding of the context, or situated knowledge of the complexities of local relations and practices, it can lead to severe consequences. There was such kind of conflict in 2010 called the "Google Maps war" between Nicaragua and Costa Rica when Google Maps published a borderline drawn in the wrong place by a US government official and Nicaraguan authorities used this data as a validation to occupy an island belonging to Costa Rica (Wilkinson 2021). A similar interpretation of the disputed border between Pakistan and India by an official working in the Office of the Geographer in the United States in the late 1960s has led to decades of fighting that continues even today high in the region of the Himalayas between Pakistan and India.

There is also a very clear difference in the accuracy and amount of data available in OSM between Global North countries and those of the Global South. This is mostly because people with the technological means, know-how, expertise, and time to edit OSM live mostly in developed countries, thus logically they will focus on mapping their own neighborhoods first. Some scholars have also argued that the Global North governments and companies help digitize the Global South only to exploit the data for their own needs and to create new customers for their economic benefit (Young 2019).

Well-founded criticism and skepticism can serve as a driving force for change and improvement as they can reveal vulnerabilities and weaknesses in proposed or existing solutions. In any case, despite the criticue, at its core, OSM is fundamentally an open-for-all approach to the lack of geospatial data, and it is in a constant search for solutions to develop better and more transparent forms of geospatial data. Anyone can map anything, and the representatives of the Global South are strongly encouraged to map everything they consider crucial and necessary. The "all-seeing eye" of satellites also sees everything in the Global North and excluding locations that are considered secret because of national security issues of a particular country, everything is mapped, often with even higher detail than in the Global South. With SAR technology even clouds are unable to provide cover from satellites. Thus, currently it is impossible to escape this exposure as this "Pandora's box" has already been opened, and what is crucial now is to try to make the OSM project as transparent, accountable, and fair as possible.

There are also challenges with data interoperability in DRR when there are many completely separate operators all working with their own separate data (Migliorini et al. 2019; Murnane et al. 2019). Data standards can vary immensely and data from separate sources often need to be integrated to work well together. Large-scale disaster response operations include actors from emergency services, politicians, healthcare services, NGOs, and different levels of government. These actors might use different languages, software, protocols, regulations, and decision-making chains. There are often significant differences in the quality of metadata, documentation, and quality of the data. Every organization has its own standards, templates, and acronyms to use when producing geospatial data and without extensive metadata and documentation, these data sets can be problematic to use by others, as utilizing them without fully understanding them can easily lead to misinterpretation of the data. These problems stem from a lack of data interoperability and clear guidelines between all of the numerous actors involved in the DRR GIS scene. In fact, these are problems that will probably persist in

international DRR projects and operations also in the near future, for the reason that there are always numerous stakeholders, and they keep changing depending on the location of the disasters in case. OSM also sometimes suffers from the sometimes-extreme disparity in their data even in neighboring houses. As OMS is based on volunteer work, it is also important to recognize that different individuals editing the map have different editing habits and the attention to detail varies from user to user, despite some moderation and existing mapping guidelines documented by the OSM Foundation.

The World Settlement Footprint 2019 Population dataset (WSF2019-Pop) has been created by the German Aerospace Agency (DLR) to combat the lack of accurate population census data all around the globe, especially in the Global South. The data is already being used by the World Bank and the Global Facility for Disaster Reduction and Recovery (GFDRR) to map the population at risk to urban hazards (Palacios-Lopez et al. 2021). Recent research has also shown that large geospatial data sets such as the WSF2019-Pop dataset can help support, implement, and monitor at least 11 out of 17 Sustainable development goals (SDGs) and play a major role in disaster risk reduction (Qui et al. 2019; Kuffer et al. 2020; Kavvada et al. 2020; Palacios-Lopez et al. 2021). A global dataset with a sufficiently accurate spatial resolution like the WSF2019-Pop data (~10m spatial resolution at the Equator) can be used as a single proxy for population modeling making disaster mapping less time-consuming which is critical in the first hours after the disaster when time is of the essence. Other global population data sets, such as WorldPop and GHS-Pop, have been found to be too coarse for accurate population estimates at local scales and have performed poorly in several studies (Smith et al. 2019).

Testing the WSF2019-Pop creation process for the entirety of Africa, Palacios-Lopez et al. achieved accuracy assessments of "reliable" for 25 to 36 countries, medium reliability for 15 countries, and poorly reliable for only two countries (Palacios-Lopez et al. 2021). Even though this method is not without errors and cannot be seen as 100 % accurate in any scenario, the WSF2019-Population data has multiple advantages over other global population products since it provides a weighting framework that is calculated independently of other geospatial layers. In addition, it has a much higher spatial resolution for more accurate local analysis, and it smooths the changes between administrative units, which eliminates visibly abrupt changes in the data. Furthermore, it is easier to update and replicate around the globe, as the model does not require the acquiring of multiple geospatial layers of equal quality, extent, spatial resolution,

and spatiotemporal coverage. As there are no other geospatial data sets involved in the production, there is no need to worry about endogeneity issues either.

EO and geospatial data can be used to monitor phenomena at varying scales from very local locations such as construction sites or mining operations to data sets covering the entire planet (E.g., Google Earth, human population extents, OSM). Combining precise local data sets with global data sets allows for the monitoring of geohazards faster and more accurately. Accurate and precise local data sets can be used to verify and correct global data sets and global data sets can be used to quickly create very accurate estimations of geographical phenomena in rapidly changing landscapes (e.g., urbanization, deforestation, ice loss). Local data sets are often more precise, but they require much more work by skilled GIS professionals to be created and many countries of the Global South do not have the necessary funding or know-how to create and update GIS databases in their countries or cities or to build and launch their own satellite constellations (Sala & Dendena 2015).

3   Study sites

Two Global South cities, Bangkok in Thailand, and Tula de Allende in Mexico were selected for this study. The criteria for selection were that both cities are located in the Global South and have suffered from severe pluvial and fluvial flooding combined with certain mistakes made in flood management and preparedness (Marks 2015; Raziel 2021; Guillén 2021). Furthermore, both cities have a river flowing through the city center and flooding in both places usually affects the poorest part of the population the most. Importantly, these sites are also very different in size, location, and infrastructure, so it puts the automatization process to the test and shows the difference in OSM data in these two areas. Comparing data from in some certain respects different, but in other aspects similar areas it is possible to see how effectively the automated process works with significantly different amounts of OpenStreetMap data available and with different sized flood extents.

ICEYE's SAR satellites were able to capture flooding in Bangkok, Thailand and Tula de Allende, Mexico in August of 2021 and ICEYE has provided flood data from these two flooding events to be used in this study from their extensive library of data on captured flooding events. The German Aerospace Agency (DLR) created completely new data from the two study sites upon request. These two study sites were selected for this thesis in cooperation with ICEYE, the World Bank, and the Resilience Academy.
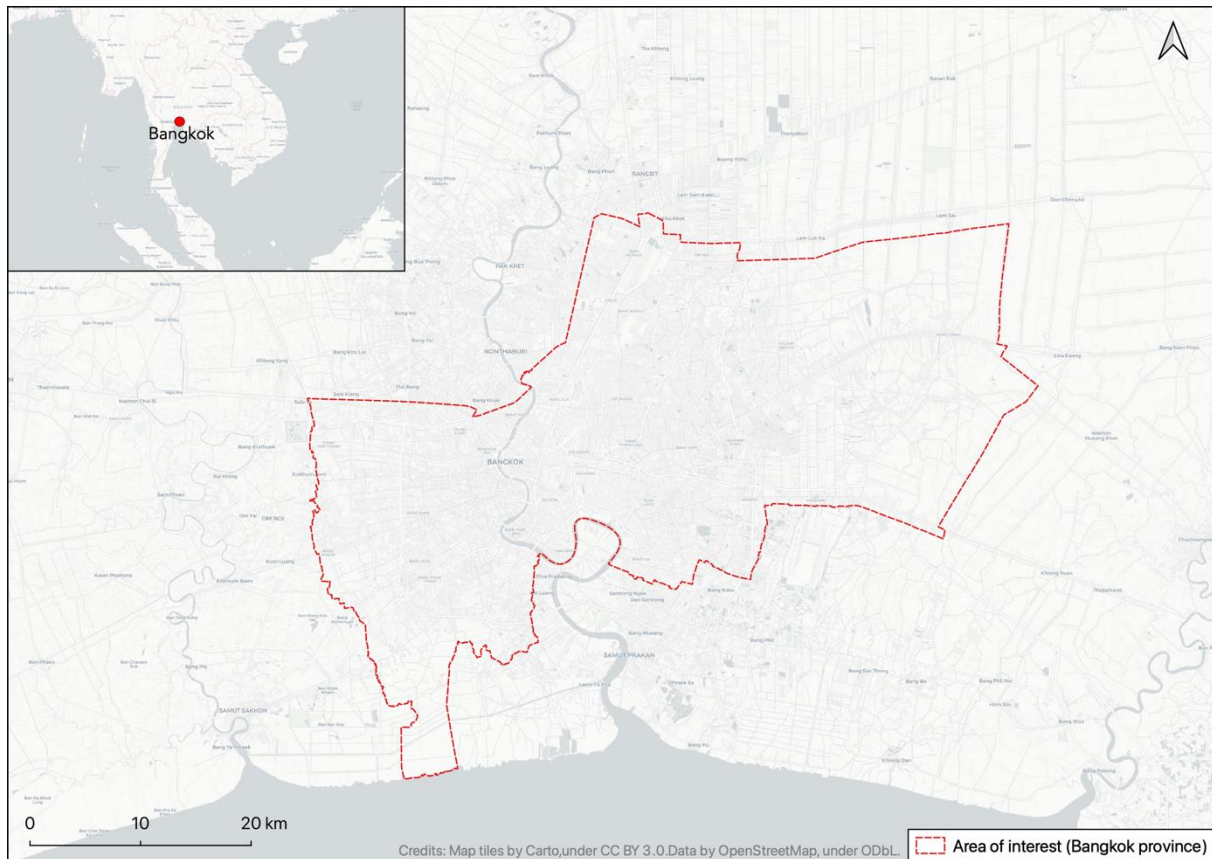
Figure 4. Map of Bangkok, Bangkok showing the extent of the study area and the extent of the data used.

Bangkok (กรุงเทพมหานคร in Thai) is the capital of Thailand and has a population of 10.9 million inhabitants which is ~15.7 percent of the entire population of Thailand (Figure 4, Thailand 2022). Bangkok is one of the most popular tourist destinations in the world year after year and it is financially, culturally, and politically the most influential and important city in Thailand and as such Bangkok can be classified as a clear primate city. It suffers from similar problems as many rapidly growing megacities in the Global South, such as Dar es Salaam in Tanzania; Jakarta in Indonesia; or Mumbai in India, that it has a considerable number of informal settlements growing at an uncontrolled rate. All of these cities also suffer from frequent severe flooding which affects the informal settlements most severely. The Bangkok metropolitan area had an estimated informal settlement population of 1.1 million inhabitants in 2010, and this number has surely only grown since then (Yap & De Wandeler 2010). People living in informal settlements are very poor and they do not have the economic means to relocate to safety from flooded areas and in the case of heavy flooding, they can only wait for the waters to recede while living in wet and unsanitary homes severely affected by floodwaters (Berquist et al. 2014). The. 2021 flood in Bangkok covered in total an area of 46 746 km^2 of land, but in order to be able to test the automatization process in an extremely large and populous city, this data has been clipped to the extent of the city limits of Bangkok (Figure 4).

Tula de Allende is a town in the State of Hidalgo in Mexico. It had a population of ~115 000 inhabitants in 2020 (Tula de Allende 2020). Tula de Allende is internationally most famous for its large archeological site, containing important artifacts and information about the Mesoamerican pre-colonial society and culture. Figure 5 presents the area of interest around Tula de Allende relevant for this study.
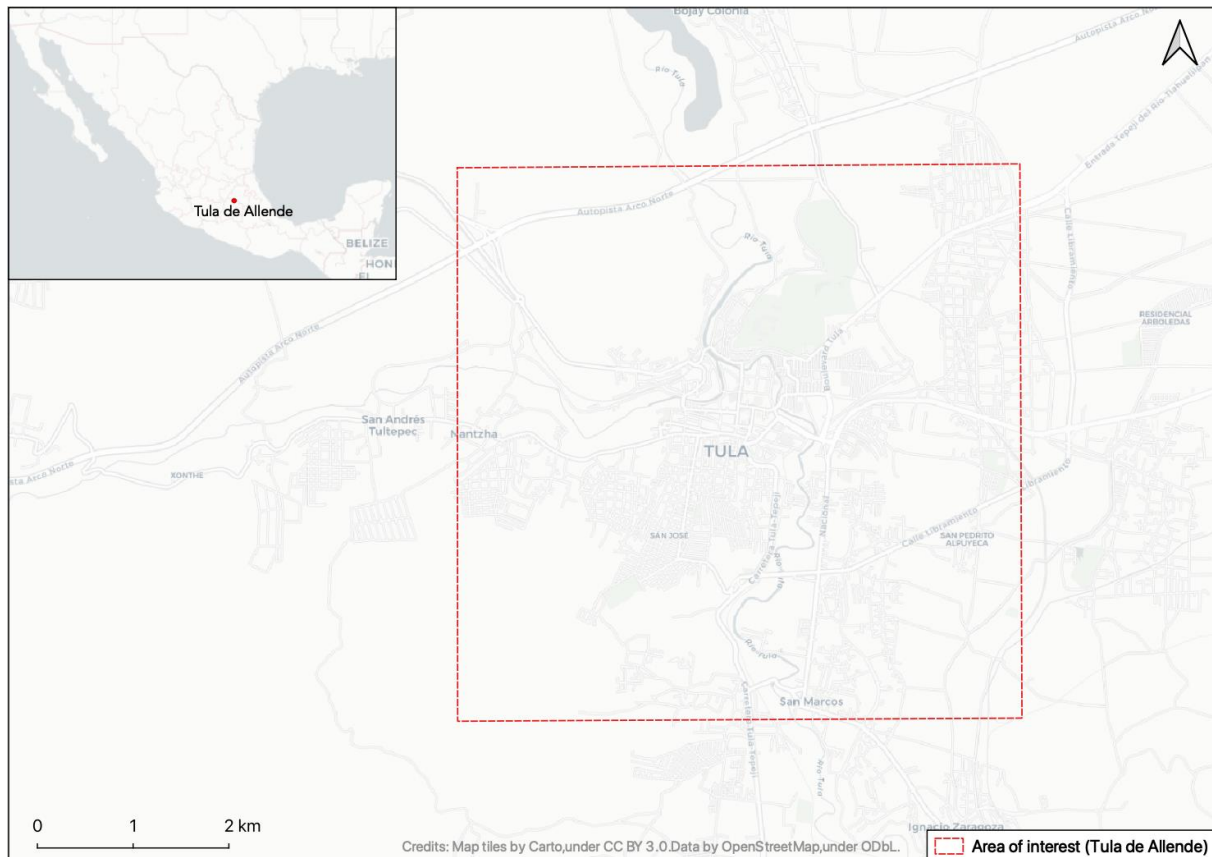
Figure 5. Map of Tula de Allende, Mexico showing the extent of the study area and the extent of the data used.

The main reason for Tula de Allende was selected as a part of this study is that it is very different in size compared to Bangkok, Thailand, while simultaneously sharing important similarities with Bangkok, as it suffered from a devastating flood in 2021. In addition, it is a city in the Global South, with very little geospatial open data available. The flooding event in Tula de Allende covered in total a territory of 30 km^2, focused especially on the surroundings of the river that runs through the city. A flood in September of 2021 led to the death of 17 people and damaged 2 000 houses as the riverbank burst and flooded the city with water (Seventeen people…2021).

4   Data and methods

4.1   Research approach

An automated flood analysis and visualization model was developed in this study using Python and Kepler.gl. The aim of the model is to automatically access, edit, analyze, and visualize flood disaster data as a solution to get faster first assessments of local flooding events using global data sets. This solution is developed particularly for cities in the Global South, for situations, where sophisticated early warning systems are missing and geospatial data about natural hazards and their effects on the local infrastructure and population are scarce.

Global geospatial data sets were used as data to make the automatization model global and repeatable anywhere in the world and to save time. The developed model calculates flood depth classes and the number of people affected by different flood depths in the area of interest. Building footprints are extracted from OpenStreetMap to get an estimation of the number of buildings and essential services (hospitals and pharmacies in this case) that have been inundated by different floodwater depths. The model was thoroughly tested and repeated to get average processing times for different segments of the model and to examine its reliability and the consistency of results. The planned use cases for this model are to get first assessments of the extent and effects of natural hazards such as floods, especially in the Global South where geospatial data and expertise is lacking. The Python coding language was chosen to perform all of the analysis and mapping as it is an open-source method, it enables the automatization of the process, and the produced code is easy to share openly for anyone to use.

4.2   Geospatial data sets

The following data was used to develop the automated model: the World Population Footprint 2019 data produced by the German Aerospace Agency (DLR), flood depth and extent data produced by ICEYE, and open-source OpenStreetMap data (Table 1).

The flood extent and depth data set used in this thesis were provided by ICEYE, a New Space company that manufactures and operates the world's largest SAR satellite constellation. ICEYE monitors global weather 24/7 to predict floods and provides clients access to near real-time SAR satellite data to respond to natural hazards more accurately and quickly and to improve decision-making (Flood Monitoring 2022). While ICEYE's data is often produced by

monitoring very local-level phenomena, its satellite constellation has a global reach, making it a producer of global geospatial data. The population data for this thesis was produced and provided by the DLR. It has been created to offer an improved global population dataset compared to existing ones. It is aimed to be used for population estimations and to monitor urbanization along with population growth globally. Finally, global open-source data from OpenStreetMap was used to get building footprints and the locations of hospitals and pharmacies.

Both the flood and population data were delivered as raster files in a GeoTIFF format. TIFF is an abbreviation from "Tagged Image File Format", which is a high-quality image format that does not lose any image quality (TIFF files. 2022). A GeoTIFF also contains geographic data embedded as tags within the file, which allows it to be plotted accurately with Python or GIS software into a world map accurately. The building footprints were extracted from OpenStreetMap using the OSMnx Python module. The data is extracted into JupyterLab as a GeoPandas GeoDataFrame which can be plotted on a map based on the geometry column which stores the coordinates of the subject. The existence of the geometry column is what differentiates GeoDataFrames from Pandas DataFrames, a much better-known Python format. A GeoDataFrame resembles an Excel spreadsheet or an Esri shapefile attribute table. GeoDataFrames can be created, accessed, edited, and manipulated with Python to find correlations, gain insights, analyze patterns, perform geospatial analysis, and highlight important aspects of data. Table 1 briefly explains the characteristics of the data used.

Table 1. Geospatial data sets used in the development of the automated flood map model and geovisualization solution.

Name	Provider	Format	Extent	Precision	CRS	Attributes	No-Data (raster) / Geometry (vector)
ICEYE flood depth and extent (Bangkok, Thailand)	ICEYE	GeoTIFF	Local	2.81m x 2.81m resolution	EPSG:32647	Band 1	3.40282e+38
ICEYE flood depth and extent (Tula de Allende, Mexico)	ICEYE	GeoTIFF	Local	2.81m x 2.81m resolution	EPSG:6369	Band 1	3.40282e+38
WSF2019-Pop (Bangkok, Thailand)	DLR	GeoTIFF	Global	8.98m x 8.98m resolution	EPSG:4326	Band 1: red_nir_max	n/a
WSF2019-Pop (Tula de Allende, Mexico)	DLR	GeoTIFF	Global	8.98m x 8.98m resolution	EPSG:4326	Band 1: red_nir_max	n/a
Building footprints	OSM	GeoData Frame	Global	Precise based on visual inspection, but incomplete	EPSG:4326	Geometry, building type, name, OSMid	MultiPolygon, Point
Essential services (Hospitals & pharmacies)	OSM	GeoData Frame	Global	Precise based on visual inspection, but incomplete	EPSG:4326	54 columns	Multipolygo, Point

4.2.1 Flood extent and depth data

Synthetic-aperture radar (SAR) satellite images were provided by the Finnish company ICEYE to be used in this thesis. The company provided flood extent data captured by their SAR satellite during 2021 floods in Bangkok, Thailand and Tula de Allende, Mexico in 2021 (Figure 6). The data has been preprocessed by ICEYE so that flood depth values have been extracted from raw SAR data and converted into a GeoTIFF (Table 1). Pixel values in the delivered GeoTIFF equal flood depth in meters. No data values are -3.40282e+38 and they have been filtered out in the Python script.
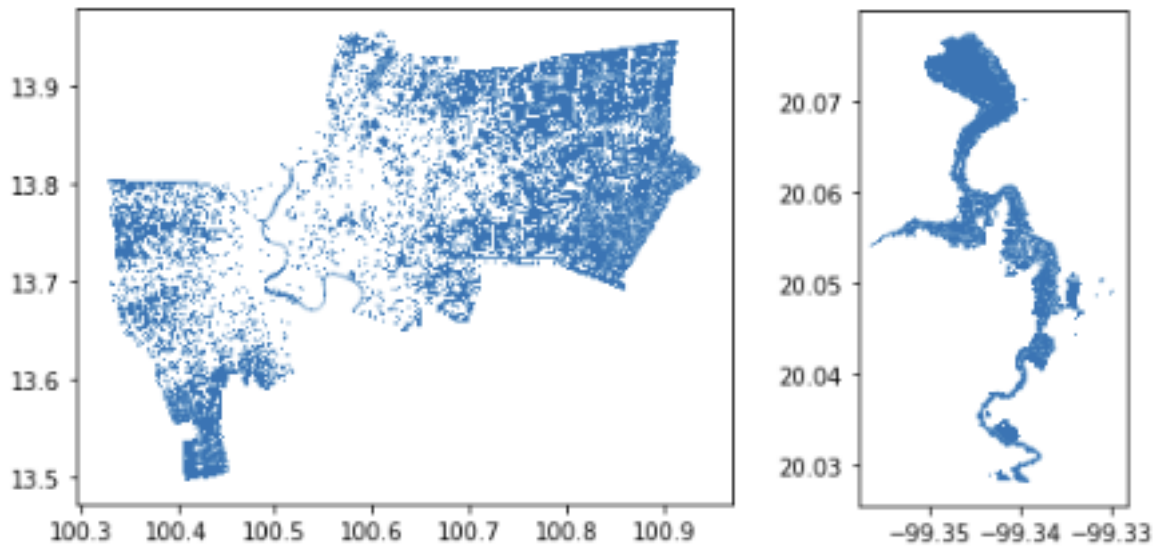
Figure 6. Flood depth data captured by ICEYE related to Bangkok, Thailand (left) and Tula de Allende, Mexico (right). The numbers around the images represent coordinates in WGS84 (EPSG:4326).

4.2.2   Global population data

The population estimation dataset for this thesis was provided by the DLR. The specific data product used in this thesis is called the World Settlement Footprint 2019 Population dataset (WSF2019-Pop). It is a dataset of population estimates and it is a preliminary internal version that is being used at the DLR and that is planned to be released as a global dataset.
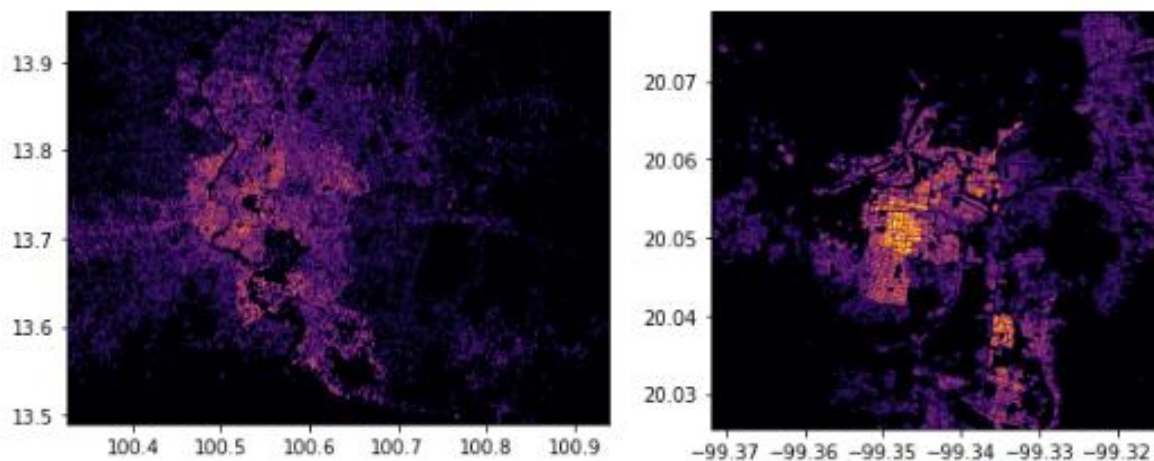

Figure 7. WSF2019-Pop data from Bangkok, Thailand (left) and Tula de Allende, Mexico (right). The lighter areas represent a higher concentration of the human population in the area. Note that the data has been mapped using a sequential colormap which means that Bangkok has significantly more population than Tula de Allende even though Tula de Allende has much brighter spots in the example images. The numbers around the images are coordinates in WGS84 (EPSG:4326).

By request, they have created two rasters that cover the study areas of this thesis's: Bangkok, in Thailand and Tula de Allende, in Mexico (Table 1, Figure 7). The main advantage of using a global data product is that it makes the mapping process considerable faster. Because there is no need to search for population data, this data can act as a single proxy for population modeling.

The WSF2019-Pop data has been created by first analyzing the percentage of the impervious surface or density in a 10m pixel (Palacios-Lopez et al. 2021). This is done by first computing the maximum temporal NDVI from Sentinel 2 optical imagery, which is an effective approximation of the presence of vegetation on the ground. The resulting maxNDVI is then compared to available imperviousness surfaces from OpenStreetMap by extracting data from OSM and rasterizing it to 10m spatial resolution and passed through support vector regression modules to produce a percentage imperviousness surface (PIS) of the pixels marked as human settlements globally at a 10m spatial resolution. The pixel values in the imperviousness layer are from 0 to 100 and represent the percentage of impervious surface, or density. The population estimates are then calculated from the open archive of the WorldPop Global Project and data provided by Center for International Earth Science Information Network (CIESIN). Population data has been standardized and categorized into administrative units (e.g., blocks, counties, or municipalities) and an estimated population count for every pixel within the administrative units is defined using the following formula:

$$Pop_{(p \in IU)} = Pop_{IU} \frac{PIS_p}{\sum_{(p \in IU)}(PIS_p)}$$

The values are disaggregated in proportion to the 0-100 values and the end result is a population dataset that depicts residential populations adjusted to the UN national total estimates as pixel values in a raster file. The spatial resolution is ~10m at the Equator and each pixel value represents the number of people per pixel. In the data that was delivered the pixel count was multiplied by 1,000 to save the file as an integer so the values have been divided by 1,000 in the Python script. Non-settlements have a value of -1 and pixels where no resident population is estimated have been set to a value of 0.

4.2.3 OpenStreetMap

In this study, OSM data has been gathered only from the study area using area the of interest polygons calculated in the script and the OSMnx Python package created by Geoff Boeing
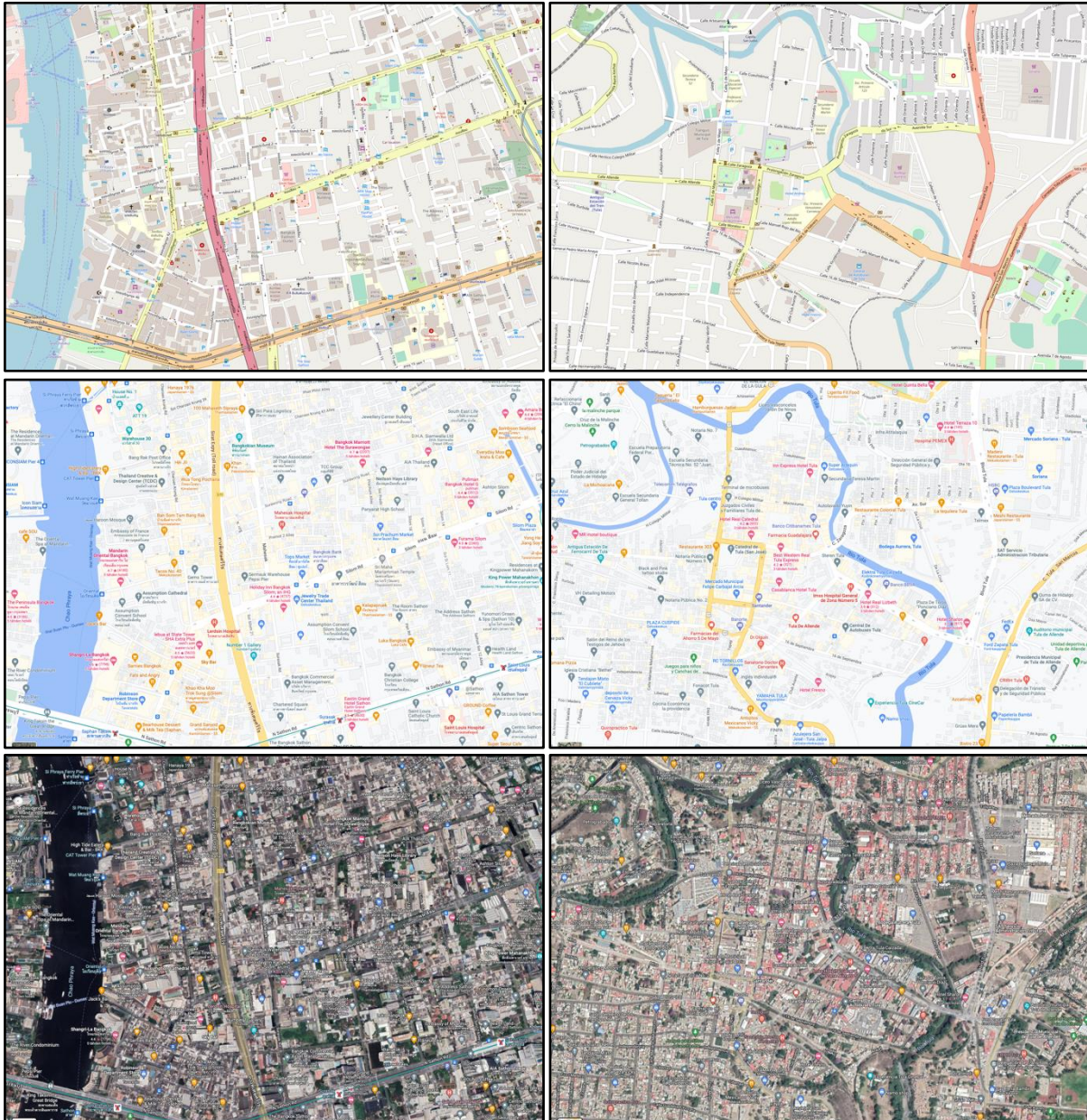
Figure 8. Visual comparison of the amount of mapped OSM data (first row), Google Maps data (second row), and Google satellite imagery (third row) between Bangkok, Thailand (left), and Tula de Allende, Mexico (right).

(Boeing 2017). The visual comparison of different data sets and the reality show that OSM is lacking many building footprints in each study site (Figure 8). The snapshots are from the city centers, which are almost always the most extensively mapped due to the high concentration of economically, strategically, and socially important locations, thus data lacking from these areas suggests that most probably the situation is even worse in more peripheral rural regions.

Overlaying the extracted OSM building footprints over satellite imagery from Bangkok, Thailand reveals that a very small percentage of buildings have been mapped, which is a re-occurring problem in the Global South (Figure 9).

Figure 9. ICEYE flood data and OSM inundated buildings overlaid on top of satellite imagery highlight the lack of mapped buildings in OSM in Bangkok, Thailand.

4.3   Automating flood analysis and visualization with Python

Python is free for anyone to use and one of the most popular and most used programming languages by data analysts and researchers in the world. It is a very powerful programming language, relatively easy to learn, and available for macOS, Windows, and Linux (General Python FAQ 2022). Python comes with a very large standard library and there are many very powerful and useful third-party Python libraries, packages, and modules that have been developed by people or groups for anyone to use for free use, which enables any almost any type of data to be created, manipulated, analyzed, and visualized using Python.

All of the analysis in this thesis was achieved using the Python coding language using, version 3.9.0 of Python (Figure 13). The Python script developed in this thesis used several open-source Python libraries, packages, and modules that require installation (Table 2).

Table 2. List of Python libraries, packages, and modules that were used in this thesis.

Name	Version	Information
Pandas	1.4.1	https://pandas.pydata.org
Geopandas	0.10.2	https://geopandas.org/en/stable/
Matplotlib	3.4.2	https://matplotlib.org
Kepler.gl	0.3.2	https://docs.kepler.gl/docs/user-guides
Shapely	1.8.1.post1	https://shapely.readthedocs.io/en/stable/
OSMnx	1.1.2	https://osmnx.readthedocs.io/en/stable/
Rasterio	1.2.10	https://rasterio.readthedocs.io/en/latest/
PyCRS	1.0.2	https://github.com/karimbahgat/PyCRS
Os	-	https://docs.python.org/3/library/os.html
Mapclassify	2.4.3	https://pypi.org/project/mapclassify/
Seaborn	0.11.2	https://seaborn.pydata.org
Pyproj	3.3.0	https://pyproj4.github.io/pyproj/stable/
Utm	0.7.0	https://pypi.org/project/utm/

The Python script was written and run using JupyterLab, a web-based interactive development environment for Jupyter notebook documents, code, and data (JupyterLab 2022). Jupyter is a completely free coding environment that has been developed by Project Jupyter, a non-profit, open-source project. JupyterLab supports over 40 programming languages and focuses on making coding, data science, scientific computing, and machine learning fast, efficient, and user-friendly.

The raster data was first read into JupyterLab and reprojected to the World Geodetic System (WGS84, EPSG:4326) coordinate reference system (CRS). The raster data of the flood captured by ICEYE's SAR satellites was polygonized (the process of creating vector polygons from the pixels of a raster layer) so that flood depth classes could be created by classifying the flood depth polygons into seven user-defined flood classes (Table 4). The floodwater depth classes were chosen based on the effects of water depths on humans and buildings according to studies (Table 3. R&D Outputs…2006; Pistrika et al. 2014; Huizinga et al. 2017).

A Python function ("*findtheutm*," see Appendix 1, page 5) was used to get accurate area calculations of flood extents. This was achieved with a function that takes latitude and longitude points from the centroid of the chosen data and then calculates in which UTM zone (Figure 10) the latitude point is in and chooses the correct Coordinate Reference System (CRS) for the area calculation. The result is then converted to km^2.
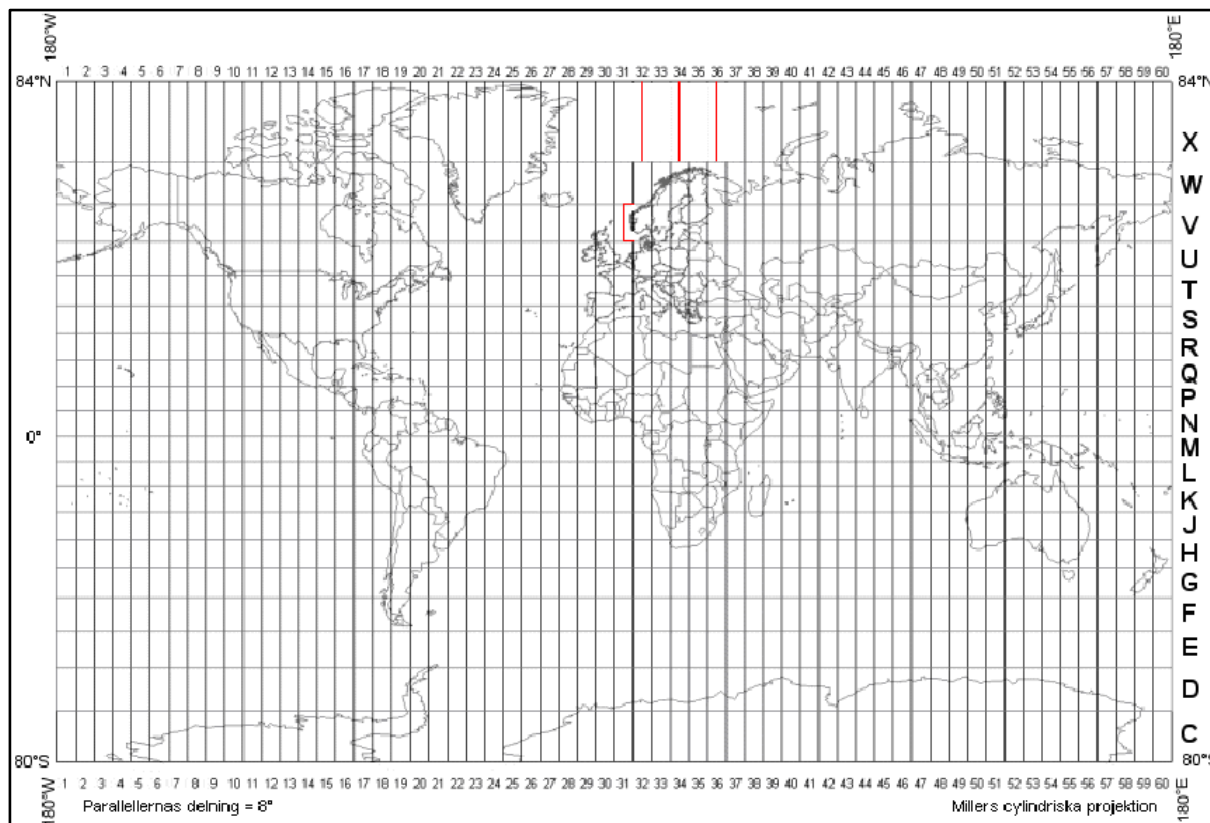
Figure 10. Map that shows the UTM zones and is the basis of how the correct CRS is chosen in the 'findtheutm' Python function. Läntmäteriet (2022)
<https://www.lantmateriet.se/contentassets/379fe00e09d74fa68550f4154350b047/utm-zoner.gif>

Using this function, the script can automatize an area calculation anywhere on Earth with minimal distortion. Using a layer projected to WGS84 (EPSG:4326) the area calculation would be extremely distorted the further away an area of interest is from the equator, and the area calculations would be in degrees instead of meters making it extremely difficult to interpret the results.

4.4 Preprocessing geospatial data sets

The OSMnx Python package makes it easy to download boundaries, building footprints, data points, and street networks from OpenStreetMap (Boeing 2017). This eliminates the time-

consuming process of finding freely available data from the flooded areas. In this thesis, using the OSMnx "ox.geometries.geometries_from_polygon()" feature, OSM data was downloaded using polygon masks created from the polygonized flood data in order to get data only from within the boundaries of the flooded area (see Appendix 1, page 4). This method was chosen to enable the script to work automatically anywhere in the world as the flood class MultiPolygons contain the geolocation information and the OSM data is always downloaded from the correct geographical area without user intervention required. The Python script gets the geographical location information and the correct coordinate system from the reprojected version of the original raster file provided by ICEYE (Figure 11).
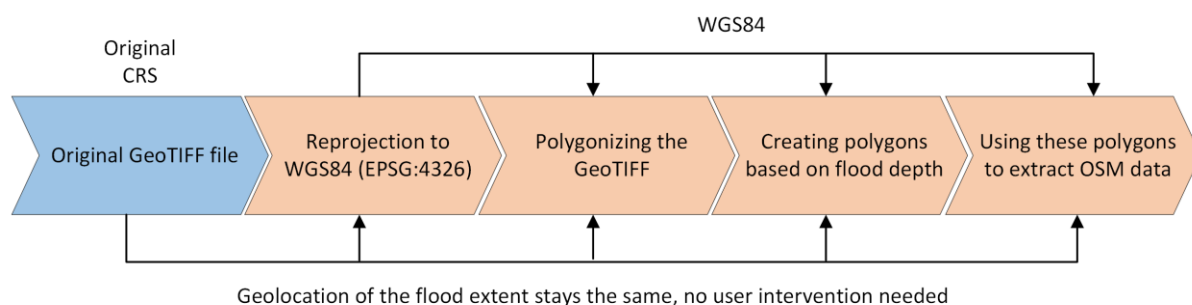

Figure 11. Flowchart visualizes the automized OSM data extraction without the need for user intervention to determine where the data needs to be downloaded from. This process can be used with data from anywhere in the world and the OSM data will always be extracted from the correct location.

This method was also used to make the processing manageable on normal laptops as the city limits of Bangkok contained over 100 000 building polygons at the time of writing this thesis and this number will hopefully rise as OSM is edited and updated in the years to come. The data was trimmed to make it lighter and easier to process by removing unnecessary columns. The raw data had over 200 columns with varying amounts of information and from this, only the "geometry", "building", and "name" columns were kept. Some of the buildings were also extracted as point and linestring features by OSMnx and these were removed to keep only the polygon footprints of buildings. The hospital and pharmacy data were also extracted using the flood polygon masks but the polygons were converted into point features used to highlight their location instead of polygons since all the building footprints had already been extracted. Icons were used to visualize the points to make them stand out more on the map and to provide such visual cues that most people are familiar with. Points are also much easier and faster to process than polygons. The hospital and pharmacy data had to be searched for using two different OSM tags since OSM data can be mapped in different ways depending on which tags the editing user decides to use. Each OSM element has a tag that consists of a key and a value. The tags describe

OSM elements; the key (the column name) describes the topic, category, or type of feature and the value provides detail for the key-specified feature. For example, the hospitals were searched for using the tags "building=hospital" and "amenity=hospital" and to find all the inundated buildings the tag "building=True" was used. In this thesis OSM data for hospitals was extracted using the following line of code:

```
hospital = ox.geometries.geometries_from_polygon(osm_aoi,
    tags={'building':'hospital', 'amenity':'hospital'})
```

Inundated buildings were extracted by first dividing the flood depth into two classes: under 0.5 meters of floodwater and over 0.5 meters of floodwater. The building footprints were then extracted from within these polygons into two variables: slightly inundated buildings and severely inundated buildings. This method was chosen based on global depth-damage curves which point to the fact that at around ≥ 0.5 meters of floodwater buildings start to suffer critical structural and interior damage (R&D Outputs…2006; Pistrika et al. 2014; Huizinga et al. 2017). The amount of damage is highly dependent on what materials are used in the construction of the building, how well the buildings are built, how fast the water is flowing, and many, many other factors, but crude estimates can be made and were considered in the analysis. However, for an unknown reason, this method did not extract all of the inundated building footprints and a third class had to be created that includes all of those building footprints that fell in-between the two classifications. This group of buildings was named "inundated buildings" and was found by extracting all of the buildings from the entire flood extent area and then searching for duplicates between all the buildings, the slightly inundated buildings, and the severely inundated buildings. It is not a case of invalid geometry in the flood classes as the validity of the MultiPolygon geometries was checked using the "geopandas.GeoSeries.is_valid" function.

4.5   Overlay analysis

Overlay analysis of the vector and raster data was used to achieve results on the impact of the floods. Overlay analysis is the process of superimposing different data sets to identify relationships between them and thus producing new information. Overlay analysis is a part of almost all spatial analysis processes and is the core of GIS projects as the end product is almost always a map that shows different information laid on top of each other in separate layers (Figure 12). In this thesis, overlay analysis was used to overlay flood depth data on top of the WSF2019-Pop and OSM building footprint data to find out how many people lived in areas affected by different water depths and how many buildings had been inundated by the floodwaters.

The affected population estimates were achieved by using overlay analysis of the flood data and the population estimation dataset. The population count per flood class was calculated with Rasterio using each
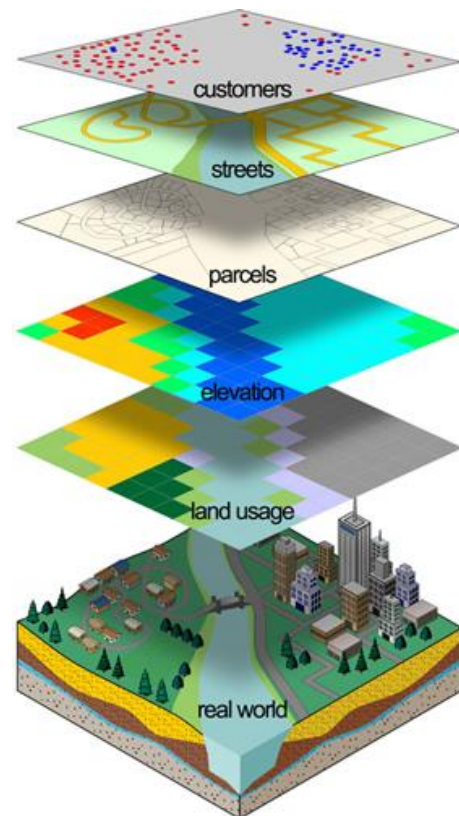
Figure 12. Overlay analysis visualized (Westfield State University, 2013, <http://wiki.gis.com/wiki/index.php/File: Gis_layers.png>)

flood class MultiPolygon as masks to clip the WSF2019-Pop raster data. This method not only returns results of the total population affected by a flooding event but also divides it by flood classes to reveal the number of people affected by extreme water depths and thus in dire need of immediate assistance. All the pixel values inside the individual masks were then summed up by flood class and inserted into a new column called "population" using the following code:

```
# Modified from: Lau & Um (2021) Using GeoSpatial Data Analytics: A Friendly
Guide to Folium and Rasterio

# URL: https://omdena.com/blog/geospatial-data-analytics/
#
# ------------------

# Creating an empty list for storing values
results = []

# Looping through flood classes in dataframe and using them as masks for
extracting values
for i in iceye_breaks['flood_class']:

    roi = iceye_breaks[iceye_breaks.flood_class == i]

    # Using the mask.mask Rasterio module for specifying the ROI
    gtraster, bound = rio.mask.mask(wsf_raster, roi['geometry'], crop=True)

    # Using values greater than 0 to get population data from the WSF2019-
Imp raster pixels
    results.append(gtraster[0][gtraster[0]>0].sum())

# Saving results in new column
iceye_breaks['population'] = results

# Dividing by 1000 because the WSF2019-Pop data has been multiplied by 1000
to save the file as integer
iceye_breaks['population'] = iceye_breaks['population'].div(1000)
```

4.6   Plotting data with Kepler Gl

Several different geospatial data visualization libraries were tested aside from Kepler.gl
including Folium, Leafmap, Plotly, IpyLeaflet, and Geemap, but ultimately Kepler.gl was
chosen as the most suitable for this thesis. It offers very easy visualization tools, processes data
very efficiently, and performs very well.

Kepler.gl is a high-performance open-source web-based application for visualizing large
geospatial data sets (Kepler.gl 2022). It is built on top of Mapbox GL and deck.gl which are
services that provide custom online maps and complex visualizations possibilities. Kepler.gl is
able to render millions of datapoints on the fly and performs very well with large data sets.

Creating a map in JupyterLab by using the Kepler.gl library allows the user to add data with a few lines of code and opens an interactive view of Kepler.gl where it is possible to edit the visualization by using only a computer mouse. Icon names for hospitals and pharmacies had to be inserted into new columns in the corresponding GeoDataFrames ('plus-alt' for hospitals and 'heart' for pharmacies) so that these string values could then be recognized by Kepler.gl to import them from the Kepler.gl icon database. Once satisfied with the visualization the configurations can be saved as a variable (variable named "config" in this script) and used with other data that is named exactly the same and is in the same format. The Kepler.gl map visualizations can also be changed while viewing the maps, but these changes are not saved. To save edited visualizations for later use, the user has to edit the configurations, call the configurations with the "map_1.config" function, and save the configurations as a variable in the script (see Appendix 1, page 7).

In this script, ICEYE flood data and OSM data were added to the map and then visualized using the Kepler.gl configurations. Latitude and longitude values extracted from the flood extent data were used to zoom the map to the flooded area, so users do not have to search for the flood on a world map. The saved visualization configurations were added to the data and the map. The maps from Bangkok and Tula de Allende were then saved as local HTML files which can be opened with a web browser if the file is downloaded to the computer in use. This was done because Kepler.gl is a client-side only application, which means that it does not send or store any data to any backends, limiting the saving and sharing of maps. The advantage of saving the map as an HTML file is that once downloaded it can be viewed and used normally by opening the file with a web browser. The map can also be shared by sending just the HTML file to another user.

5   Results

5.1   Automated flood analysis and visualization model

The automated flood analysis and visualization model was created by progressing through the analysis and visualization phases step by step and tying it all together into a model that works without user intervention by using previous outputs as inputs in the next process (Figure 13).
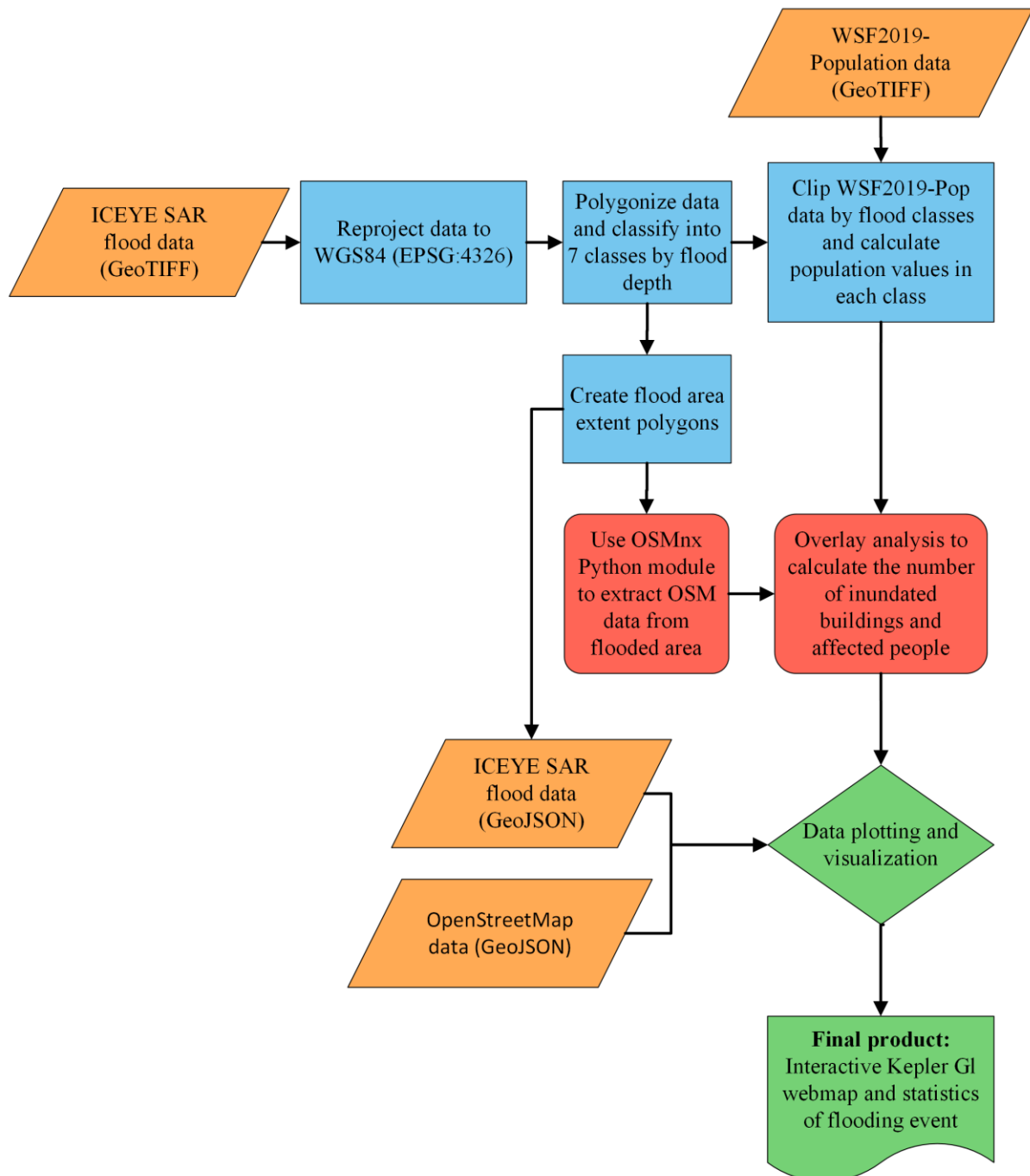

Figure 13. Workflow map visualizing the Python code process. The flow chart contains the GIS data used, the most important steps, and the end result of the Python script.

The maps from Bangkok and Tula de Allende produced by this script (see Appendix 1) can be viewed by going to: *https://seafile.utu.fi/d/ea2fcc03a4d147c58da0/* and downloading the two .html files and then opening them in a browser. The files should open in the default browser where the user can then interact with the maps.

The code is presented in its entirety in Appendix 1, and it can also be found in the following GitHub repository: *https://github.com/resilienceacademy/-flood-automapping*. Running the script on the computer in use will produce an interactive map as a .html file, a bar chart of the affected population by flood classes, and print information related to the number of buildings inundated and of essential services (hospitals and pharmacies) inundated. Running the script requires access to either the two raster data sets used in this thesis (ICEYE flood depth and WSF2019-Population data) or raster data that is similar.

The script was created and tested using a 2017 13-inch MacBook Pro that has a 2,3 GHz Dual-Core Intel Core i5 processor, 8GB of memory, and an Intel Iris Plus Graphics 640 1536 MB graphics card and is using macOS Monterey Version 12.3.1. In 2022 this is not a particularly powerful computer, but it performed fairly well in processing the script and by far the longest time processing data in Bangkok was to extract the OSM data which is more dependent more on internet speeds and OSM server requests limitations (Table 3).

Table 3. Processing times and averages for both study areas. The script performed fairly consistently in the 5 test runs per study area that were used to gather this information.

	Time in seconds (Bangkok, Thailand)					
Process	1st test	2nd test	3rd test	4th test	5th test	Average
Importing modules	8.21	7.92	7.3	6.03	4.39	6.77
Raster transformations	0.846	0.776	0.571	0.62	0.272	0.617
Vector transformations	108	95	90	91	89	94.6
Raster clipping	27	25.7	27.3	26.3	25.7	26.4
Getting OSM data	492	460	453	456	441	460.4
Results	26.2	25.3	26.3	26.6	24.7	25.82
Kepler.gl	32.6	14.073	15.193	15.176	16.065	18.621
Total time	694.856	628.769	619.66	621.726	601.127	633.228
	Time in seconds (Tula de Allende, Mexico)					
Process	1st test	2nd test	3rd test	4th test	5th test	Average
Importing modules	6.95	6.53	3.04	3.94	4.53	4.998
Raster transformations	0.272	0.55	0.314	0.442	0.31	0.378
Vector transformations	18.7	18.6	18.1	19.4	19.9	18.94
Raster clipping	1.24	1.15	1.35	1.14	1.64	1.304
Getting OSM data	21.8	10.2	10.2	11.3	10.8	12.86
Results	1.71	2.00	1.98	2.42	2.55	2.132
Kepler.gl	1.525	1.338	0.962	0.882	0.93	1.127
Total time	52.197	40.368	35.946	39.524	40.660	41.739

The data from the two study sites were consistently mapped to the correct location despite the different coordinate reference systems in the source data. This means that the reprojection process in the code works automatically without the user having to worry about projecting the data correctly, a somewhat regular headache in GIS work. The area calculations are also accurate, despite the geographic location of the flood changing as the "*findtheutm*" function (see Appendix 1, page 5) finds the correct CRS to use for the area in question. The final map also automatically zooms to the correct area of interest without user interaction. These functions were tested by running the script with data from the two study sites, located on the opposite sides of the Earth.

As expected, the longest processing time by the model in the study site of Bangkok was to extract the over 12,000 building footprints of inundated buildings. In Tula de Allende there were only 36 OSM building footprints in the flooded area, and it took only 12.86 seconds on average to extract the data compared to 7 minutes and 40 seconds on average in Bangkok. In Tula de Allende the longest average processing time was taken by the vector transformations which include polygonizing the raster and breaking it down by flood depths. There is a

significant difference in the total average processing time between the two study areas with 10 minutes and 33 seconds for Bangkok and only 41.739 seconds for Tula de Allende. This is logically explained easily by the massive difference in the flood extent, 718.76 km^2 in Bangkok and only 2.2 km^2 in Tula de Allende, and the massive difference in the amount of OSM data from these two areas. The visualization configurations transfer over without issue and as such the visualization of the flood and different buildings are identical in both study sites. There are also no issues with producing the bar chart or the prints of the number of inundated buildings and essential services (hospitals and pharmacies) (see Appendix 1, page 6).

Compared to the delivery times promised by the CEMS Raping Mapping program (ready in 2 to 3 hours after the first imagery) even the 10 minutes and 33 seconds of processing time in Bangkok by this script is significantly faster. In smaller areas like Tula de Allende, the map is practically instantly ready and depends only on the time it takes ICEYE to process raw SAR imagery into a product that can be used by this model. The data used in this model does not have to be flood data produced by ICEYE, the only requirements are that it has been geolocated and it is a raster with values that correspond to flood depths. The same applies to the population estimates, any population data that is in a raster format that has been geolocated and has pixel values corresponding to population estimates can be used to get results with this model.

5.2 Model accuracy and reliability in the cases of Bangkok and Tula de Allende

5.2.1 Bangkok, Thailand

The total flooded area in Bangkok was 718.76 km^2 covering an area where over 1.4 million people live and over 13 000 buildings were inundated with floodwater (Figure 14, Table 4). The overlay analysis of the flood depths and WSF2019-Pop pixels showed that there were approximately 1,440,977 people were affected by the flooding (Table 5).
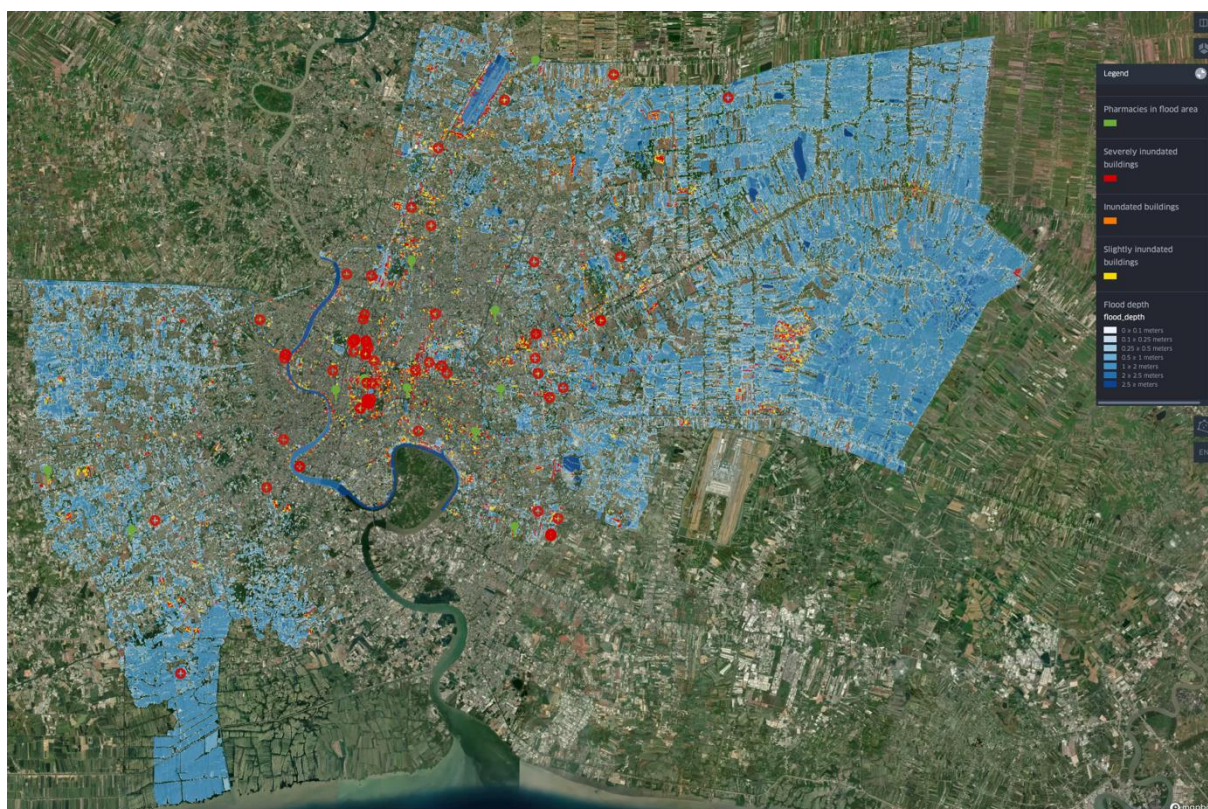
Figure 14. The flood map of Bangkok, Thailand created using the Python script. Creating this map does not need any user input apart from changing the file path to the raster of a flooding event and the WSF2019-Pop data.

Table 4. Flood classes defined in this thesis, their corresponding floodwater depths, and how many square kilometers the floodwaters covered in Bangkok, Thailand.

Flood class	Flood depth	Flooded area (718.76 km^2 in total)	Population affected
0	$0 \geq 0.1$m	40.62 km^2	175,192
1	$0.1 \geq 0.25$m	66.04 km^2	251,923
2	$0.25 \geq 0.5$m	104.54 km^2	328,273
3	$0.5 \geq 1$m	208.25 km^2	397,039
4	$1 \geq 2$m	261.46 km^2	248,873
5	$2 \geq 2.5$m	23.00 km^2	22,462
6	2.5m $\geq$	14.92 km^2	18,477

Of these people, 685,989 live in areas that were inundated by over 0.5 meters of water according to the flood zones calculated from the ICEYE SAR data, which indicates that their buildings were severely inundated. The homes of people living in at least a second floor of a high-rise building would not be inundated, but it could be extremely dangerous for these people to exit their building while the floodwaters are high. According to several studies, standing still in waters of 1.5m or more is impossible for adults (depending on the height of the person) and if the water is flowing with high velocity, as it usually does during floods, it could be impossible to stand in water depths of just 0.3m if the velocity of water flow is high (R&D Outputs…2006). The flooding in August of 2021 in Bangkok, Thailand was so extensive that 1,013,859 people were living in areas that had flood depths of over 0.25 meters during the flooding event.

Table 5. Bar chart from the number of people affected by the flood in Bangkok, Thailand. This bar chart is produced automatically by the Python script.
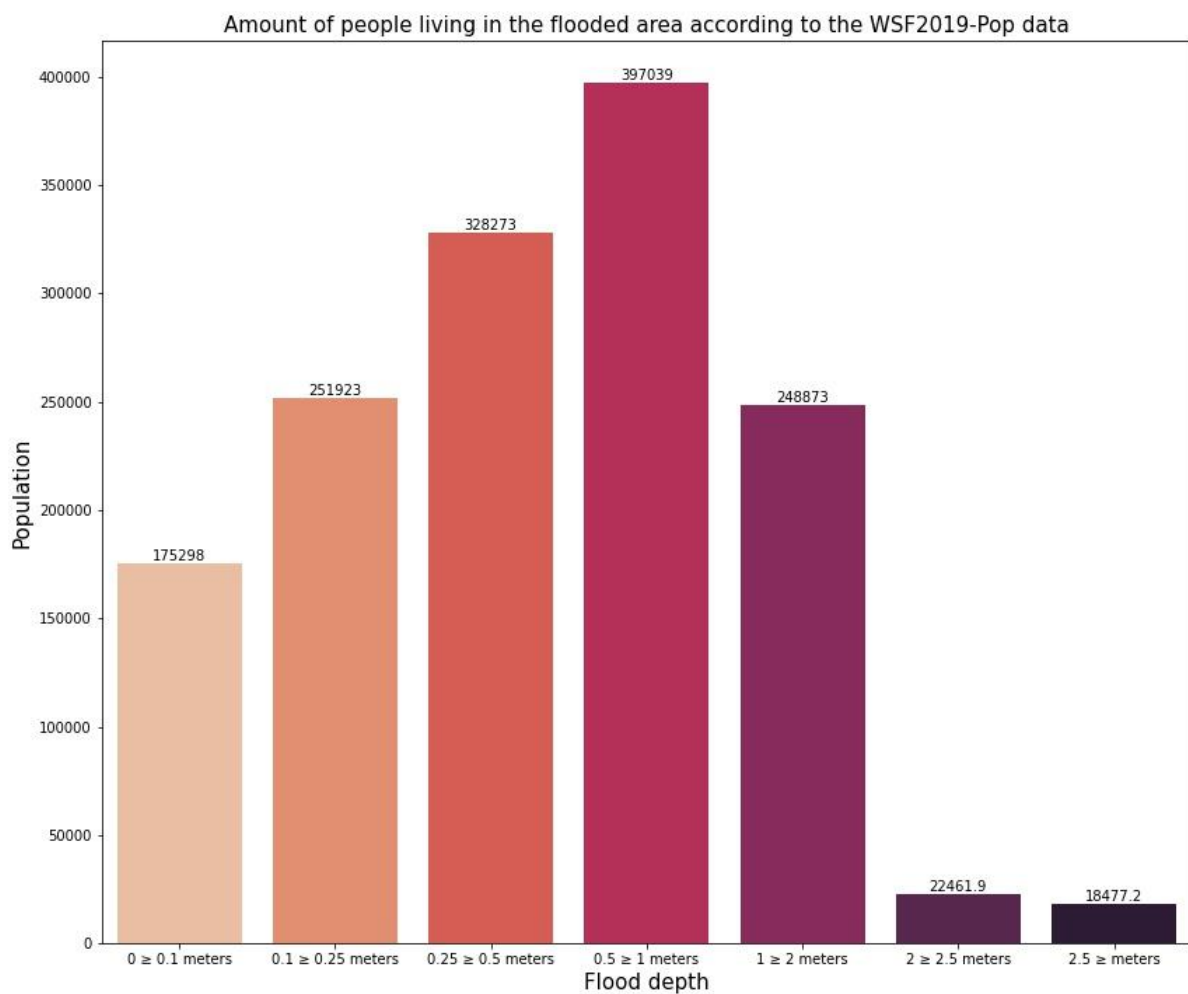
The people severely affected most likely had little or no chance to go outside during this flooding event and the water levels that were captured by ICEYE would have caused first responders significant trouble to reach these people in urgent need of assistance.

According to the overlay analysis of the flooded area, there were 5,112 buildings inundated by over 0.5 meters of water, 6260 buildings inundated by a maximum of 0.5 meters of water, and 1,786 buildings that fell in-between the two categories, bringing the total number of inundated buildings to 12,843 based on the OSM data during the writing of this thesis. Of these buildings, 74 were hospital buildings and 11 were pharmacies. This does not mean that 74 separate hospitals in Bangkok were inundated as hospitals usually consist of multiple separate buildings. Most of the mapped buildings in OSM in the area of Bangkok are strategically, societally, or economically important big buildings. The OSM data in the area of Bangkok lacks many buildings, which makes the situation much direr in reality, especially in the poor neighborhoods that are often overlooked in favor of mapping more affluent and economically important areas. These results might change considerably even in a short span of time as OSM is constantly changing and updating by nature.

5.2.2   Tula de Allende, Mexico

The total flooded area in Tula de Allende was ~2.20 km^2, covering an area where 4,908 people live and 36 buildings were inundated with floodwater (Figure 15, Table 6). The overlay analysis of the flood depths and WSF2019-Pop pixels showed that there were approximately 4,908 people who were affected by the flooding (Table 7). Of these people, 3,956 lived in areas that were inundated by over 0.5 meters of water according to the flood classes calculated from the ICEYE SAR data.

Figure 15. The flood map of Tula de Allende, Mexico created using Python. Creating this map does not need any user input apart from changing the file path to the raster of a flooding event and the WSF2019-Pop data.

Table 6. Flood classes defined in this thesis, their corresponding floodwater depths, and how many square kilometers the floodwaters covered in Tula de Allende, Mexico.

Flood class	Flood depth	Flooded area (2.20 km^2 in total)	Population affected
0	$0 \geq 0.1$m	0.05 km^2	151
1	$0.1 \geq 0.25$m	0.08 km^2	284
2	$0.25 \geq 0.5$m	0.13 km^2	517
3	$0.5 \geq 1$m	0.26 km^2	1,195
4	$1 \geq 2$m	0.42 km^2	1,664
5	$2 \geq 2.5$m	0.14 km^2	368
6	2.5m $\geq$	1.11 km^2	729

Of the 36 inundated buildings in Tula de Allende, two were hospital buildings. There were no inundated pharmacy buildings, which also proves that the script does not cause an error if there is no data from the area of interest. The OSM data in the area of Tula de Allende, Mexico lacks a majority of buildings, like in the case of Bangkok, which makes the number of inundated buildings much higher in reality. The results show a severe lack of data in these Global South cities and reveals the need to increase the pace of mapping OSM data in these areas. The results are from the time of writing this thesis as OSM data might and hopefully will change over time when more data is added.

Table 7. Bar chart from the number of people affected by the flood in Tula de Allende, Mexico. This bar chart is produced automatically by the Python script.
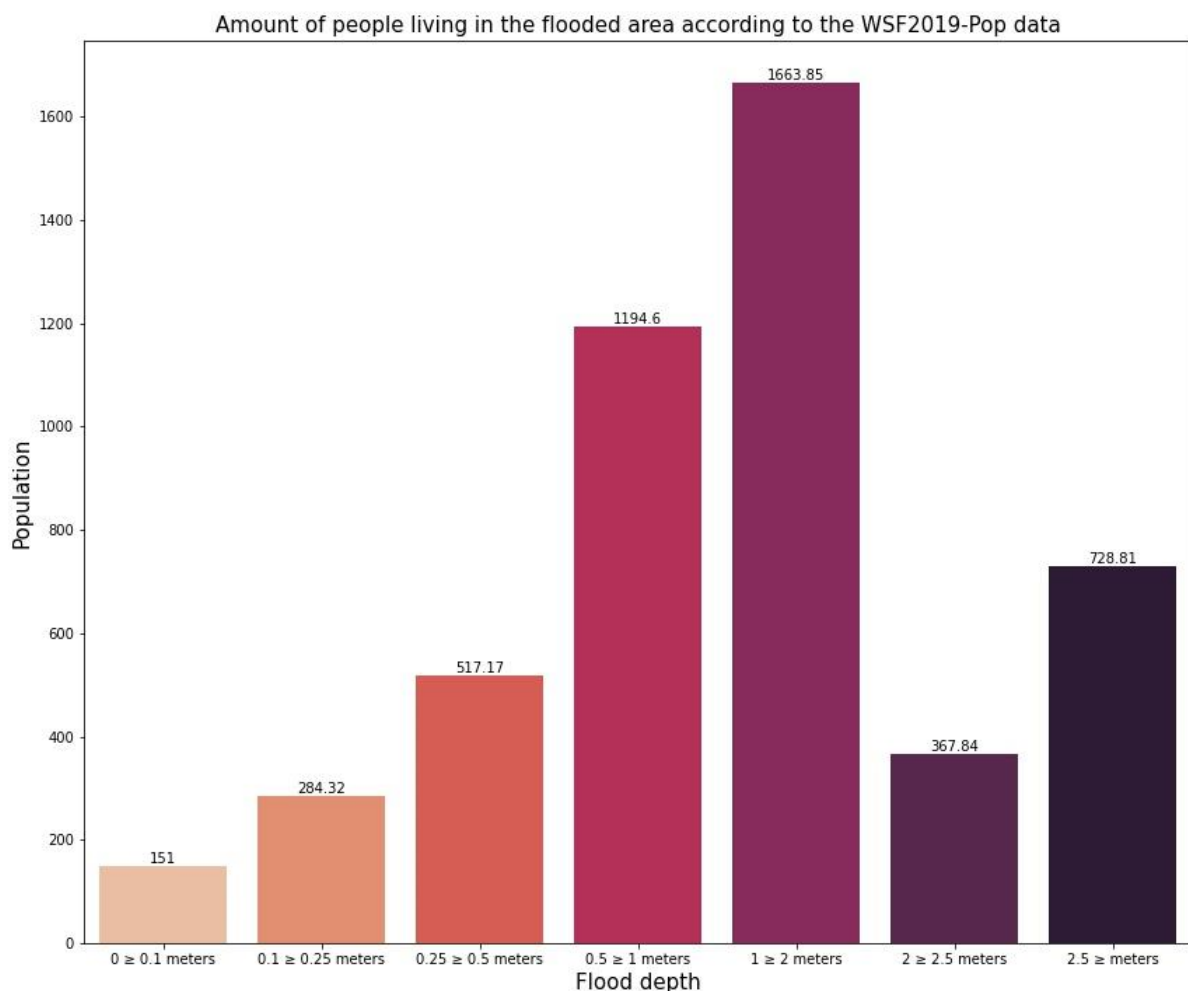
The results of produced by the model are produced in the exact same format for both case study cities and there are no issues with the model when changing locations despite a geodesic distance of over 15,600 kilometers between the study cites. Based on the results, it can be

concluded that the model works consistently and as planned. It produces key information of the effects of natural hazards that are relevant to first response teams and humanitarian aid organizations considerably faster than other rapid mapping services such as the Copernicus Emergency Management Service (CEMS) and the United Nations Institute for Training and Research - Operational Satellite Applications Programme (UNITAR/UNOSAT) Rapid Mapping Service.

One function that was missing from all of these libraries was the ability to easily and accurately overlay raster files that have no data values. Also, the ability to upload the visualization to an online server for sharing would have been very useful, but there are so many costs involved in running a server that services with this option are subscription-based with monthly costs.

6 Discussion

6.1 The value of automated models in disaster risk reduction operations

The results of this thesis show that the automatization of analysis and visualization processes is possible and that it can speed up crisis mapping and increase situational awareness in disaster response efforts. The model can be run repeatedly as new SAR imagery of the disaster is acquired to produce up-to-date maps and the progression of a flooding event can be mapped with high precision and in a consistent format from start to finish. A timeline of the flooding event can then be analyzed after the disaster has passed to study how it progressed and how effects of similar natural hazards can be mitigated in future disaster risk reduction schemes and in the development of early warning systems. The model is open-source and free to use for everyone, especially in the Global South as in the developed countries most re-occurring natural hazards have been mapped. An example of this is the flood map service produced by the Finnish Environment Institute (SYKE) where the entirety of Finland has been mapped in regard to flood risk zones of yearly floods all the way to 1,000-year floods (Tulvakarttapalvelu 2022).

Creating an operational automated model is extremely time consuming and should only be done when working with extremely large amounts of geospatial data sets or with repetitive tasks. Automating parts of the analysis operations allows workers more time to focus on other tasks, such as acquiring new data and getting the information to the people in charge. Automated models reduce the dependence of experienced workers, and reduce the delay caused by data exchange, workflow construction and human error (Qi et al. 2017). Visual guidelines produced in cooperation with the crisis management community and experienced cartographers ensures that the crisis mapping is as efficient as possible and automated models will produce this predetermined and approved visual look perfectly every time when the model is configured to these requirements (Divjak & Lapaine 2018).

Previous studies have mainly focused on automating the early warning systems and flooding forecasting (Rodda 2005; Nguyen et al. 2019; Sinagra et al. 2020; Sarafanov et al. 2021). There is a lack of automated flood models that predict and analyze the effects of an on-going flooding event which is equally important as early warning systems rarely reach the most vulnerable people in society and these people often cannot or will not relocate from harm's way even if they were aware of approaching natural hazards. People will persist to live in high-risk areas as

these people have no alternatives and even if residents move out, new residents will move in to fill the voids (Panman et al. 2018). Ensuring that first responders and global humanitarian aid is focused on the most critically affected areas, and urgently reaches the most vulnerable and most in need of help, saves countless lives in times of crisis and may give these people, who have nothing to lose besides their lives, some hope for the future and assurance that they have not been completely forgotten.

6.2   Advantages and challenges of automated analysis and visualization

Using a Python script instead of a desktop GIS software (e.g., ArcGIS Pro by Esri, QGIS) to analyze the data and produce a map as an output enables the process to be automatized to an extent. The script does not have any triggers based on real-life measurements or parameters from external data sources, so it has to always be opened and run by a user. However, using this script, the user only has to change the raster files of the flood data and the population data. In the future, the WSF2019-Population data will hopefully be released as a GeoTIFF covering the entire globe which would eliminate the need to change the population raster data as the script calculates the extent of the flooded area and clips the population raster calculations to fit the extent of the captured flooding event. Automatizing the analysis and mapping process saves the user a lot of time and standardizes the output eliminating human errors from the process.

The output data can also be saved as a local file, using the Geopandas "geopandas.GeoDataFrame.to_file()" function if the user wants to select pieces of data for further analysis on desktop GIS software, as this script is intended to be used as a fast, first estimate of the scope of a disaster caused by a natural hazard. Data can also be saved for the creation of a historical geospatial database of flooding events. The HTML map produced by this script is also interactive, which provides crucial extra value compared to a static PDF or image, as users can zoom in very close and examine the entire flood event in much higher detail without being hampered by the visualization and image quality limitations of static maps. The basemap has been set to satellite imagery to allow for visual inspection of the real-world surroundings and to see where OSM data is lacking. Due to the lack of mapped OSM data, the estimation of inundated buildings in the two study sites is very imprecise. However, OSM is still the best readily available source of building footprints, and in a city of the Global North like Helsinki or London, where OSM data matches reality much more accurately, the estimation would be highly reliable. Thus, the need to map regions that are highly susceptible to natural

hazards in the Global South is very relevant and urgent, as urbanization increases, and new informal settlements are constructed in many of these rapidly expanding cities. The satellite imagery basemap used in the Kepler.gl map compensates for some of the weaknesses in using OSM data for building footprints.

The challenges of using Python automatization as the method of analysis include the possible future deprecations of certain Python modules which then require the script to be updated. There is also a possibility of Python libraries not being compatible with each other following future updates. Fixing these problems and incompatibilities requires knowledge of Python coding as does installing all of the Python libraries required to run the script. Lack of internet access or expensive internet contracts also limits the use of downloading OSM data and the downloading of Python libraries for installation or updating.

Performing accurate area calculations on a global map, mapping the data precisely to the correct geographical location, and focusing the interactive map on the correct geographical location was a challenging task to automatize, as this usually requires extensive knowledge of coordinate reference systems by the user and user input to match the data to the correct CRS. Successfully automatizing this process also eliminates user error from this part of the analysis and mapping which is important as the use of an incorrect CRS displaces the data very easily and can cause extreme distortions to calculations and analysis. Mapping a flood to a wrong country, calculating the flood extent incorrectly, or calculating the affected population inadequately would cause misinterpretations of data and possibly inadequate or excessive initial reactions to disasters caused by natural hazards.

Even though this automated flood mapping model was created with the intention of helping countries in the Global South to gain access to geospatial analysis and benefit from GIS, this model, like all the mapping of disasters caused by natural hazards, requires access to high-quality satellite data. Even if it is possible to edit and analyze this data using only open-source methods, it is impossible to escape the need for highly sophisticated and extremely expensive space technology to map the extent and depth of the floodwaters. Even though ICEYE provided their SAR data for this thesis free of cost, ICEYE is a private company and operates by selling their data to customers. For the purpose of openly sharing this model for further use, this is a major issue as up-to-date, preprocessed flood data from ICEYE is not open-source data. One option is to use SAR data from the Sentinel-1 satellite constellation operated by the European

Space Agency (ESA). It is free of charge to the general public, scientific and commercial users and can be delivered within an hour of reception for emergency response (Data Products 2022). The drawback is that Sentinel-1 data is not as precise as ICEYE's SAR data, and the revisit time is not as frequent as ICEYE's whose constellation is capable of persistent monitoring of the entire globe. ICEYE can also deliver data that is pre-processed and ready to use with the model developed in this thesis, while Sentinel-1 delivers raw SAR data or less processed data. Another option is to use flood depth and extent data that has been procured by other means. As long as it is in a geolocated raster format, it can be used as input data in this thesis. The model developed can also be quickly modified to accommodate flood data that does not have flood depth data but does show the extent of the flooding event.

To quickly acquire data in the Global South, affordable drones that fly a few hundred meters above the area can possibly be used, but using drones requires expertise and calm weather for acquiring needed images, considerable computer storage for storing the images, and a lot of time to process the images into a usable format. Countries in the Global South remain highly dependent on data that is only commercially available often with a high price, and usually available for free only long after the most urgent need has passed. Another option for them to get the needed data is through the certain humanitarian aid projects, provided to them for free in the form of raw data or as a result of analysis carried out to be used by them.

Models help us understand complex systems and are an especially effective approach to working with large-scale systems such as the Earth's climate. Once created, models can be used over and over without having to manually perform each individual step of the analysis and visualization. An individual step of a model can also be edited afterwards to examine different scenarios or to adjust it to new data. However, a model is also always a simplification of complex systems and a representation of some aspects of the world based on the model's creator's understanding of the relationships in question. Models are also only as good as the data it is based upon, meaning that the lack of OSM data is the biggest weakness of this model. It also has to be considered that when the situation with regards to lack of OSM data gets better and more data is available, it will also slow down the model as there will be more data to extract and process. Extracting OSM data from the entire extent of the province of Bangkok already takes 7 minutes and 40 seconds and only a marginal percent of buildings have been mapped in that area.

Possible future developments of this model could include adding road networks from OSM to highlight inundated sections of the road networks. This would help first responders in planning how to reach the affected area most efficiently and to plan efficient and safe evacuation routes. The inundated sections could be exported as shapefiles to be used as routing obstacles in network analyst tools like the one produced by Esri. It would also be better if the map could be uploaded automatically to an internet server for easier sharing of the flood map. This would eliminate the need to send or download any files between users. Making the map into a dashboard would also enable the adding of charts to the visualization to aid with the interpretation of the flooding event. After some historical data collection from several floods, it would be extremely interesting to try to automatize the flood model to produce early warnings of potential floods. This could probably be done by connecting to a feed of meteorological data sources, defining the parameters that would trigger the predictive analysis (e.g., a certain amount of rainfall in one hour), and by producing estimations based on historical flood data.

6.3   Geospatial data and proficiency in disaster risk reduction

Everything happens somewhere and that is the essence of GIS. Like the world-renowned geographic information system company Esri's slogan says, GIS is "The Science of Where" (Dangermond 2017). A vast majority of data is geospatial as it is tied to a certain place, no matter how feebly, which has most likely led to the famous quote that floats around in the field of geospatial work and research: "80 % of all data is geographic" (Dempsey 2012). If you imagine hearing about a flood event or any kind of disaster, the first questions often are: "Where did this happen? How many people are hurt? Can it have an affect where I live?"

In fact, all of these questions are geographic in nature and can be answered with the approaches and methods used in geospatial science. The flood can be predicted with an early warning analysis, the flood can be monitored with satellite imagery, such as SAR images, the number of people affected can be estimated using census data, estimates by the first responders at the scene, or existing GIS analyses of human populations (e.g., WSF data, WorldPop, High-Resolution Settlement Layer), while the extent of the flood or otherwise affected area can be mapped. Even if you tried your best, it would be impossible to organize disaster responses without using or creating any kind of geospatial data. Finding and focusing on the essential data for each use case is extremely important and valuable. Approaches focusing on digital data often tend to get enamored and swept away by all the available data and the result is a mishmash

of too much data drowning away any clear conclusions and points of action to focus on. Finding the real problems and relevant issues and focusing on highlighting them and on the resolutions of defined problems is what GIS analysis needs to be about. GIS is not a solution in itself, but rather an invaluable tool for discovering information and visualizing it in ways that are not possible by any other means. This is why it is extremely important to optimize and maximize the use of geospatial scientific knowledge and methods in disaster risk reduction, response, and recovery.

The Resilience Academy project, Ramani Huria community-mapping project in Dar es Salaam, Tanzania, and the Open Cities projects have produced excellent results by training university students and local community members to collect, create, and edit geospatial data (Käyhkö et al. 2018; Msilanga et al. 2021; About Open Cities Africa 2022; About Open Cities South Asia 2022; Ramani Huria 2022). There are multiple benefits to these approaches. They create new, urgently needed geospatial data relevant for the countries of the Global South, equip university students with geospatial skills, competencies, and abilities to increase future employment opportunities, and provide these countries with professionals who have the skills to push for sustainable transformations with information-based decision-making. As per Tobler's first law of geography that states "near things are more related than distant things" (Tobler 1970), people are most likely to voluntarily map areas closer to where they live in. Having more geospatial proficiency in the Global South, would provide OSM with data that has been produced with the crucial knowledge of local context and its rapidly changing landscapes. This thesis also showcases many of the possibilities and advantages of using only Python for GIS analysis and the need to also include Python coding also in advanced-level GIS courses at the universities in the countries of the Global South.

7 Conclusion

This thesis aimed at developing an automated global geospatial dataset analysis in disaster risk reduction in the Global South, by using the Python programming language to produce an automated flood analysis and visualization model. The model ingests raster data of a flooding event and of population estimates provided by the user and produces ready to use GIS data, statistics of the effects of a disaster caused by flooding, and an interactive map of the scope of the disaster. By using global geospatial data sets, the model can produce analysis from anywhere in the world in minutes to be used by first response efforts and global humanitarian aid. The combination of commercially produced SAR satellite data by ICEYE, open-source population estimation data by the German Aerospace Agency (DLR) and volunteered geographic information (VGI) from the OpenStreetMap (OSM) project, this model can be seen as a combination of approached from the two paradigms of disaster risk reduction: the hazard paradigm and the vulnerability paradigm.

The automated flood analysis and visualization model developed in this thesis works as planned. It performs consistently and the accuracy is only dependent on the accuracy of the data used. With minimal user input, the room for human error has been minimized and the results are standardized. This model is suitable to be used to create the first assessment of a flooding event anywhere in the world, as long as it is understood and accepted that the population estimates are never 100 percent correct and that the accuracy of the number of inundated buildings is highly dependent on the amount of OSM data from the affected area. The model provides essential information for decision-makers; the geographical extent of the flood, floodwater depths and their geographical location, an estimation of the total number of people affected and how severely their homes have been inundated, and how severely other kinds of buildings and essential services have been inundated. Using global data sets or quickly acquiring new local data with satellites that have a global reach eliminates the most time-consuming part of GIS analysis, which is how to find suitable data. Automating the analysis and visualization with Python also eliminates the other most time-consuming parts: cleaning up the data and visualizing all the separate data sets.

The model is aimed to be used at organizations with a global reach and influence and sufficient knowledge of GIS and Python. A GIS expert working only locally might have easy access to better data that can be used and analyzed with higher precision and accuracy and can probably

contact relevant authorities with estimates at almost equal speeds to the model created in this thesis. However, in countries of the Global South, this expertise and data are often lacking, and humanitarian aid organizations working on a global scale often depend on global geospatial data sets for needed information and analysis. Geospatial approaches to disaster risk reduction must be collaborative efforts by global and local stakeholders in order to utilize the newest, most efficient technology and methods while taking into consideration the complex needs and concerns of local inhabitants most affected by disasters caused by natural hazards.

Acknowledgments

References

About OpenStreetMap (2022) OpenStreetMap Wiki 14.01.2022 <*https://wiki.openstreetmap.org/wiki/About_OpenStreetMap*>. 18.01.2022.

About Open Cities Africa (2022) Open Cities Africa. <*https://opencitiesproject.org/about/*>. 04.05.2022.

About Open Cities South Asia (2022) Open Cities. <*https://opencitiesproject.github.io/about/*>. 04.05.2022.

About Resilience Academy (2022) Tanzania Resilience Academy. <*https://resilienceacademy.ac.tz/about-us/*>. 04.02.2022.

About Us (2022) Dar Ramani Huria. <*https://ramanihuria.org/en/about-us/*>. 01.05.2022.

Alfieri L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K. & Feyen L. (2016) Global projections of river flood risk in a warmer world. *Earth's Future* 5 171–181. https://doi.org/10.1002/2016EF000485

Al-Tahir, R. & R.S. Mahabir (2014) Applications of Remote Sensing and GIS Technologies in Flood Risk Management. In Chadee, D.D., Sutherland, J.M. & Agard, J.B. (eds.) *Flooding and Climate Change Sectorial Impacts and Adaptation Strategies for the Caribbean Region*, 138–150. Nova Publishers. https://www.researchgate.net/publication/261000276_Applications_of_Remote_Sensing_and_GIS_Technologies_in_Flood_Risk_Management

Aman, H., Irani, P. & Amini, F. (2014) Revisiting Crisis Maps with Geo-Temporal Tag Visualization. *IEEE Pacific Visualization Symposium*. https://doi.org/10.1109/PacificVis.2014.55

Anderson, J., Sarkar, D. & Palen, L. (2019) Corporate Editors in the Evolving Landscape of OpenStreetMap. *International Journal of Geo-Information* 8(5) 232–250. https://doi.org/10.3390/ijgi8050232

Berquist, M., Daniere A. & Drummond L. (2014) Planning for global environmental change in Bangkok's informal settlements. *Journal of Environmental Planning and Management* 58(10) 1711–1730. https://doi.org/10.1080/09640568.2014.945995

Birkmann, J., Sauter, H., Jamshed, A., Sorg, L., Fleischhauer, M., Sandholz, S., Wannewitz, M., Greiving, S., Bueter, B., Schneider, M. & Garschagen, M. (2020) Strengthening risk-informed decision-making: scenarios for human vulnerability and exposure to extreme events. *Disaster Prevention and Management* 29(5) 663–679. https://doi.org/10.1108/DPM-05-2020-0147

Boeing, G. (2017) OSMnx: New Methods for Aquiring, Constructing, Analyzing, and Visualizing Complex Street Networks. *Computers, Environment and Urban Systems* 65 126–139. https://doi.org/10.1016/j.compenvurbsys.2017.05.004

Building Footprints (2022) Microsoft. <*https://www.microsoft.com/en-us/maps/building-footprints*>. 12.04.2022.

Carlitz, R.D. & McLellan, R. (2020) Open Data from Authoritarian Regimes: New Opportunities, New Challenges. *Perspectives on Politics* 19(1) 160–170. https://doi.org/10.1017/S1537592720001346

Coates, R. (2022) Infrastructural events? Flood disaster, narratives and framing under hazardous urbanization. *The International Journal of Disaster Risk Reduction* 74. https://doi.org/10.1016/j.ijdrr.2022.102918

Completing the picture: SAR Product Guide (2021) Version 4.2, ICEYE. <*https://www.iceye.com/hubfs/Downloadables/ICEYE_SAR_Product_Guide_2021_V4.0.pdf*>. 22.04.2022.

Dangermond, J. (2017) The Science of Where: Our Promise. Esri Technology <*https://www.esri.com/about/newsroom/arcuser/the-science-of-where-our-promise/*>. 12.05.2022.

Data Products (2022) The European Space Agency <*https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/data-products*>. 18.05.2022.

Dempsey, C. (2012) Where is the Phrase "80% of Data is Geographic" From? GIS Lounge 28.10.2012 <*https://www.gislounge.com/80-percent-data-is-geographic/*>. 12.05.2022.

Depietri, Y., Renaud, F.G. & Kallis, G. (2011) Heat waves and floods in urban areas: a policy-oriented review of ecosystem services. *Sustainability Science* 7 5–107. https://doi.org/10.1007/s11625-011-0142-4

Disaster risk reduction (2021) United Nations Department of Economic and Social Affairs. <*https://sdgs.un.org/topics/disaster-risk-reduction*>. 21.10.2021.

Divjak, A.K. & Lapaine, M. (2018) Crisis Maps–Observed Shortcomings and Recommendations for Improvement. *International Journal of Geo-Information* 7(11) 436. https://doi.org/10.3390/ijgi7110436

Doocy, S., Daniels, A., Murray, S. & Kirsch, T.D. (2013) The Human Impacts of Floods: a Historical Review of Events 1980-2009 and a Systematic Literature Review. *PLOS Currents Disasters* 1 1–14. https://doi.org/10.1371/currents.dis.f4deb457904936b07c09daa98ee8171a

Epic of Atra-Hasis (17th century BCE) *Myths from Mesopotamia: Creation, The Flood Gilgamesh, and Others. A new translation by Stephanie Dalley*. World Classics, Oxford. <*https://geha.paginas.ufsc.br/files/2017/04/Atrahasis.pdf*>

Evans, W.P. (2020) Geospatial Data in Authoritarian Contexts. EthicalGEO 12.06.2020. <*https://ethicalgeo.org/geospatial-data-in-authoritarian-contexts/*>. 11.05.2022.

Extreme flood events once again drive high losses in 2021, yet 75% of flood risks remain uninsured, Swiss Re Institute reveals (2022) Swiss Re Group 30.03.2022 <*https://www.swissre.com/press-release/Extreme-flood-events-once-again-drive-high-losses-in-2021-yet-75-of-flood-risks-remain-uninsured-Swiss-Re-Institute-reveals/3269ad99-b743-4398-82e3-534a87783910*>. 12.04.2022.

FAQ (2022) OpenStreetMap Wiki. <*https://wiki.openstreetmap.org/wiki/FAQ*>. 18.1.2022.

Falco, E., Zambrano–Verratti, J. & Kleinhans, R. (2020) Web-based participatory mapping in informal settlements: The slums of Caracas, Venezuela. *Habitat International* 94 1–10. https://doi.org/10.1016/j.habitatint.2019.102038

Fekade, W. (2000) Deficits of formal urban land management and informal responses under rapid urban growth, an international perspective. *Habitat International* 24(2) 127–150. https://doi.org/10.1016/S0197-3975(99)00034-X

Ferro-Famil, O. & Pottier, E. (2016) 1 – Synthetic Aperture Radar Imaging. *Microwave Remote Sensing of Land Surface, Techniques and Methods* 1–65. https://doi.org/10.1016/B978-1-78548-159-8.50001-3

Financial Management of Flood Risk (2016) OECD 1–134 <*https://www.oecd.org/daf/fin/insurance/Financial-Management-of-Flood-Risk.pdf*>

Flood Monitorin (2022) ICEYE <*https://www.iceye.com/solutions/flood-monitoring*>. 11.05.2022.

Folium (2022) Folium documentation. <*https://python-visualization.github.io/folium/*>. 12.04.2022.

Fuentes, E. (2019) Why GPS Coordinates Look Wrong on Maps of China. ServiceObjects 19.09.2019 <*https://www.serviceobjects.com/blog/why-gps-coordinates-look-wrong-on-maps-of-china/*>. 11.05.2022.

Furin, M. (2018) Disaster Planning. *Medscape* 30.07.2018 <*https://emedicine.medscape.com/article/765495-overview*>. 4.11.2021.

Gaillard, J.C. & Mercer, J. (2012) From knowledge to action: Bridging gaps in disaster risk reduction. *Progress in Human Geography* 37(1) 93–114. https://doi.org/10.1177/0309132512446717

General Python FAQ (2022) Python <*https://docs.python.org/3/faq/general.html*>. 06.04.2022.

Gies, E. (2018) Sponge Cities Can Limit Urban Floods and Droughts. *Scientific American* 319(6) 80–85. 01.12.2018 <*https://www.scientificamerican.com/article/sponge-cities-can-limit-urban-floods-and-droughts/*>. 18.2.2022.

Goh, K. (2019) Urban waterscapes: The hydro-politics of flooding in a sinking city. *International Journal of Urban and Regional Research*. 42(2) 250–271. https://doi.org/10.1111/1468-2427.12756

Gray, A. (2018) Bangkok is sinking. Here's how a new park can protect the city from flooding. World Economic Forum18 10.09.2018 <*https://www.weforum.org/agenda/2018/09/bangkok-has-created-a-sponge-park-to-combat-future-flooding/*> 05.05.2022.

Greater Impact: How Disasters Affect People of Low Socioeconomic Status (2017) Disaster Technical Assistance Center Supplemental Research Bulletin, SAMHSA. <*https://www.samhsa.gov/sites/default/files/dtac/srb-low-ses_2.pdf*>. 29.04.2022.

Guillén, B. (2021) El Gobierno reconoce en un informe que la inundación de Tula se debió a la descarga en exceso de agua del Valle de México. El País 17.10.2021 <*https://elpais.com/mexico/2021-11-17/el-gobierno-reconoce-en-un-informe-que-la-inundacion-de-tula-se-debio-a-la-descarga-en-exceso-de-agua-del-valle-de-mexico.html*>. 29.04.2022.

Hernaiz, R.S.M. (2019) Enhancing geospatial preparedness for disaster management through the work of development organisations. Doctoral Thesis in Information Management. NOVA Information Management School, Universidade Nova de Lisboa. <*https://run.unl.pt/bitstream/10362/98700/1/D0055.pdf*>

Huizinga, J., de Moel, H. & Szewczyk, W. (2017) Global flood depth-damage fuctions: Methodology and the database guidelines. *Joint Research Center (JRC) Technical Report* 1–110. EUR 28552, https://doi.org/10.12760/16510

Hyndman, D. & Hyndman, D. (2017) *Natural Hazards and Disasters* 5th ed. 540. Cengage Learning. Boston, MA, USA.

ICEYE Imagery Archive – 18,000 SAR Satellite Image Thumbnails (2022) ICEYE <*https://www.iceye.com/lp/iceye-18000-public-archive*>. 11.05.2022.

Independent Evaluation of UNOSAT Rapid Mapping Service – Final Report (2018) UNITAR <*https://unitar.org/sites/default/files/media/file/independent_evaluation_of_unosat_rapid_mapping_service_final_report.pdf*>. 12.05.2022.

Jackson, G., McNamara, K. & Witt, B. (2017) A framework for natural disaster vulnerability in a small island in the Southwest Pacific: a case study of Emae Island, Vanuatu. *International Journal of Disaster Risk Science* 8 358–373. https://doi.org/10.1007/s13753-017-0145-6

JupyterLab (2022) Jupyter. <*https://jupyter.org*> 11.04.2022.

Kavvada, A., Metternicht, G., Kerblat, F., Mudau, N., Haldorson, M., Laldaparsad, S., Friedl, L., Held, A. & Chuvieco, E. (2020) Towards delivering on the Sustainable Development Goals using Earth observations. *Remote Sensing of Environment* 247 1–8. https://doi.org/10.1016/j.rse.2020.111930

Kepler.gl (2022) Kepler.gl. <*https://docs.kepler.gl*>. 22.04.2022.

Kitajima, N., Seto, R., Yamazaki, D., Zhou, X., Ma, W. & Kanae, S. (2021) Potential of a SAR Small-Satellite Constellation for Raping Monitoring of Flood Extent. *Remote Sensing* 13(10) 1959. https://doi.org/10.3390/rs13101959

Klijn, F., Asselman, N. & Wagenaar, D. (2018) Room for Rivers: Risk Reduction by Enhancing the Flood Conveyance Capacity of The Netherlands' Large Rivers. *Geosciences* 8(6) 1–20. https://doi.org/10.3390/geosciences8060224

Krellenberg, K., Welz, J., Link, F. & Barth, K. (2017) Urban vulnerability and the contribution of socio-environmental fragmentation: theoretical and methodological pathways. *Progress in Human Geography* 41(4) 408–431. https://doi.org/10.1177/2F0309132516645959

Kuffer, M., Thomson, D.R., Boo, G., Mahabir, R., Grippa, T., Vanhuysse, S., Engstrom, R., Ndugwa, R., Makau, J., Darin, E., Porto de Albuquerque, J. & Karabia, C. (2020) The Role of Earth Observation in an Intergrated Deprived Area Mapping "System" for Low-to-Middle Income Countries. *Remote Sensing* 12(982) 1–26. https://doi.org/10.3390/rs12060982

Käyhkö, N., William, C., Mayunga, J., Makame, M.O., Mauya, E. & Järvi, A. (2018) Building Geospatial Competences in Tanzanian Universities With Open Source Solutions. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-4/W8 93–99. https://doi.org/10.5194/isprs-archives-XLII-4-W8-93-2018

Learn Synthetic Aperture Radar (SAR) by Example (2021) GISGeography 29.10.2021 *<https://gisgeography.com/synthetic-aperture-radar-examples/>*. 21.10.2021.

Major OpenStreetMap Consumers (2022) OSMWiki 11.02.2021 *<https://wiki.openstreetmap.org/wiki/Major_OpenStreetMap_Consumers>* 22.04.2022.

Make cities and human settlements inclusive, safe, resilient, and sustainable (2021) United Nations *<https://sdgs.un.org/goals/goal11>* 20.10.2021.

Mapathon (2022) Kartta.nyt *<https://blog.edu.turku.fi/karttanyt/mapathon/>*. 04.02.2022.

Marchezini, V. (2019) The power of localism during the long-term disaster recovery process. *Disaster Prevention and Management* 28(1) 143–152. https://doi.org /10.1108/DPM-05-2018-0150

Marks, D. (2015) The Urban Political Ecology of the 2011 Floods in Bangkok: The Creation of Uneven Vulnerabilities. *Pacific Affairs* 88(3) 623–651. https://doi.org/10.5509/2015883623

Maxar is a leading space technology and intelligence company (2022) Maxar. *<https://www.maxar.com/about>*. 27.04.2022.

Meier, P. (2011) New information technologies and their impact on the humanitarian sector. *International Review of the Red Cross* 93(884) 1239–1263. https://doi.org/10.1017/S1816383112000318

Met office award for 200 years of continuous weather observations at Oxford (2015) School of Geography and the Environment (SoGE) 15.05.2015 *<https://www.geog.ox.ac.uk/news/articles/150515-met-office-award.html>*. 29.03.2022.

Migliorini, M., Hagen, J.S., Mihaljevic, J., Mysiak, J., Rossi, J–L., Siegmund, A., Meliksetian K. & Sapir, D.G. (2019) Data interoperability for disaster risk reduction in Europe. *Disaster Prevention and Management* 28(6) 804–816. https://doi.org/10.1108/DPM-09-2019-029

Molle, F. (2009) River-basin planning and management: The social life of a concept. *Geoforum* 40(3), 484–494. https://doi.org/10.1016/j.geoforum.2009.03.004

Molle, F., Mollinga, P. & Wester, P. (2009) Hydraulic bureaucracies and the hydraulic mission: Flows of water, flows of power. *Water Alternatives* 2(3), 328–349.

Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I. & Papathanassiou (2013) A Tutorial on Synthetic Aperture Radar. *IEEE Geoscience and Remote Sensing Magazine* 1(1) 6–43. https://doi.org/10.1109/MGRS.2013.224830

Mosavi, A., Ozturk, P. & Chau, K. (2018) Flood prediction using machine learning models: a literature review. *Water* 10(1536) 1–40. https://doi.org/10.3390/w10111536

Msilanga, M., Käyhkö, N., Mbise, M,. Ngereja, Z., Makame, M.O. & Mauya, E. (2021) Resilience Academy Student Internship Model As A Innovative Way To Enhance Geospatially Literate Future Work Force In Africa. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLVI-4/W2-2021 115–117. https://doi.org/10.5194/isprs-archives-XLVI-4-W2-2021-115-2021

Murnane, R.J., Allegri, G., Bushi, A., Dabbeek, J., de Moel, H., Duncan, M., Fraser, S., Galasso, C., Giovando, C., Henshaw, P., Horsburgh, K., Huyck, C., Jenkins, S., Johnson, C., Kamihanda, G., Kijazi, J., Kikwasi, W., Kombe, W., Loughlin, S., Løvholt, F., Masanja, A., Mbongoni, G., Minas, S., Msabi, M., Msechu, M., Mtongori, H., Nadim, F., O'Hara, M., Pagani, M., Phillips, E., Rossetto, T., Rudari, R., Sangana, P., Silva, V., Twigg, J., Uhinga, G. & Verucci, E. (2019) Data schemas for multiple hazards, exposure and vulnerability. *Disaster Prevention and Management* 28(6) 752–763. https://doi.org/10.1108/DPM-09-2019-0293

Navigating OpenStreetMap: How to make the most of an increasingly valuable mapping asset (2020) Accenture. <*https://wiki.openstreetmap.org/w/images/8/80/Accenture-Navigating-OpenStreetMap.pdf*>. 22.04.2022.

Nguyen, K.L., Nguyen, D.L., Le Hoang, T.,Nguyen, T.H., Cao, D.T., Vo Ngoc, G.T., Tran, T.N., Tran, N.A. & Jaehak, J. (2019) Automated procedure of real-time flood forecasting in Vu Gia – Thu Bon river basin, Vietnam bby intergrating SWAT and HEC-RAS models. *Journal of Water and Climate Change* 10(3) 535–545. https://doi.org/10.2166/wcc.2018.015

Nomura, S., Parsons, A.J.Q., Hirabayashi, M., Kinoshita, R., Liao, Y. & Hodgson, S. (2016) Social determinants of mid-to-long-term disaster impacts on health: A systematic review. *International Journal of Disaster Risk Reduction* 16 53–67. https://doi.org/10.1016/j.ijdrr.2016.01.013

Norton, B., Bosomworth, K., Coutts, A., Williams, N., Livesley, S., Trundle, A., Harris, R. & McEvoy, D. (2013) Planning for a cooler future: green infrastructure to reduce urban heat. *Climate Adaptation for Decision-makers. Victorian Centre for Climate Change Adaptation (VCCCAR)* 1–29

Nygren, A. (2016) Socially differentiated urban flood governance in Mexico: Ambiguous negotiations and fragmented contestations. *Journal of Latin American Studies* 48(2), 335–365. https://doi.org/10.1017/S0022216X15001170

Palacios-Lopez, D., Bachofer, F., Esch, T., Marconcini, M., MacManus, K., Sorichetta, A., Zeidler, J., Dech, S., Tatem, A.J. & Reinartz, P. (2021) High-Resolution Gridded Population Data sets: Exploring the Capabilities of the World Settlement Footprint 2019 Imperviousness Layer for the African Continent. *Remote Sensing* 13(1142) 1–26. https://doi.org/10.3390/rs13061142

Panman, A., Madison, I. & Falisse, J-B. (2018) Why do people live in flood-prone areas? Reflections from Dar es Salaam. World Bank Blogs 14.12.2018 <*https://blogs.worldbank.org/nasikiliza/why-do-people-live-in-flood-prone-areas-reflections-from-dar-es-salaam*>. 19.05.2022.

Penrose, K., Caldas de Castro, M., Werema, J. & Ryan, E.T. (2010) Informal urban settlements and cholera risk in Dar es Salaam, Tanzania. *PLoS Neglected Tropical Diseases* 4(3) 1–11. https://doi.org/10.1371/journal.pntd.0000631

Persistent Monitoring by ICEYE (2021) ICEYE. <*https://www.iceye.com/persistent-monitoring*>. 30.09.2021

Persistent monitoring enables a totally new level of strategical decision-making (2022) ICEYE. <*https://www.iceye.com/the-applications*>. 30.03.2022.

Pistrika, A., Tsakiris, G. & Nalbantis, I. (2014) Flood Depth-Damage Functions for Built Environment. *Environmental Processes* 1 553–572. https://doi.org/10.1007/s40710-014-0038-2

Puteh, S.E.W., Siwar, C., Hod, R., Nawi, A.M., Idris, I.B., Ahmad, I.S., Idris, N.D.M., Alias, N.A. & Taha, M.R. (2019) Burden of health-related issues and community empowerment in Malaysia's east coast flood. *Improving Flood Management, Prediction, and Monitoring: Case Studies in Asia* 20 19–26. https://doi.org/10.1108/S2040-726220180000020011

Qui, L., Du, Z., Zhu, Q. & Fan, Y. (2017) An integrated flood management system based on linking environmental models and disaster-related data. *Environmental Modelling & Software* 91 111–126. https://doi.org/10.1016/j.envsoft.2017.01.025

Qui, Y., Zhao, X., Fan, D. & Li, S. (2019) Geospatial Disaggregation of Population Data in Supporting SDG Assessments: A Case Study from Deqing County, China. *International Jouranl of Geo-Information* 8(356) 1–17. https://doi.org10.3390/ijgi8080356

Rapid Mapping (2022) Copernicus Emergency Management Service. <*https://emergency.copernicus.eu/mapping/ems/rapid-mapping-portfolio*>. 24.04.2022.

Raziel, Z. (2021) Conagua y Sacmex descargaron e inundaron con aguas negras a Tula; se planeó así para salvar al Valle de México. Animal Politico 11.10.2021 <*https://www.animalpolitico.com/2021/11/conagua-sacmex-inundaron-aguas-negras-tula-valle-mexico-cdmx/*>. 29.04.2022.

Resilience Academy (2022) Frontpage, Resilience Academy <*https://resilienceacademy.ac.tz*>. 12.05.2022.

R&D Outputs: Flood Risks To People: Phase 2 FD2321/TR2 – Guidance document (2006) HR Wallingford, Defra / Environmental Agency 1–91. <*https://www.gov.uk/flood-and-coastal-erosion-risk-management-research-reports/flood-risks-to-people-phase-2-managing-risks-and-dangers*>

Rodda, H.J.E. (2005) The Development and Application of a Flood Risk Model for the Czech Republic. *Natural Hazards* 24 207–220. https://doi.org/10.1007/s11069-004-4549-4

Rosen, J. (2021) Shifting ground. Science 26.02.2021 <*https://www.science.org/doi/full/10.1126/science.371.6532.876*>. 18.05.2022.

Rowland-Shea, J., Doshi, S., Edberg, S. & Fanger, R. (2020) The Nature Gap: Confronting Racial and Economic Disparities in the Destruction and Protection of Nature in America. Center for American Progress 21.07.2020 <*https://www.americanprogress.org/article/the-nature-gap/*> 21.1.2022.

Sala, S. & Dendena, B. (2015) Geographical Information Systems in the Global South. *The International Encyclopedia of Digital Communication and Society* 1–12. https://doi.org/10.1002/9781118767771.wbiedcs089

Sadafumi, T. (2017) Japan's Coming Population Implosion. Nippon.com 27.07.2017 <*https://www.nippon.com/en/currents/d00336/*>. 20.1.2022

Sakijege, T., Sartohadi, J., Marfai, M.A., Kassenga, G.R. & Kasala, S.E. (2014) Assessment of adaptation strategies to flooding: a comparative study between informal settlements of Keko Machungwa in Dar es Salaam and Sangkrah in Surakarta, Indonesia. *Jàmbá: Journal of Disaster Risk Studies* 6(1) 1–10. http//dx.doi.org/10.4102/jamba.v6i1.131

Sakijege, T., Lupala, J. & Sheuya, S. (2012) Flooding, flood risk and coping strategies in urban informal residential areas: The case of Keko Machungwa, Dar es Salaam, Tanzania. *Jàmbá: Journal of Disaster Risk Studies* 4(1) 1-10. https://doi.org/10.4102/jamba.v4i2.46

Sarafanov, M., Borisova, Y., Maslyaev, M., Revin, I., Maximov, G. & Nikitin, N.O. (2021) *Water* 12(24) 3482. https://doi.org/10.3390/w13243482

Schröter, E., Kiefl, R., Neidhardt, E., Gurczik, G., Dalaff, C. & Lechner, K. (2020) Trialing Innovative Technologies in Crisis Management–"Airborne and Terrestrial Situational Awareness" as Support Tool in Flood Response. *Applied Sciences* 10(11) 3743. https://doi.org/10.3390/app10113743

Schwartz, L., Hansen, K., Alari, A., Ilango, S.D., Bernal, N., Basu, R., Gershunov, A. & Benmarhnia, T. (2021) Spatial variation in the joint effect of extreme heat events and ozone respiratory hospitalizations in California. *PNAS* 118(22) 1–9. https://doi.org/10.1073.pnas2023078118

Sendai Framework for Disaster Risk Reduction 2015–2030 (2015) United Nations. <*https://www.preventionweb.net/files/43291_sendaiframeworkfordrren.pdf*>

Seventeen people, most with COVID-19, die in flooding of Mexican hospital (2021) Reuters 08.09.2021 <*https://www.reuters.com/world/americas/17-people-die-mexican-hospital-due-severe-flooding-2021-09-07/*>. 31.03.2022.

Silva, C.d.S., Vegas, I., Panagopoulos, T. & Bell, S. (2018) Environmental Justice in Accessibility to Green Infrastructure in Two European Cities. *Land* 7(134) 1–23. 10.3390/land7040134fykm1s

Silva, N. & Mena, C. (2020) Identifying the underlying risk factors of local communities in Chile. *Disaster Prevention and Management* 29(5) 681–696. https://doi.org/10.1108/DPM-04-2020-0105

Sinagra, M., Nasello, C., Tucciarelli, T., Barbetta, S., Massari, C. & Moramarco, T. (2020) A Self-Contained and Automated Method for Flood Hazard Maps Prediction in Urban Areas. *Water* 12(5) 1266. https://doi.org/10.3390/w12051266

Smith, A., Bates, P.D., Wing, O., Sampson, C., Quinn, N. & Neal, J. (2019) New estimates of flood exposure in developing countries using high-resolution population data. *Nature Communications* 10(1814) 1–7. https://doi.org/10.1038/s41467-019-09282-y

Sudden Onset (2022) UNICEF. <*https://www.unicef.org/innovation/sudden-onsets*>. 28.04.2022.

Thailand (2022) The World Factbook, CIA. <*https://www.cia.gov/the-world-factbook/countries/thailand/*>. 31.03.2022.

Thatcher, J., O'Sullivan, D. & Mahmoudi, D. (2016) Data colonialism through accumulation by dispossession: New metaphors for daily data. *Environment and Planning D: Society and Space* 34(6) 990–1006. https://doi.org/10.1177/0263775816633195

The Emergency Management Service – Mapping (2022) Copernicus, European Comission <*https://emergency.copernicus.eu/mapping/ems/emergency-management-service-mapping*>. 24.04.2022.

The Msimbazi opportunity (2019) VE-R. <*https://www.ve-r.nl/projects-all/ecology/the-msimbazi-opportunity/*>. 17.2.2022

The power of maps (2022) Esri 23.03.2022. <*https://learn.arcgis.com/en/projects/the-power-of-maps/*>. 18.05.2022.

Tierney, K.J. (2012) Disaster Governance: Social, Political, and Economic Dimensions. *Annual Review of Environment and Resources* 37 341–363. https://doi.org/10.1146/annual-environ-020911.095618

Tierney, K.J. (2007) From the Margins to the Mainstream? Disaster Research at the Crossroads. *Annual Review of Sociology* 33 503–525. https://doi.org/10.1146/annurev.soc.33.040406.131743

TIFF files. (2022) Adobe. <*https://www.adobe.com/creativecloud/file-types/image/raster/tiff-file.html*>. 11.04.2022.

Tobler, W. (1970) A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46(Supplement) 234–240

Tomás, R. & Li, Z. (2017) Earth Observations for Geohazards: Present and Future Challenges. *Remote Sensing* 9(3) 1–10. https://doi.org/10.3990/rs9030194

Transforming Tanzania's Msimbazi River from a Liability into an Opportunity (2019) The World Bank 12.08.2019 <*https://www.worldbank.org/en/news/feature/2019/08/12/transforming-tanzanias-msimbazi-river-from-a-liability-into-an-opportunity*>. 17.02.2022

Tula de Allende (2020) DataMÉXICO. <*https://datamexico.org/en/profile/geo/tula-de-allende*> 31.03.2022.

Tulvakarttapalvelu (2022) Finnish Environmental Insitute (SYKE) <*https://paikkatieto.ymparisto.fi/tulvakartat/Viewer/Viewer.html?configBase=https://paikkatieto.ymparisto.fi/Geocortex/Essentials/REST/sites/Tulvakarttapalvelu/viewers/HTML5/virtualdirectory/Resources/Config/Default/*>. 18.05.2022.

Uitto J.I., & Shaw, R. (2016) Sustainable Development and Disaster Risk Reduction. Tokyo, *Springer*

UNOSAT Geospatial Catalogue to Support Humanitarian Operations (2022) UNOSAT <*https://unosatgis.cern.ch/agol/UNOSAT%20Geospatial%20Catalogue/UNOSAT%20Geospatial%20Catalogue%20to%20Support%20Humanitarian%20Operations.pdf*>. 12.05.2022

UNOSAT Raping Mapping Service (2022) United Nations Institute for Training and Research (UNITAR) <*https://www.unitar.org/maps/unosat-rapid-mapping-service*>. 12.05.2022.

Venter, Z.S., Shackleton, C.M., Van Staden, F., Selomane, O. & Masterson, V.A. (2020) Green Apartheid: Urban green infrastructure remains unequally distributed across income and race geographies in South America. *Landscape and Urban Planning* 203 1–12. https://doi.org/10.1016/j.landurbplan.2020.103889

Watson, V. (2020) Digital Visualisation as a New Driver of Urban Change in Africa. *Urban Planning* 5(2) 35–43. 10.17645/up.v5i2.2989

What is Coherent Change Detection (2022) BAE Systems. <*https://www.baesystems.com/en-us/definition/what-is-coherent-change-detection*>. 22.04.2022.

What is Disaster Risk Reduction? (2021) United Nations International Strategy for Disaster Reduction. <*https://eird.org/esp/acerca-eird/liderazgo/perfil/what-is-drr.html*>. 21.10.2021.

What is Synthetic Aperture Radar? (2020) NASA 16.04.2020 <*https://earthdata.nasa.gov/learn/backgrounders/what-is-sar*>. 30.09.2021.

Wilkinson, F. (2021) How a tiny line on a map led to conflict in the Himalaya. *National Geographic* 18.02.2021 <*https://www.nationalgeographic.com/magazine/article/how-a-tiny-line-on-a-map-led-to-conflict-in-the-himalaya-feature*>. 19.01.2022.

Win, T.L. (2017) Bangkok struggles to protect slum dwellers as floods worsen. Reuters 19.07.2017 <*https://www.reuters.com/article/us-thailand-floods-bangkok-idUSKBN19A0KL*>. 26.1.2022

Yap, K.S. & De Wandeler, K. (2010) Self-help housing in Bangkok. *Habitat International* 34(3) 332–341. https://doi.org/10.1016/j.habitatint.2009.11.006

Young, J.C., Lynch, R., Boakye–Achampong, S., Jowaisas, C., Sam, J. & Norlander, B. (2020) Volunteer geographic information in the Global South: barriers to local implementation of mapping projects across Africa. *GeoJournal* 86 2227–2243. https://doi.org/10.1007/s10708-020-10184-6

Young, J.C. (2019) The new knowledge politics of digital colonialism. *Environment and Planning A: Economy and Space* 51(7) 1424–1441. https://doi.org/10.1777/0308518X19858998

Zieleniec, A. (2018) Lefebvre's Politics of Space: Planning the Urban as Oeuvre. *Urban Planning* 3(3) 5–15. https://doi.org/10.17645/up.v3i3.1343

Appendices

Appendix 1. Automated flood analysis and visualization model code

Automating Global Geospatial Data Set Analysis

Visualizing flood disasters in the cities of the Global South

This code will turn raster data into an interactive local Kepler.gl web map displaying the extent of flooding events and its effects on the local population and infrastructure.

The script is used as the research method in the Master's thesis of Ohto Nygren for the Department of Geography at the University of Turku, Finland.

Flood data has been provided by the new space company ICEYE and the world settlement footprint (WSF) population data that was used was provided by the German Aerospace Agency (DLR). The script has been optimized for these data sets, but similar data can most likely be used with the script.

The OSMnx module is used to pull data from OpenStreetMap (OSM) to get estimations of inundated buildings, with an emphasis on hospitals and pharmacies.

Estimations provided by the script are not completely accurate due to the lack of mapped OSM data, but flood depths captured by ICEYE's SAR satellites are very accurate and the population data created by the DLR is very accurate for a global dataset.

The main idea is to provide quick first estimations via Python automatization of the extent of flood damage and the number of people affected by a flooding event to first responders and international humanitarian aid.

```python
import pandas as pd
import geopandas as gpd
import matplotlib
import matplotlib.pyplot as plt
import osmnx as ox
import shapely
from shapely.geometry import Point, LineString, Polygon, shape
from shapely.geometry import box
from shapely import speedups
import rasterio as rio
import rasterio.features
from rasterio.plot import show
from rasterio.plot import show_hist
from rasterio.features import shapes
from rasterio.mask import mask
from fiona.crs import from_epsg
import pycrs
import os
import mapclassify
import seaborn as sns
import utm
from pyproj import CRS
```

Raster transformations

```python
# Filepath for WSF raster file
wsf_fp =
r'/Users/ohtonygren/Yliopisto/Gradu/Data/WSF/Tula_WSF2019population_10m.tif
'
wsf_raster = rio.open(wsf_fp)

# Filepath for ICEYE raster file
iceye_fp =
r'/Users/ohtonygren/Yliopisto/Gradu/Data/ICEYE/ICEYE_tula_wd_v_2_r_01_ws_1.
tif'
iceye_orig = rio.open(iceye_fp)

from rasterio.warp import calculate_default_transform, reproject,
Resampling

# Getting CRS from the WSF data
dstCrs = wsf_raster.crs

# Calculate transform array and shape of reprojected layer
transform, width, height = calculate_default_transform(iceye_orig.crs,
dstCrs, iceye_orig.width, iceye_orig.height, *iceye_orig.bounds)

# Working of the meta for the destination raster
kwargs = iceye_orig.meta.copy()
kwargs.update({'crs': dstCrs, 'transform': transform, 'width': width,
'height': height})

iceye_raster = rio.open(iceye_fp, 'w', **kwargs)

# Reproject and save raster band data
for i in range(1, iceye_raster.count +1):
    reproject(
        source=rio.band(iceye_orig, i),
        destination=rio.band(iceye_raster, i),
        src_crs=iceye_raster.crs,
        dstCrs=dstCrs,
        resampling=Resampling.nearest)

# Close destination raster
iceye_raster.close()

print('Progress: Raster reprojection done.')

# Reopening the raster that is now projected to EPSG: 4326
iceye_wgs84 = rio.open(iceye_fp)
```

Vector transformations

```python
# Polygonizing the raster file
mask = None
with rio.Env():
```

```python
    with rio.open(iceye_fp) as src:
        image = src.read(1) # first band
        results = (
        {'properties': {'raster_val': v}, 'geometry': s}
        for i, (s, v)
        in enumerate(
            shapes(image, mask=mask, transform=src.transform)))

geoms = list(results)

# Creating a GeoDataFrame from the polygonized raster
iceye_breaks = gpd.GeoDataFrame.from_features(geoms)

# Get indexes where 'raster_val' column has value 'min'
indexNames = iceye_breaks[iceye_breaks['raster_val'] ==
iceye_breaks['raster_val'].min()].index
# Delete 'min' rows indexes from dataframe as they are Null values in the
raster file
iceye_breaks.drop(indexNames, inplace=True)

print('Progress: Polygonizing done.')

# Rounding up unnecessary decimals
iceye_breaks['raster_val'] = iceye_breaks['raster_val'].round(3)

# Removing all rows with the value of zero
iceye_breaks[iceye_breaks.raster_val != 0.000]

# Classifying the data by flood hazard levels
breaks = mapclassify.UserDefined(iceye_breaks['raster_val'], bins=[0.1,
0.25, 0.5, 1, 2, 2.5])
iceye_breaks['flood_class'] = iceye_breaks[['raster_val']].apply(breaks)

# Dissolving the data into the 6 classified breaks
iceye_breaks = iceye_breaks.dissolve(by='flood_class', as_index=False)
```

Raster clipping and affected population estimation

```python
# Modified from: Lau & Um (2021) Using GeoSpatial Data Analytics: A
Friendly Guide to Folium and Rasterio
# URL: https://omdena.com/blog/geospatial-data-analytics/
#
# ------------------

# Creating an empty list for storing values
results = []

# Looping through flood classes in dataframe and using them as masks for
extracting values
for i in iceye_breaks['flood_class']:

    roi = iceye_breaks[iceye_breaks.flood_class == i]
```

```python
    # Using the mask.mask Rasterio module for specifying the ROI
    gtraster, bound = rio.mask.mask(wsf_raster, roi['geometry'], crop=True)

    # Using values greater than 0 to get population data from the WSF2019-
Imp raster pixels
    results.append(gtraster[0][gtraster[0]>0].sum())

# Saving results in new column
iceye_breaks['population'] = results

# Dividing by 1000 because the WSF2019-Pop data has been multiplied by 1000
to save the file as integer
iceye_breaks['population'] = iceye_breaks['population'].div(1000)
```

Getting OSM data from flooded area

```python
# Splitting the flood areas into > 1m and < 1m flood depth areas
flood = iceye_breaks.from_features(iceye_breaks[0:3])
severe_flood = iceye_breaks.from_features(iceye_breaks[3:7])

# Creating a constant value to use for dissolve
iceye_breaks['dissolve'] = 1
flood['dissolve'] = 1

# Dissolving all the flooded polygons into flooding extents for the
geometries_from_polygon feature
flood_aoi = iceye_breaks.dissolve(by='dissolve')
flood_area = flood.dissolve(by='dissolve')

# Turning the GeoDataFrames into a shapely polygons for OSMnx
osm_aoi = flood_aoi.iloc[0]['geometry']
flood_shape = flood_area.iloc[0]['geometry']
sos_aoi = severe_flood.iloc[0]['geometry']

# Getting OpenStreetMap data from the flooded area
hospital = ox.geometries.geometries_from_polygon(osm_aoi,
tags={'building':'hospital', 'amenity':'hospital'})
pharmacy = ox.geometries.geometries_from_polygon(osm_aoi,
tags={'building':'pharmacy', 'amenity':'pharmacy'})
sos_build = ox.geometries.geometries_from_polygon(sos_aoi,
tags={'building': True})
low_flood = ox.geometries.geometries_from_polygon(flood_shape,
tags={'building': True})
buildings = ox.geometries.geometries_from_polygon(osm_aoi,
tags={'building': True})

# Removing unnecessary columns
sos_build = sos_build[['geometry', 'building', 'name']]
low_flood = low_flood[['geometry', 'building', 'name']]
buildings = buildings[['geometry', 'building', 'name']]

# Removing buildings that are point data
sos_build = sos_build.loc[sos_build.geometry.type=='Polygon']
```

```
low_flood = low_flood.loc[low_flood.geometry.type=='Polygon']
buildings = buildings.loc[buildings.geometry.type=='Polygon']

# Removing the dissolve column
del iceye_breaks['dissolve']
del flood['dissolve']

# Reseting the index of OSM data
low_flood = low_flood.reset_index()
sos_build = sos_build.reset_index()
buildings = buildings.reset_index()

# Removing duplicate buildings from low flood area
low_flood.osmid.isin(sos_build.osmid)
~low_flood.osmid.isin(sos_build.osmid)
low_flood = low_flood[~low_flood.osmid.isin(sos_build.osmid)]

# Removing duplicate buildings
buildings.osmid.isin(sos_build.osmid)
~buildings.osmid.isin(sos_build.osmid)
buildings = buildings[~buildings.osmid.isin(sos_build.osmid)]
buildings.osmid.isin(low_flood.osmid)
~buildings.osmid.isin(low_flood.osmid)
buildings = buildings[~buildings.osmid.isin(low_flood.osmid)]

# Removing unnecessary columns
del low_flood['element_type']
del sos_build['element_type']
del buildings['element_type']
```

RESULTS

```
Reprojecting the data to a CRS that uses a metric system instead of degrees
for the buffering
hospital = hospital.to_crs(epsg=3857)
pharmacy = pharmacy.to_crs(epsg=3857)

# Calculating the centroids of all the points & polygons in the pharmacy
dataset
hospital['geometry'] = hospital['geometry'].centroid
pharmacy['geometry'] = pharmacy['geometry'].centroid

# Reprojecting the data back to a CRS that is better for plotting
hospital = hospital.to_crs(epsg=4326)
pharmacy = pharmacy.to_crs(epsg=4326)

# Getting x and y coordinates of hospitals for mapping
hospital['x'] = hospital['geometry'].apply(lambda geom: geom.x)
hospital['y'] = hospital['geometry'].apply(lambda geom: geom.y)

# Getting x and y coordinates of pharmacies for mapping
pharmacy['x'] = pharmacy['geometry'].apply(lambda geom: geom.x)
pharmacy['y'] = pharmacy['geometry'].apply(lambda geom: geom.y)
```

```python
# Creating a list of coordinate pairs
hospitals = list(zip(hospital['y'], hospital['x']))
pharmacies = list(zip(pharmacy['y'], pharmacy['x']))

# Adding icon columns for kepler gl visualization later on
hospital['icon'] = 'plus-alt'
pharmacy['icon'] = 'pin'

# Setting a coordinate system to the data
iceye_breaks.crs = 'EPSG:4326'

# Function modified from: https://gis.stackexchange.com/questions/429601/
def findtheutm(aGeometry):
    """A function to find a coordinates UTM zone"""
    x, y, parallell, latband = utm.from_latlon(aGeometry.centroid.y,
aGeometry.centroid.x)
    if latband in 'CDEFGHJKLM':
#https://www.lantmateriet.se/contentassets/379fe00e09d74fa68550f4154350b047
/utm-zoner.gif
        ns = 'S'
    else:
        ns = 'N'
    crs = "+proj=utm +zone={0} +{1}".format(parallell, ns)
#https://gis.stackexchange.com/questions/365584/convert-utm-zone-into-epsg-
code
    crs = CRS.from_string(crs)
    _, code = crs.to_authority()
    return int(code)

epsg = findtheutm(iceye_breaks.geometry.iloc[0])

# Calculating the area of flood extents in km2
iceye_breaks['area'] = iceye_breaks.to_crs(epsg).area / 10**6

# Creating a new column
iceye_breaks['flood_depth'] = ''

# Creating labels for the flood depths of the different flood classes
iceye_breaks.loc[0, 'flood_depth'] = '0 ≥ 0.1 meters'
iceye_breaks.loc[1, 'flood_depth'] = '0.1 ≥ 0.25 meters'
iceye_breaks.loc[2, 'flood_depth'] = '0.25 ≥ 0.5 meters'
iceye_breaks.loc[3, 'flood_depth'] = '0.5 ≥ 1 meters'
iceye_breaks.loc[4, 'flood_depth'] = '1 ≥ 2 meters'
iceye_breaks.loc[5, 'flood_depth'] = '2 ≥ 2.5 meters'
iceye_breaks.loc[6, 'flood_depth'] = '2.5 ≥ meters'

# Barplot of people affected by flood
fig, ax = plt.subplots(figsize=(12,10))
sns.barplot(y='population', x='flood_depth', palette='rocket_r',
data=iceye_breaks, ax=ax)
ax.set_title('Amount of people living in the flooded area according to the
WSF2019-Pop data', fontsize=15)
```

```python
ax.bar_label(ax.containers[0])
ax.set_ylabel('Population', fontsize=15)
ax.set_xlabel('Flood depth', fontsize=15)

plt.tight_layout()
plt.savefig('flooded_population.jpg')

# Printing some results
print(len(sos_build), 'buildings in the flooded area are severely inundated
and', len(low_flood) + len(sos_build) + len(buildings), 'buildings have
flooded in total according to existing OpenStreetMap data.')
print('There are',(len(pharmacy)), 'pharmacies within the flood danger
area.')
print('There are',(len(hospital)),'hospitals within the flood danger
area.')
print('The total flooded area is', iceye_breaks['area'].sum().round(2),
'km2.')
```

Kepler.gl

```python
from keplergl import KeplerGl
map_1 = KeplerGl(height=1000)

kepler_flood = iceye_breaks.to_json()
sos_j = sos_build.to_json()
low_j = low_flood.to_json()
build_j = buildings.to_json()
hospitals = hospital.to_json()
pharmacies = pharmacy.to_json()

map_1.add_data(data=hospitals, name='Hospitals in flood area')
map_1.add_data(data=pharmacies, name='Pharmacies in flood area')
map_1.add_data(data=sos_j, name='Severely inundated buildings')
map_1.add_data(data=build_j, name='Inundated buildings')
map_1.add_data(data=low_j, name='Slightly inundated buildings')
map_1.add_data(data=kepler_flood, name='Flood depth')

# Setting a CRS that uses meters
flood_aoi.crs = 'EPSG:3857'

# Getting the lat and lon data of building vectors for automatical zooming
to AOI in interactive map
flood_aoi['lon'] = flood_aoi.centroid.x
flood_aoi['lat'] = flood_aoi.centroid.y

# Getting coordinates for map zoom
lat = flood_aoi['lat'].values[0]
lon = flood_aoi['lon'].values[0]

config = {'version': 'v1',
 'config': {'visState': {'filters': [],
   'layers': [{'id': '6blfkxe',
     'type': 'icon',
```

```
    'config': {'dataId': 'Hospitals in flood area',
     'label': 'Hospitals in flood area',
     'color': [218, 0, 0],
     'highlightColor': [252, 242, 26, 255],
     'columns': {'lat': 'y', 'lng': 'x', 'icon': 'icon', 'altitude':
None},
     'isVisible': True,
     'visConfig': {'radius': 35,
      'fixedRadius': False,
      'opacity': 0.8,
      'colorRange': {'name': 'Global Warming',
       'type': 'sequential',
       'category': 'Uber',
       'colors': ['#5A1846',
        '#900C3F',
        '#C70039',
        '#E3611C',
        '#F1920E',
        '#FFC300']},
      'radiusRange': [0, 50]},
     'hidden': False,
     'textLabel': [{'field': None,
       'color': [255, 255, 255],
       'size': 18,
       'offset': [0, 0],
       'anchor': 'start',
       'alignment': 'center'}]},
    'visualChannels': {'colorField': None,
     'colorScale': 'quantile',
     'sizeField': None,
     'sizeScale': 'linear'}},
   {'id': 'cacpee',
    'type': 'icon',
    'config': {'dataId': 'Pharmacies in flood area',
     'label': 'Pharmacies in flood area',
     'color': [82, 163, 83],
     'highlightColor': [252, 242, 26, 255],
     'columns': {'lat': 'y', 'lng': 'x', 'icon': 'icon', 'altitude':
None},
     'isVisible': True,
     'visConfig': {'radius': 50,
      'fixedRadius': False,
      'opacity': 0.8,
      'colorRange': {'name': 'Global Warming',
       'type': 'sequential',
       'category': 'Uber',
       'colors': ['#5A1846',
        '#900C3F',
        '#C70039',
        '#E3611C',
        '#F1920E',
        '#FFC300']},
       'radiusRange': [0, 50]},
```

```
        'hidden': False,
        'textLabel': [{'field': None,
          'color': [255, 255, 255],
          'size': 18,
          'offset': [0, 0],
          'anchor': 'start',
          'alignment': 'center'}]},
     'visualChannels': {'colorField': None,
      'colorScale': 'quantile',
      'sizeField': None,
      'sizeScale': 'linear'}},
   {'id': 'fnd5s2d',
    'type': 'geojson',
    'config': {'dataId': 'Severely inundated buildings',
     'label': 'Severely inundated buildings',
     'color': [245, 153, 153],
     'highlightColor': [252, 242, 26, 255],
     'columns': {'geojson': '_geojson'},
     'isVisible': True,
     'visConfig': {'opacity': 0.2,
      'strokeOpacity': 0.8,
      'thickness': 0.5,
      'strokeColor': [218, 0, 0],
      'colorRange': {'name': 'Global Warming',
       'type': 'sequential',
       'category': 'Uber',
       'colors': ['#5A1846',
        '#900C3F',
        '#C70039',
        '#E3611C',
        '#F1920E',
        '#FFC300']},
      'strokeColorRange': {'name': 'Global Warming',
       'type': 'sequential',
       'category': 'Uber',
       'colors': ['#5A1846',
        '#900C3F',
        '#C70039',
        '#E3611C',
        '#F1920E',
        '#FFC300']},
      'radius': 10,
      'sizeRange': [0, 10],
      'radiusRange': [0, 50],
      'heightRange': [0, 500],
      'elevationScale': 5,
      'enableElevationZoomFactor': True,
      'stroked': True,
      'filled': True,
      'enable3d': False,
      'wireframe': False},
     'hidden': False,
     'textLabel': [{'field': None,
```

```
            'color': [255, 255, 255],
            'size': 18,
            'offset': [0, 0],
            'anchor': 'start',
            'alignment': 'center'}]},
    'visualChannels': {'colorField': None,
     'colorScale': 'quantile',
     'strokeColorField': None,
     'strokeColorScale': 'quantile',
     'sizeField': None,
     'sizeScale': 'linear',
     'heightField': None,
     'heightScale': 'linear',
     'radiusField': None,
     'radiusScale': 'linear'}},
  {'id': 'fr0h2g',
   'type': 'geojson',
   'config': {'dataId': 'Inundated buildings',
    'label': 'Inundated buildings',
    'color': [255, 152, 51],
    'highlightColor': [252, 242, 26, 255],
    'columns': {'geojson': '_geojson'},
    'isVisible': True,
    'visConfig': {'opacity': 0.2,
     'strokeOpacity': 0.8,
     'thickness': 0.5,
     'strokeColor': [239, 93, 40],
     'colorRange': {'name': 'Global Warming',
      'type': 'sequential',
      'category': 'Uber',
      'colors': ['#5A1846',
       '#900C3F',
       '#C70039',
       '#E3611C',
       '#F1920E',
       '#FFC300']},
     'strokeColorRange': {'name': 'Global Warming',
      'type': 'sequential',
      'category': 'Uber',
      'colors': ['#5A1846',
       '#900C3F',
       '#C70039',
       '#E3611C',
       '#F1920E',
       '#FFC300']},
     'radius': 10,
     'sizeRange': [0, 10],
     'radiusRange': [0, 50],
     'heightRange': [0, 500],
     'elevationScale': 5,
     'enableElevationZoomFactor': True,
     'stroked': True,
     'filled': True,
```

```
        'enable3d': False,
        'wireframe': False},
      'hidden': False,
      'textLabel': [{'field': None,
        'color': [255, 255, 255],
        'size': 18,
        'offset': [0, 0],
        'anchor': 'start',
        'alignment': 'center'}]},
     'visualChannels': {'colorField': None,
      'colorScale': 'quantile',
      'strokeColorField': None,
      'strokeColorScale': 'quantile',
      'sizeField': None,
      'sizeScale': 'linear',
      'heightField': None,
      'heightScale': 'linear',
      'radiusField': None,
      'radiusScale': 'linear'}},
   {'id': 'xcowcw',
    'type': 'geojson',
    'config': {'dataId': 'Slightly inundated buildings',
     'label': 'Slightly inundated buildings',
     'color': [255, 228, 102],
     'highlightColor': [252, 242, 26, 255],
     'columns': {'geojson': '_geojson'},
     'isVisible': True,
     'visConfig': {'opacity': 0.2,
      'strokeOpacity': 0.8,
      'thickness': 0.5,
      'strokeColor': [246, 218, 0],
      'colorRange': {'name': 'Global Warming',
       'type': 'sequential',
       'category': 'Uber',
       'colors': ['#5A1846',
        '#900C3F',
        '#C70039',
        '#E3611C',
        '#F1920E',
        '#FFC300']},
      'strokeColorRange': {'name': 'Global Warming',
       'type': 'sequential',
       'category': 'Uber',
       'colors': ['#5A1846',
        '#900C3F',
        '#C70039',
        '#E3611C',
        '#F1920E',
        '#FFC300']},
      'radius': 10,
      'sizeRange': [0, 10],
      'radiusRange': [0, 50],
      'heightRange': [0, 500],
```

```
        'elevationScale': 5,
        'enableElevationZoomFactor': True,
        'stroked': True,
        'filled': True,
        'enable3d': False,
        'wireframe': False},
      'hidden': False,
      'textLabel': [{'field': None,
        'color': [255, 255, 255],
        'size': 18,
        'offset': [0, 0],
        'anchor': 'start',
        'alignment': 'center'}]},
     'visualChannels': {'colorField': None,
      'colorScale': 'quantile',
      'strokeColorField': None,
      'strokeColorScale': 'quantile',
      'sizeField': None,
      'sizeScale': 'linear',
      'heightField': None,
      'heightScale': 'linear',
      'radiusField': None,
      'radiusScale': 'linear'}},
    {'id': 'e7iibuv',
     'type': 'geojson',
     'config': {'dataId': 'Flood depth',
      'label': 'Flood depth',
      'color': [77, 193, 156],
      'highlightColor': [252, 242, 26, 255],
      'columns': {'geojson': '_geojson'},
      'isVisible': True,
      'visConfig': {'opacity': 0.8,
       'strokeOpacity': 0.8,
       'thickness': 0.5,
       'strokeColor': [119, 110, 87],
       'colorRange': {'name': 'ColorBrewer Blues-7',
        'type': 'singlehue',
        'category': 'ColorBrewer',
        'colors': ['#eff3ff',
         '#c6dbef',
         '#9ecae1',
         '#6baed6',
         '#4292c6',
         '#2171b5',
         '#084594']},
       'strokeColorRange': {'name': 'Global Warming',
        'type': 'sequential',
        'category': 'Uber',
        'colors': ['#5A1846',
         '#900C3F',
         '#C70039',
         '#E3611C',
         '#F1920E',
```

```
              '#FFC300']},
        'radius': 10,
        'sizeRange': [0, 10],
        'radiusRange': [0, 50],
        'heightRange': [0, 500],
        'elevationScale': 5,
        'enableElevationZoomFactor': True,
        'stroked': False,
        'filled': True,
        'enable3d': False,
        'wireframe': False},
      'hidden': False,
      'textLabel': [{'field': None,
        'color': [255, 255, 255],
        'size': 18,
        'offset': [0, 0],
        'anchor': 'start',
        'alignment': 'center'}]},
    'visualChannels': {'colorField': {'name': 'flood_depth',
      'type': 'string'},
      'colorScale': 'ordinal',
      'strokeColorField': None,
      'strokeColorScale': 'quantile',
      'sizeField': None,
      'sizeScale': 'linear',
      'heightField': None,
      'heightScale': 'linear',
      'radiusField': None,
      'radiusScale': 'linear'}}],
  'interactionConfig': {'tooltip': {'fieldsToShow': {'Hospitals in flood
area': [{'name': 'addr:street',
        'format': None},
      {'name': 'name', 'format': None},
      {'name': 'name:en', 'format': None}],
      'Pharmacies in flood area': [{'name': 'addr:street', 'format': None},
      {'name': 'name', 'format': None},
      {'name': 'name:en', 'format': None}],
      'Severely inundated buildings': [{'name': 'name', 'format': None}],
      'Inundated buildings': [{'name': 'name', 'format': None}],
      'Slightly inundated buildings': [{'name': 'name', 'format': None}],
      'Flood depth': [{'name': 'area', 'format': None},
      {'name': 'flood_class', 'format': None},
      {'name': 'flood_depth', 'format': None},
      {'name': 'population', 'format': None}]},
    'compareMode': False,
    'compareType': 'absolute',
    'enabled': True},
   'brush': {'size': 0.5, 'enabled': False},
   'geocoder': {'enabled': True},
   'coordinate': {'enabled': True}},
  'layerBlending': 'normal',
  'splitMaps': [],
  'animationConfig': {'currentTime': None, 'speed': 1}},
```

```python
   'mapState': {'bearing': 0,
    'dragRotate': False,
    'latitude': lat,
    'longitude': lon,
    'pitch': 0,
    'zoom': 10.650982495801145,
    'isSplit': False},
   'mapStyle': {'styleType': 'satellite',
    'topLayerGroups': {},
    'visibleLayerGroups': {'label': True,
     'road': True,
     'border': False,
     'building': True,
     'water': True,
     'land': True,
     '3d building': False},
    'threeDBuildingColor': [9.665468314072013,
     17.18305478057247,
     31.1442867897876],
    'mapStyles': {}}}}

map_1.save_to_html(data={'Hospitals in flood area' : hospitals,
                          'Pharmacies in flood area' : pharmacies,
                          'Severely inundated buildings' : sos_j,
                          'Inundated buildings' : build_j,
                          'Slightly inundated buildings' : low_j,
                          'Flood depth' : kepler_flood},
                   config=config,
                   file_name='Kepler_flood_map.html')
```