Master Thesis

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

Flood risk assessment of heavy rainfall events in data scarce regions

case study of Carrefour, Haiti

by

Andreas Reithofer
U106713

A thesis submitted in partial fulfilment of the requirements of
the degree of Master of Science – MSc

Advisor:
Dr. Korbinian Breinl
Salzburg, 16.04.2023

Science pledge

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

Roitham am Traunfall, 16.04.2023 Mag. Andreas Reithofer

Dedicated to the people suffering from floods
and harsh living conditions in Ayiti, the Flower of High Land,
my brave wife and lovely daughters, who always believed in me,
my dear father in heaven, my dear mother and my dear brother,

and all other hard-working mums and dads,
who have the courage, will and discipline
to complete a postgraduate study program
parallel to daily working business and family business.

Believe in yourself, in your endurance and your power
to reach the top of the via ferrata,
as if you don´t – nobody else will.

Acknowledgments

I would like to acknowledge the Paris Lodron University of Salzburg for the comprehensive Geographic Information Science and Systems Master of Science program. The expertise I have gained in this studies is certainly invaluable.

My sincere thanks to Dr. Korbinian Breinl and Dr. Lorenz Wendt and, who guided me through the present research. Your commitment to supervision was extremely helpful. Also the participation of the MSF GI staff, Edith Rogenhofer and Léo Tremblay, in various web meetings had been a valuable contribution to this work, beginning at the selection of the survey area.

My special thanks to DI Günter Humer and DI Michael Hofer from Humer consulting engineers, where I have worked for 7 years and allowed me to use the “Hydro-As-2D” software as one of the two flood calculation tools used for this work as well as to Dr. Stefano Bagli and Eng. Paolo Mazzoli from GECOsistema srl -Environmental, Climate Risk and Geospatial Intelligence, with whom I had the honour to collaborate together in various projects and who supported me the best way they could in enabling the “pilot area Haiti” on the "SaferPLACES" platform.

Also my special thanks to Prof. Paul Bates from University of Bristol for getting in touch with me for various questions and inquiries on the methodology of LISFLOOD-FP, and your valuable help in getting to possible solutions for the problems I experienced in our AOI Carrefour.

And last but not least I want to thank everybody in my personal environment, especially my wife Ulrike and my daughters Elisabeth, Katharina and Ines for the patience they had with me at countless “holidays” and weekends where I had to sit in front of the computer again instead of being able to spend my free-time with them. Thank you for your endurance and your belief in me and my capability to get this work done!

Table of Contents

Table of Figures

Index of Tables

List of used abbreviations

ARF	Area Reduction Factor
CMORPH	Climate Prediction Center Morphing Technique
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
ERA5	European ReAnalysis Data, 5th generation
FVM	Finite Volumes Method
GHCN-D	Global Historical Climatology Network-Daily
GPEX	Global Precipitation Extremes (dataset by TU Delft, Netherlands)
HDSC	Hydro-meteorological Design Studies Center
HFSA	Hierarchical Filling-&-Spilling Algorithm
IDF	Intensity-Duration-Frequency
LiDAR	Light Detection And Ranging
MSF	Medicins Sans Frontières
NOAA	National Oceanic and Atmospheric Administration
PPFE	Point Precipitation Frequency Estimates
SWE	Shallow-Water-Equations

Abstract

This work deals with the possibilities and limitations for carrying out a flood risk assessment in regions with no or scarce available input data (at the example of the city of Carrefour in Haiti) for for a calculation software and also compares the performance of two very different tools (“Hydro-As-2D” with “Sturzflut” extension for adding pluvial rainfall into the model, hierarchical filling-and-spilling algorithm “Safer-RAIN”) against each other.

The work evaluates a series of different global precipitation datasets (namely CMORPH, MSWEP and ERA5-Land as well as GPEX) originating from remote sensing techniques a) against each other and b) against station data, led by the US National Oceanic and Atmospheric Administration, namely GHCN-D and the point precipitation frequency estimates from the Hydro-meteorological Design Studies Center (HDSC) of the US National Weather Service NOAA.

They are built of the same database like GHNC-D, but with the difference that hourly and sub-hourly durations (until the 5-min interval) are included as well. This matters in the case of this work, as the relevant rainfall duration for the AOI is one hour maximum, presumably below that. So IDF statistics for 3 hours and upwards (like GPEX offers) alone will not do the full job for this area.

While the first three of them are mere precipitation datasets, the GPEX dataset differs from the others mentioned here in form of offering IDF curves, providing information on the intensity, duration and frequency (IDF) of extreme precipitation events. Via an R script, the information in the very big netCDF files could be extracted, evaluated and compared with station data from the neighbouring island of Puerto Rico and the nearest station in the Dominican Republic. This served also as validation for assigning a certain rainfall sum R a certain return period T.

After evaluating these possible data sources for rainfall against each other resp. against the station data, a pluvial flood calculation has been carried out in both toolsets and for three different rainfall scenarios (50mm, 90mm and 130 mm, each in 1h). Regarding the flood extent and the average water depth, differences in the results are very apparent and are due to the different methodologies used in the toolsets.

As an outlook, a proposal is made on how these tools could help local Government authorities and citizens to be prepared for future heavy rainfall events by means of the calculation of several rainfall intensities and their specific impact on the urban area (delineation of the most flood-prone an areas, the expected water depth, conclusion on safe areas) and maximise the reusability of the collected data.

1. Introduction

According to the Washington based World Resources Institute, in a global view, in the last 40 years, flood events have caused over $1 trillion in economic losses and more than 220,000 people lost their lives[1]. Global flood damages have increased sharply in this period, mainly driven by growth of population and economic activities in flood-prone areas (Winsemius et al., 2016). Alone in the year 2021, the amount of global economic loss caused by flooding reached 105 billion U.S. dollars, according to a report of Hamburg-based researcher Madhumitha Jaganmohan[2]. In the European Union, according to data of the European Environmental Agency EEA, the total reported economic losses caused by weather and climate-related extremes between 1980 and 2020 amount to approximately EUR 487 billion (at 2020 values) accounting for about 80% of the monetary losses over that period.[3]

Haiti, rated as the poorest country in the Western hemisphere, is especially vulnerable to a series of natural disasters, namely earthquakes, draught and floods, the latter due to Hurricanes crossing the island as well as ordinary heavy precipitation events. In 2016, from Oct 4th until Oct 6th, cat-5 hurricane Matthew has struck Haiti hard, especially the Tiburon peninsula in the south-west of the country. About 2.1 million people had been affected with an official death toll of 546 citizens, 200.000 buildings being destroyed and an estimated monetary damage of 2.8 billion US$[4].

But not only hurricanes strike the island of Hispaniola, together with the exceptionally high relief energy (in most areas, terrain rises very steeply from the coastline to the back country) and the excessive deforestation in Haiti, heavy precipitation events due to tropical thunderstorms, which can occur in the “rainy seasons” from April to June and September to November, are a severe threat to the often informal settlements, too. Primary forest, which is defined by the FAO as “*naturally regenerated forest of native species, where there are no clearly visible indications of human activities and the ecological processes are not significantly disturbed*”, has shrunk in Haiti from already low 4.4% in 1988 to almost zero (0.32%) in the year 2016, (Hedges et al., 2018)

Main reasons for deforestation in Haiti are the enormous pressure on land-use for self-sufficient farming in combination with the use of charcoal as primary energy source. The nearly complete loss of primary forest is fatal during heavy precipitation events, where the small, ephemeral creeks (“Ravines”) quickly turn into dangerous torrents, carrying significant amounts of sediment due to erosive processes in the watersheds (Heimhuber et al., 2015).

But the frequent floods are – together with all other numerous problems of natural, political and economic origin – a heavy burden and constant threat to the residents living in flood prone areas on often insufficient building fabric. Together with the lacking financial resources of the Haitian Government as well as the municipalities to protect the buildings in an adequate way, it is essential to at least identify the most flood prone spots, to enable the governmental bodies as well as the

1 Press release from April 23rd 2020, https://www.wri.org/news/release-new-data-shows-millions-people-trillions-property-risk-flooding-infrastructure (access on 2022-07-23)
2 https://www.statista.com/statistics/510894/natural-disasters-globally-and-economic-losses/ (access on 2022-08-10)
3 https://www.eea.europa.eu/ims/economic-losses-from-climate-related (access on 2022-08-10)
4 https://www.nhc.noaa.gov/data/tcr/AL142016_Matthew.pdf, https://reliefweb.int/report/haiti/haiti-iom-appeal-haiti-hurricane-matthew-9-october-2016 (access on 2022-08-11)

rescue services and Non-Governmental aid organisations like Medicins Sans Frontières (MSF), who run a hospital in the city of Carrefour (situated near the capital Port-au-Prince) which had already been endangered by flooding in the past.

This is also the personal background of this work – Dr. Lorenz Wendt from the GeoHUM research laboratory in Salzburg arranged the contact with MSF for me, and in regular discussions, the area of interest for this thesis work had been defined. The focus lies not only on the flood calculation (under tightened conditions, which means dealing with scarce input data) itself, but also on the comparison of a merely GIS based flood routing algorithm called “Safer_RAIN” and a conventional, commercial 2d hydraulic simulation software (“Hydro-As-2d”).

Figure 1: Flood damage in Haiti, August 2022. Photo: Haiti Civil Protection (https://floodlist.com/america/haiti-floods-august-2022, access on 2022-08-11)

1.1. The specific situation in Haiti

1.1.1 Data Scarcity

I still remember very well a term my Columbian professor told us to memorize when I was a young student at the University of Economics in Vienna: “La escasez de los recursos”, the scarcity of resources, human beings almost always have to deal with. In the case of Haiti, input data for accurate and validated flood risk calculation is scarce. There are not many gauges for discharge and precipitation, and the ones that do exist are to the time being not publicly available. High-resolution Digital Elevation Models out of airborne LIDAR data or drone flights in most areas simply do not exist.

Nevertheless, accurate high-resolution terrain models with a pixel width of 2m or below are really crucial for an accurate delineation of flood prone areas both for GIS-based algorithms like the hierarchical filling-and-spilling approach (“Safer_RAIN”, (Samela et al., 2020) used in this work as well as for hydrodynamic flood calculation software using the shallow water equations like Hydro-As-2d, Flo-2d, Jflow, Mike, Lisflood-FP and others (Dimitriadis et al., 2016).

For being able to compare two different computational approaches, 2d hydrodynamic with GIS data only as input and the merely GIS-based approach “Safer_RAIN”, I was able to use LIDAR point cloud data of Carrefour itself plus the surrounding areas near the coast. This is due to data gained for a humanitarian research mission carried out by the world bank some years ago. The data is available via the platform opentopography.org, an open data portal for LIDAR data around the globe, point cloud data as well as Digital Terrain Models (DTMs) and Digital Surface Models (DSMs).

The data has been made publicly available only last year, before, there had been nothing more than SRTM or ASTER data, which is comparably poor in accuracy as well as resolution, especially for flood calculation (i.e. ASTER data is surface data, so the tree cover is part of the digital elevation model. This is of course suboptimal for flood simulation).

Data availability is significantly poorer when it comes to precipitation data, discharge data and / or high-resolution satellite images for validation of flood extent, if it comes to real life rain events which happened in the past. Data of in-situ meteorological stations could not be obtained for this work, even though I had been in contact with MSF in the form of regular discussions and progress reports at least every 2 months for almost the whole year 2021. The GI department of MSF enquired data from the Haitian government officials, but it was not possible to obtain any station data. For this reason, the only solution was to use one of the global precipitation data models, where data is gained via remote sensing (mostly radar data) and compare it with available station data “nearby” (which means either the Dominican Republic or Puerto Rico).

High-resolution satellite images, soil data and land cover as well as data on building outlines (for setting up the hydraulic 2d model as well as the GIS-based calculation for Safer_RAIN) are hard to acquire. Open Street Map data, which is an excellent data source for traffic routes and buildings at least in Europe and North America, works well for the road network in Haiti, but not for buildings.

Buildings are eventually available via Open Street Map, although the quality of data is very inconsistent in these areas – there are districts with good quality data, and in other, hardly any buildings have been digitized. This is especially the case in the Bidonvilles, roughly translatable

with slums, which often originate out of spontaneous disaster refugee camps (Hannemann, 2014). But the integration of buildings is essential for accurate calculations of flood routing and -extent, regardless of which method (GIS vs. hydrodynamic) we use.

1.1.2 Deforestation & pressure for land use

Deforestation is long since a big issue in Haiti. To understand the reasons for this, it is necessary to undertake a small crash course into Haitian history. Christopher Columbus founded his first settlement, “La Navidad”, in the year 1492 on the North coast of today´s Haiti, then inhabited by the indigenous Taíno peoples.

By the early 17th century, French colonists started planting sugar cane, coffee and other crops in a large scale, which made the island extremely profitable for the white upper class as well as for Mainland France. In the 1780s, the French colony of “Saint Domingue” produced about half of all the sugar and coffee consumed in Europe and the Americas (Ferrer, 2012) and became central to France’s economic strength, yielding about 500 million livres in revenues a year, with more than 400.000 slaves imported from West Africa (Cantir, 2017). “*On the eve of the revolution, the French colony produced more than the Spanish and British colonies combined, with its 793 sugar, 3117 coffee, 789 cotton, 3150 indigo, 54 cacao, 182 guildiveries, and 370 fours a chaux plantations and more than 1,500 ships that transported goods to Europe”* (Obregón 2018: p 601). For all the plantations, significant amounts of forest areas had to be cleared.

After the first (and only) successful slave revolution in the world taking place in Haiti 1791, black leader Jean-Jacques Dessalines declared sovereignty from France in 1804. After 21 years of political unrest and 10 years of unsuccessful negotiations, King Charles X of France set “the French part of Santo Domingo” (Today´s Haiti and the Dominican Republic had been united in one republic at that time) a last ultimatum with extremely harsh economic conditions and the physical threat of 14 army ships docked in Port-au-Prince in exchange for a formal recognition of political independence (Obregón 2018).

Haitian President Boyer could only chose between war and economic surrender, with the French crown demanding the enormous sum of 150 million francs indemnity for the lost plantations payable in yearly rates. The indemnity amounted to five times France´s total annual budget and ten times as much as the United States paid Napoleon for purchasing the state of Louisiana (F50 million, (Obregón, 2018).

This sum was effectively unpayable and severely hampered economic development (until 80% of Haiti's government spending had been debt repayment, as money had also been borrowed from French and later US Banks at high interest rates, called “Haiti´s double-debt”) and, most important in this context, also led to further deforestation. Lumber was one of the most valuable income sources, and any source of revenues was desperately needed for debt repayment.

Other factors for the vast deforestation in Haiti are the massive use of charcoal as the predominant source of energy for cooking and – if needed – heating and the land use pressure. According to a world bank study from 2018, Charcoal consumption in the capital Port-au-Prince with an estimated

population of 2.6 million inhabitants for the urban agglomeration in the year 2018[5], sums up for almost 440,000 tons per year, accounting for almost half of the entire consumption in Haiti which is about 950.000 tons per year (Tarter et al. 2018).

Population in Haiti has tripled between the years 1960 and 2021[6], and the vast majority of the rural population still relies on subsistence farming. But in contrast to population growth, land resources are finite and fertile arable land is scarce – so again, forest areas had been cut down as a result.

Today, estimations for forest cover in Haiti vary quite a lot, between 3% for the year 2006 (Dolisca et al., 2007) and 21% for the year 2015 (Pauleus and Aide, 2020), depending not only on the definition of the word “forest” (tree height, percentage of canopy) but also which type of tree cover is considered to count as forest (primary forests only versus inclusion of secondary forests and shrubland). Table 1 gives a good overview on different studies on forest cover and their underlying forest definition, carried out in the timespan from 1972-2015.

One can easily see that the definition of the term forest is not at all uniform and therefore has a huge impact on the estimated percentage of forest cover. Also the source of information (usually freely available satellite images with LANDSAT as pioneer and therefore leading source of information) and its spatial resolution and temporal availability has influence on the estimations.

But generally, it can be stated that one can see the difference in tree cover density between Haiti and the neighbouring Dominican Republic in most areas at one glance in Google Earth (see Figure 3 on page 18). This fact is also reflected by world bank statistics, where a reported forest cover of 44.4% in the Dominican Republic in the year 2020 is juxtaposed by a forest cover of 12.6% in Haiti. Following the world bank statistics, forest area in Haiti has even shrunk from already low 13.9% in the year 1990, whereas it has significantly grown from 33% in the Dominican Republic in the year 1990[7] (see Figure 2).

And alone from that, one can guess that surface flow originating from heavy rainfall events does pose a problem, as terrain is generally very steep and mountainous in the whole Island of Hispaniola. Together with rather unfavourable geologic conditions (mostly limestone, which leads to shallow soil horizons vulnerable to erosion as soon as the protective vegetation cover is gone), the ongoing process of deforestation is a big issue in Haiti.

5 Source: https://reliefweb.int/report/world/worlds-cities-2018 (access on 2022-07-07)
6 Source: https://data.worldbank.org/indicator/SP.POP.TOTL?end=2021&locations=HT-DO&start=1960&view=chart (access on 2022-07-07)
7 Source: https://data.worldbank.org/indicator/AG.LND.FRST.ZS?locations=HT-DO (access on 2022-07-07)

Sources	Location	Forest cover (%)	Methods	Agency	Date	Forest definition
Lanly (1982)	Tropics (include Haiti)	1.73	Landsat 1 & 2	FAO	1972 & 1978	Broadleaved woody vegetation with more than 10% canopy cover.
Ortiz-Chour, Davis & Pugliese (2000)	Tropics (include Haiti)	3.2	Compilation of National Statistics.	FAO	2000	Woody vegetation with >10% crown cover including plantation; excludes trees for agricultural production.
Food and Agriculture Organization of the United Nations (2010)	Global Forest	4	National forest inventories compiled by the FAO	FAO	2010	Woody vegetation with >10% of tree crown cover >5-m tall. Closed forests with tree crown coverage >60%.
Álvarez-Berríos et al. (2013)	Great Antilles	8	MODIS Vegetation Index image; random forest classifier	NA	2000 & 2010	Trees and shrubs with >80% cover.
Hansen et al. (2013)	Global Forest	16.17	Wall to wall 30 m resolution Landsat images.	University of Maryland	2013	>25% canopy (2013), >5 m height, tree cover.
Churches et al. (2014)	Haiti	29.4 & 32.4	Modified normalization method using pseudo-invariant polygons, Landsat 5 image classification.	NA	2010	Vegetation greater than 5 m in height with a canopy cover of ≥10%. Includes mangroves. Exclude fruit plantations.
Hedges et al. (2018)	Haiti	0.32	Landsat 5, 7 & 8	N /A	1984–2016	Old growth with closed forest cover canopy (≥70% tree cover) with little to no human intervention.
This study	Haiti	26.4 & 21.3	Wall to wall 30 m resolution Landsat images using GEE	NA	2000 & 2015	Dominant tree cover class with moderate canopy (≥10%; 80% or more cells in the training polygons are trees).

Table 1: Different studies on forest cover estimations and forest definition (Pauleus and Aide, 2020, p. 15)

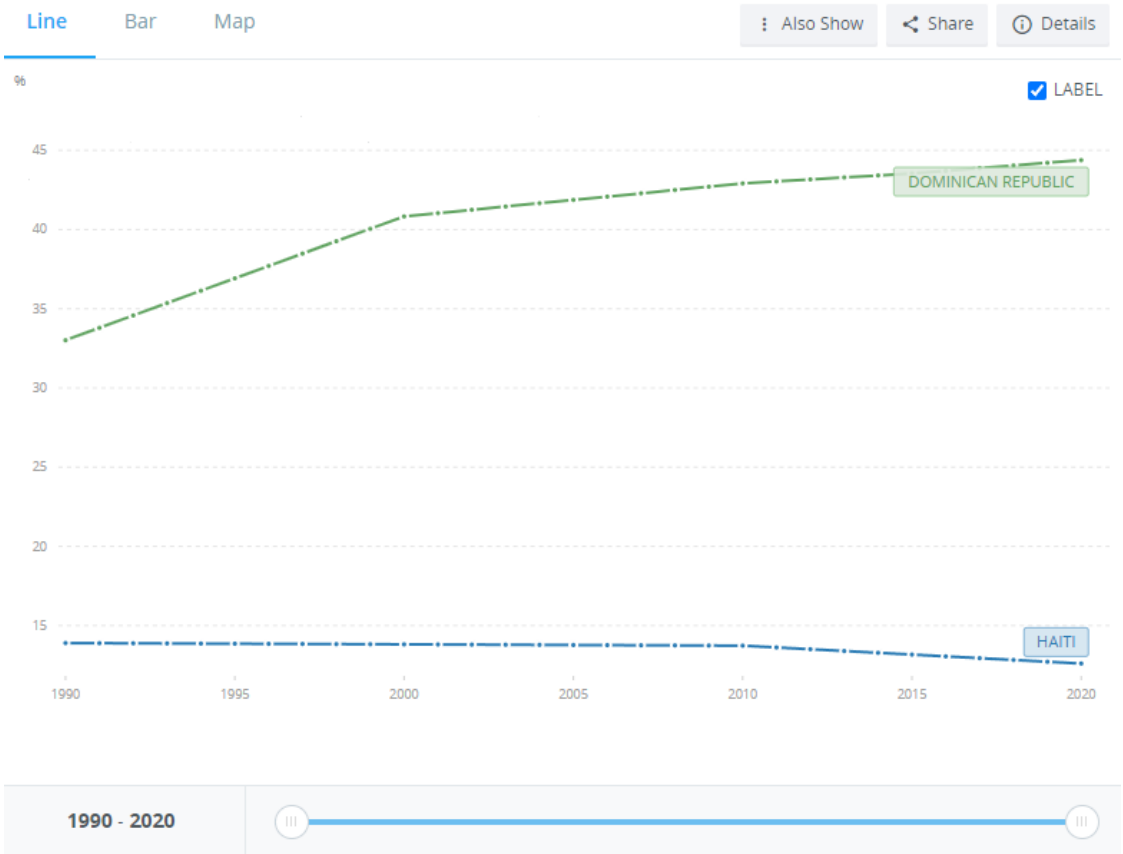


Figure 2: Forest area (% of land area) in Haiti, Dominican Republic (World bank, https://data.worldbank.org/indicator/AG.LND.FRST.ZS?locations=HT-DO, access on 2022-07-07)

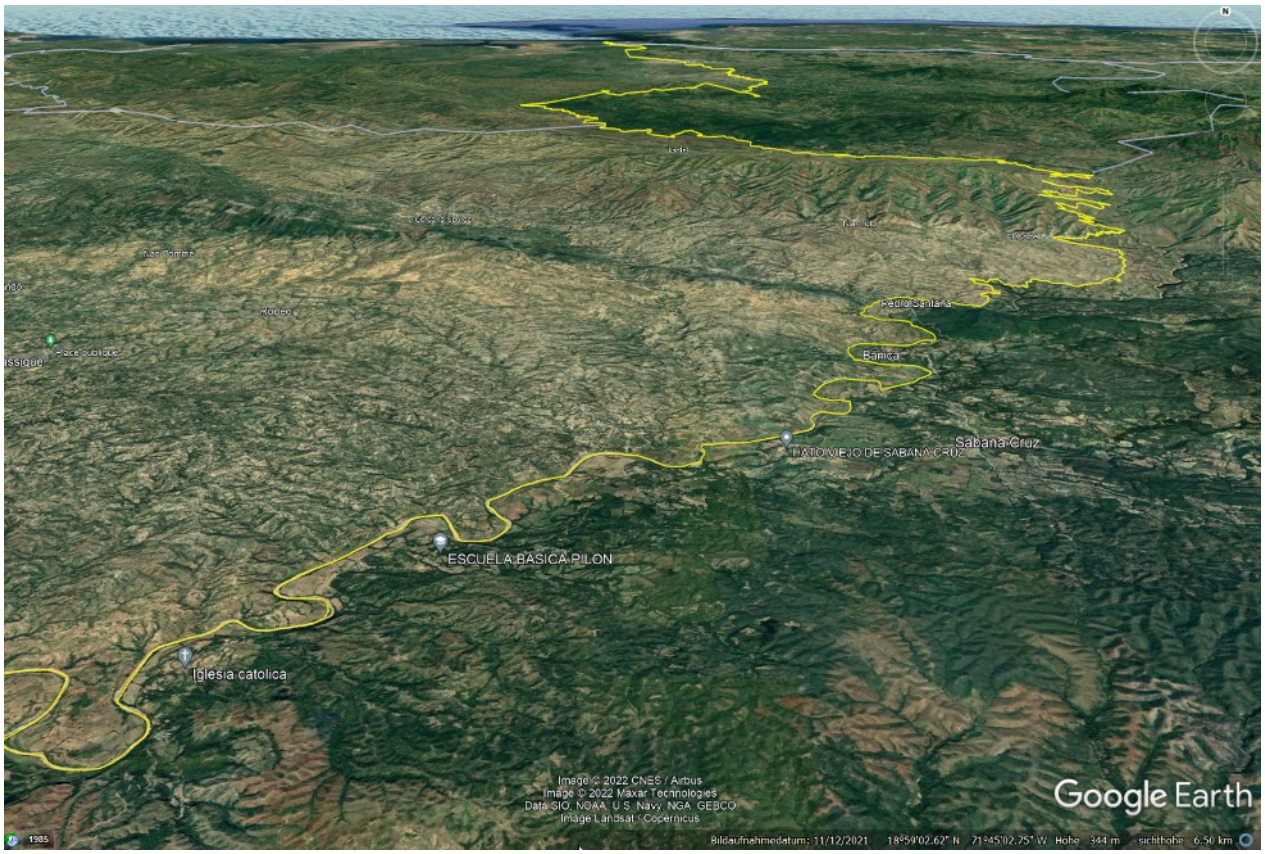


Figure 3: Google Earth Satellite Images in the border region between Haiti and the Dominican Republic (image centre at ~ 19°04’ N / 71°44' W, left: Haiti, right: Dominican Republic, access on 2022-09-15)

1.1.3 Hurricanes & heavy rainfall events

Flood hazards and flood damage in Haiti have mainly two origins: First, the hurricane season, and second, the rainy season which is for the most areas in fact twofold – the first one occurring from March to May and the second one occurring from August to October (see(Klar et al., 2014)Figure 4). While hurricanes are large-scale and slowly moving and thus mostly affecting a big area, the two rainy seasons account for the occurring small-scale heavy rainfall events.

Being located in the centre of the North Atlantic hurricane belt, Haiti is frequently exposed to strong tropical storms and hurricanes. During the cyclonic season, Haiti is primarily subject to tropical waves and local and regional disturbances (fairly organized storms), which are sometimes influenced by the Inter-tropical Convergence Zone (ITC) and carried from east to west by trade winds. These low pressure systems can lead to tropical cyclogenesis and evoke cyclonic systems producing abundant rainfall and strong winds, which can be qualified in terms of their intensity and structure (Mora et al. 2010).

In addition to these synoptic-scale disturbances, Haiti can also be affected by orographic-convective systems (storms) at regional or local level, strengthened by local effects like the mountainous landscape. This often leads to very severe flooding and landslides and / or debris flow caused by spatially limited, but nevertheless heavy and intense rainfalls in a limited area (Mora et al., 2010).

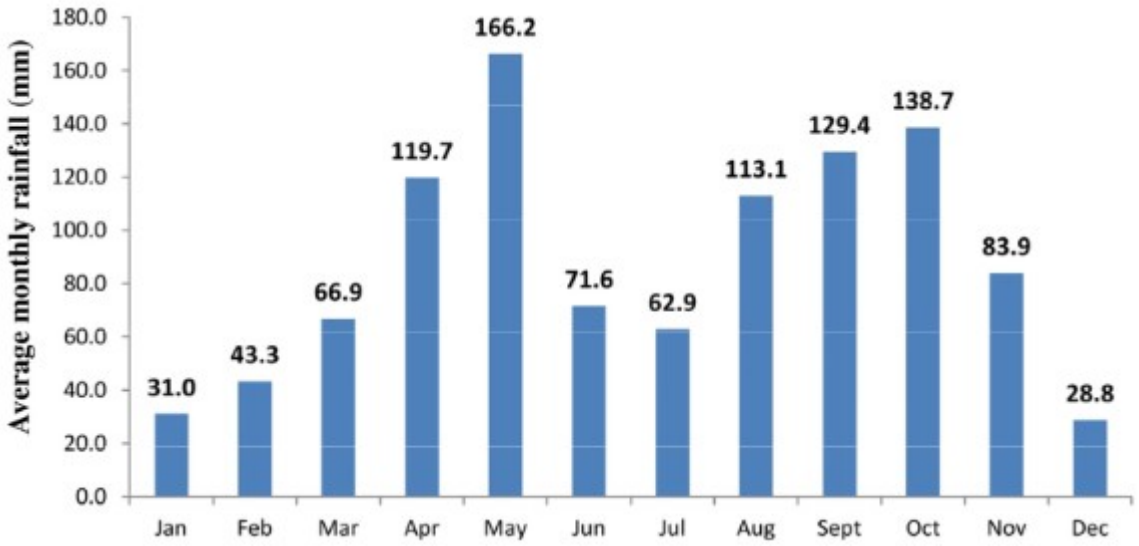


Figure 4: Average monthly rainfall at Port-au-Prince (station Damien, n= 68, 1927 until 2002, Heimhuber et al., 2015, p. 3839)

In recent decades, Haiti has also been deeply affected by global climate change causing longer dry seasons, periods of extended drought, erratic seasonal rains and increased frequency and severity of tropical storms (Smucker et al., 2017). Haiti has a long row of hurricanes crossing the country, as well as heavy precipitation events in the two rainy seasons, which become more and more precarious due to the low percentage of forests, high soil erosion and generally poor soil conditions, a high percentage of imperviousness in urban areas (very high building density) and the steep relief. All these factors together account for an extremely high runoff coefficient in most parts of the country, which makes high precipitation amounts even worse.

The widespread absence of bigger rivers (with the exception of the Artibonite river), together with the factors mentioned above make Haiti extremely vulnerable to pluvial floods, which fill the otherwise dry *ravines* (gullies) with water very quickly and also lead to extensive spontaneous overland flow. As a result, we have recurring reports on pluvial flood damage causing casualties and damage to the usually bad building substance (see Figure 5).

Figure 5: Report on a flash flood event in Ouest Departement of Haiti (https://floodlist.com/america/haiti-flash-floods-ouest-department-june-2019 , access on 2022-07-15)

Table 2 shows some selected hurricane and / or heavy rain events in Haiti and their consequences:

Rain events Hispaniola and surrounding islands									
date	place	type	Dur h	mm total	peak discharge m³/s	Persons dead	Persons injured	flooded buildings	destroyed buildings
2021-04-05	North West Haiti	Heavy rain	96			3		1270	149
2020-08-22	South-Eastern Dom. Rep.	Tropical storm "Laura"	24	200		11		1000	
2019-09-18	Quest Dep.	Heavy rain				2		250	30
2019-07-02	Quest Dep.	Heavy rain							
2019-06-08	Carrefour	Heavy rain				2		500	17
2017-11-16	Sud, Nord-Ouest Haiti	Heavy rain				5		10000	
2017-09-21	North Dom. Rep.	Hurricane "Maria"	48	500				3700	110
2017-09-20	Guadeloupe	Hurricane "Maria"	24	246					
2017-09-06	Haiti	Hurricane "Irma"							
2017-05-16	North Dom. Rep.	Heavy rain	24	147				655	
2017-04-24	Les Cayes / Sud Haiti	Heavy rain				2			80
2016-11-05	Nord Haiti	Heavy rain	24	105		10			
2016-05-08	Eastern Dom. Rep.	Heavy rain / thunderstorms	24	300					
2008-09-06	Gonaive	Hurricane "Ike"	24	216	315				
2008-09-01	Gonaive	Hurricane "Hanna"				529			
2008-08-26		Hurricane "Gustave"				75			

Table 2: Selected hurricane and / or heavy rain events in Haiti (data source: https://floodlist.com, manual evaluation of reported flood incidents concerning Haiti, access on 2022-07-15)

1.1.4 Steep Slopes & soil erosion

Before the arrival of Columbus, the people of the Taíno called their land "Ayiti", meaning "Flower of high land", which can also be translated as "Mountainous land"[8]. Indeed, Hispaniola as a whole and Haiti in particular is a very mountainous island with steep slopes, high mountains and deep gorges and valleys, for that reason, soil erosion originating from loss of forest areas counts double – as well as the steep terrain intensifies erosion on its own. In most areas of Haiti, terrain rises steep and quickly as one goes inland coming from the sea and this is also the case in the area of interest of this work, the city of Carrefour. At the same time, the widespread absence of sewage systems and waste collection impose big hygienic problems in case of flooding due to heavy rainfall as well.

As already pointed out in chapter 1.1.2, Haiti has almost completely lost its primary forest cover. Though there exist secondary forests and shrubland, excessive tree harvesting for charcoal and the widespread agricultural use of steep slopes have radically transformed the natural landscape in Haitian watersheds. These two factors radically decreased the infiltration capacity and exposed the population to both acute erosion upstream and severe flooding downstream.

Another hazard caused by deforestation together with slope gradients of over 20° (which occur very often in Haiti) are landslides and debris flow / debris avalanches. Forest canopies serve as natural buffers against wind and rain and the roots of trees help keeping the granular soil from shifting. Deforestation has led to landslides becoming a major concern, especially during the rainy season. Figure 6 shows the correlation between slope angle and occurred landslides – the number of pixels with landslides starts to rise significantly at slope angle of ~15° (26%) and reaches its top at around 35° (70%). Many hill slopes in Haiti fall into this category.

8 https://haitihub.wordpress.com/2015/10/08/how-haiti-got-its-name-a-reminder-this-columbus-day/

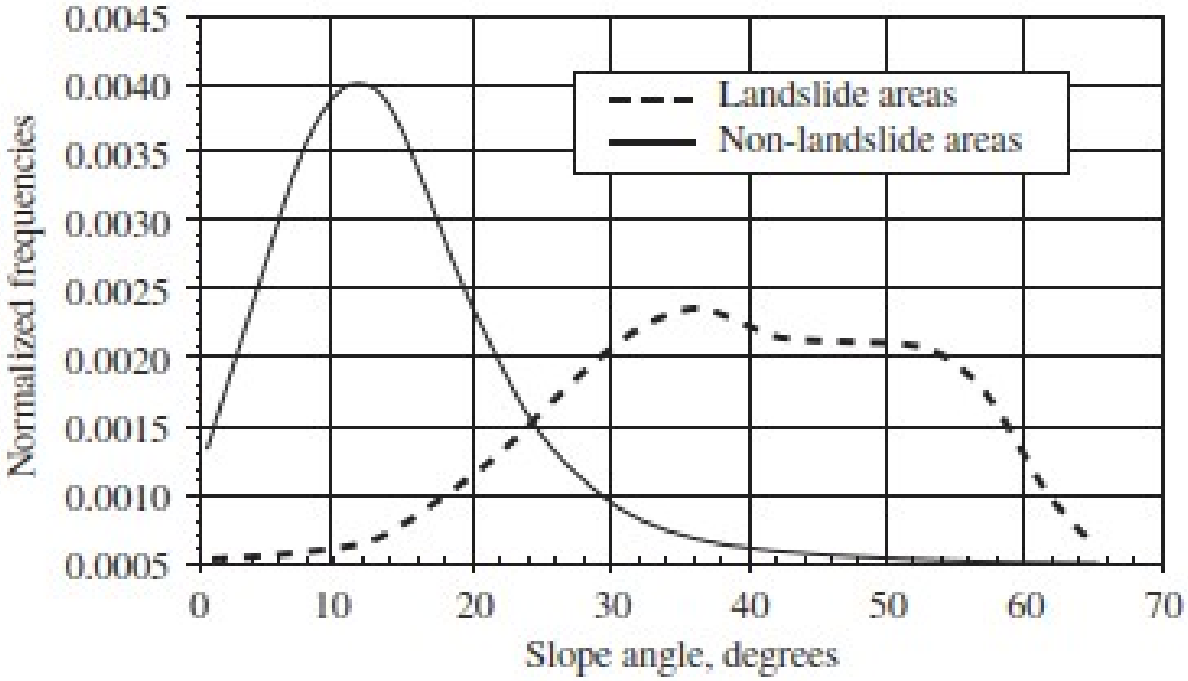


Figure 6: Empirical distribution functions of slope gradient in pixels from landslide area vs. non-landslide area (Davis, Chung and Ohlmacher 2006, p. 1123)

1.1.5 Humanitarian situation & context

Since the enormous reparation payments Haiti had been urged into by the former colonial power France in the 19th century, Haiti has always been a poor country in economic context. This seems somewhat paradox, as the French colony of Saint Domingue was called “pearl of the Antilles” before – but this was valid for the white upper class of French land owners only. In fact, Haiti´s natural resources of rich forest vegetation and fertile soil had been ruthlessly exploited and the enormous economic wealth of a minority as well as the French crown had been gained on the back of millions of enslaved women and men deported from West Africa for forced labour on the sugarcane and other plantations.

In July 2021, the elected Haitian president Jovenel Moïse has been assassinated and since that, the state of the in many aspects already dysfunctional political system has become even worse. Gang crime[9], corruption, an underdeveloped and heavily indebted economy[10], widespread poverty and an analphabetism rate of over 30% of the total population[11], sparse agricultural resources in combination with a rapidly growing population and high vulnerability to natural disasters, namely hurricanes and earthquakes, sound like the “perfect storm” of challenges for a country.

The multidimensional challenges for Haiti also extends to the lack of spatial planning, combined with uncontrolled and rapidly increasing urban expansion, often in shanty towns (“Bidonvilles”) in

9 Source: https://www.derstandard.at/story/2000127786881/bandenkrieg-bringt-haiti-an-den-rand-der-anarchie (Newspaper article in Austrian newspaper “Der Standard”, access on 2022-07-28)
10 Source: https://admin.lenouvelliste.com/images/kiosk/2022-07-23/1658553495_7.jpg (access on 2022-07-28)
11 Source: Haiti-Education: Cri d’alarme face au taux d’analphabétisme, https://www.alterpresse.org/spip.php?article20589 (access on 2022-07-28)

the surroundings of the capital city of Port-au-Prince. By 2017, approximately 56% of Haiti's total population of then 10.3 million people (2021: 11.5 million[12])lived in the capital city of Port-au-Prince (Posner and Georgakakos 2017).

Often, the unorganised informal settlements emerge out of natural disasters like the catastrophic earthquake of 2010, where according to official estimates 316,000 people lost their lives, over one million were left homeless and whole cities have been levelled[13]. It is obvious that these informal settlements are especially vulnerable to natural disasters like hurricanes and floods, as the building substance is mostly very bad and the localities of the Bidonvilles are usually unfavourable, either on steep slopes or near the coast or ravines, making them especially vulnerable to landslides and floods.

Medicins Sans Frontières, shortly MSF, run several hospitals in the Port-au-Prince, Carrefour and Leogane area (as well as in other cities), some of them regularly in danger of being flooded. This work should be a contribution for raising awareness and estimating the potential damage and endangered lives in case of a severe flood. Additionally, as central research question, it should be examined, whether cost-efficient and fast online tools for assessing (here: pluvial) flood risk could be an alternative to cost- and time intensive hydraulic 2d calculations.

12 Source: https://data.worldbank.org/indicator/SP.POP.TOTL?end=2021&locations=HT&start=1960&view=chart (access on 2022-07-28)

13 Source: https://www.britannica.com/event/2010-Haiti-earthquake (access on 2022-07-28)

1.2. Aims and objectives

1.2.1 Methodological approach

How can one get reliable flood delineation results in an economically deprived region like Haiti? Where input data is scarce in general and data sources for the calibration and, even more important, the validation, of the calculated results are even more scarce or they are simply not available?

This is indeed a methodological problem, where much time and brain work has run into this academic thesis. *How can I assess the validity of my results, if I cannot get any in-situ data from rain gauges and / or discharge volumes?*

In my 7-year engineering practice in the consulting engineer office IBH, pluvial as well as fluvial flood calculations have been part of my daily business. But, unlike in Haiti, for model calibration, we used discharge curves, rain gauges and the national dataset of design precipitation[14], edited by the Austrian federal government.

Basically, the workflow for obtaining accurate and methodologically valid results of flood extents (and water depth) for different return periods respectively different rainfall intensities is described in Figure 7 as well as in detail in the following main chapters 2(Study area and data) and 3 (Methods and concepts).

The **main input** for both modelling approaches (Hydro-As-2D with “Sturzflut” extension as well as the GIS-based hierarchical filling-and-spilling algorithm “Safer_RAIN”) for this work is

- **high-resolution digital terrain model** (2m pixel size, originating from LIDAR data, in this case it was built directly out of the original point cloud)
- **land cover data** (gained out of OpenStreetMap and satellite images)
- **building outlines** (gained directly out of the LIDAR point cloud, as OSM data is very inconsistent here)
- **rainfall data** (as no in-situ data sources have been available, derived from the global datasets MSWEP & ERA5-Land and NOAA station / IDF data for Puerto Rico)

Roughness coefficient (important parameter for flow velocity) has been set according to land cover, while **infiltration** rate has been set to zero to account for the very low potential for infiltration in case of heavy precipitation events in combination with a high gradient, poor soil conditions, a very low tree cover and a high degree of imperviousness in the urban areas.

Concerning the simulated **precipitation**, it is good practice to calculate at least three different scenarios based on different return periods. In this case I have calculated three different rainfall scenarios with 50, 90 and 130mm in one hour. But how can I assess a certain return period, if neither in-situ rainfall gauge data nor design precipitation data are available for the area of interest?

14 https://ehyd.gv.at/

For **model validation** (see Chapter 3.2)**,** my thesis advisor Korbinian Breinl first suggested to use remote sensing (radar) datasets, like the **Copernicus ERA5-Land** climate reanalysis dataset[15]. Although this is a well proven and globally available dataset, it has shown (see chapter 2.3.6 and3.2.1) that in the case of the station Maria Montez International Airport, Dominican Republic (the nearest rain gauge to Haiti), the rainfall sum for 24h for Hurricane Matthew was about 2-3 times lower than the measured (point) precipitation, so even regarding an area reduction factor (ARF, see chapter 3.2.2) of ~0.9 for the grid cell size of ERA5-Land and the duration of 24h, this does not explain the huge delta between the two datasets.

To compare data of more than one station alone (which might lead to a chance finding), **NOAA point precipitation frequency estimates (PPFE)** for the orographically similar region around the city of Ponce in the neighbouring island of Puerto Rico[16] was collected, evaluated and compared against the remote sensing datasets MSWEP and ERA5-Land. For this reason, COOP hourly precipitation data[17] of Ponce (South coast of Puerto Rico) has been collected, compared to the time series (annual maxima) of the PPFE data and after a successful matching between COOP and PPFE data, juxtaposed with the remote sensing datasets MSWEP and ERA5-Land (see Chapter 3.2.1 and 3.2.3 for more detailed information on this).

Model calibration (see Chapter 3.3) could be done either via comparing the discharge curve (or at least the peak value on river X of a certain rain event) of a gauge somewhere at the bottom of the catchment or a smaller subcatchment to the discharge curve in the 2d model or via comparing the flood extent to aerial or satellite images; in case of absence of both also to (geotagged) ground photographs. Unfortunately, neither suitable satellite images nor geotagged ground photographs have been available either for this work, despite intensive research from my side and from side of MSF. With other words: model calibration like it is usually done could not be carried out in this work.

Satellite images have to be very high in resolution to be useful for this aspect; unfortunately, we have to deal with the problem that pluvial floods occur generally very quickly and do not last long (unlike fluvial floods in big rivers, for example). The steep terrain with hills and mountains rising directly behind the coast even accelerates the flow velocity, so the chance to get a photo from an event – with little or no clouds – is indeed very small. Although I found a high resolution image of a flood event in Carrefour from the year 2019, the image had a cloud cover of 90% and thus was not suitable for any flood extent analysis.

Nevertheless, despite the experienced problems in calibrating both flood models (the hydrodynamic 2d model as well as the GIS-based Safer_RAIN model), it was possible to calculate a reasonable and accurate flood extent for certain amounts of precipitation, at least with the Hydro-As-2d model. The filling-and-spilling algorithm, on which Safer_RAIN is based, has methodological problems with our area of interest, which I was not aware of when the decision from MSF fell on the city of Carrefour – more on this in Chapter 4.3.

15 https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview (access on 2023-03-04)
16 https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_pr.html (access on 2022-11-15)
17 https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00988 (access on 2022-11-15)

The **two last major steps** in the flowchart consist of

- the **GIS intersection** of (available) OSM layers of critical infrastructure (schools, hospitals) with the calculated flood extent and the maximum water depth for the final flood risk assessment and
- the **comparison** of the results (water depth, flood extent) of the hydrodynamic commercial software Hydro-As-2d and the GIS-based filling-and-spilling algorithm “Safer_RAIN” as well as the elaboration of each models strengths and weaknesses and its suitability for pluvial flood calculations in data scarce regions in general and the chosen area of interest in particular.

To sum it up, though the terrain model was pretty good in resolution, I was experiencing the phenomenon of “data scarcity” in many other aspects, from rain gauges to discharge curves, from adequate satellite images to ground photos, which means the only possible way to go was a pragmatic approach of evaluating the data sources which are available (in this case: NOAA point precipitation data for Puerto Rico and the globally available ERA5-Land dataset) and use them for a basic validation of the model.

The usual calibration process, where the 2d model is fine-tuned via rain gauge data and discharge curves, could not be applied in this case. For this reason, I left the old version of the flowchart (see Figure 7) intentionally in the diagram and only crossed the part of the “normal” flood model calibration with red lines, to point out the difference between the usual workflow for this aspect and the chosen approach for this work, as a consequences of the external limitations (data scarcity).

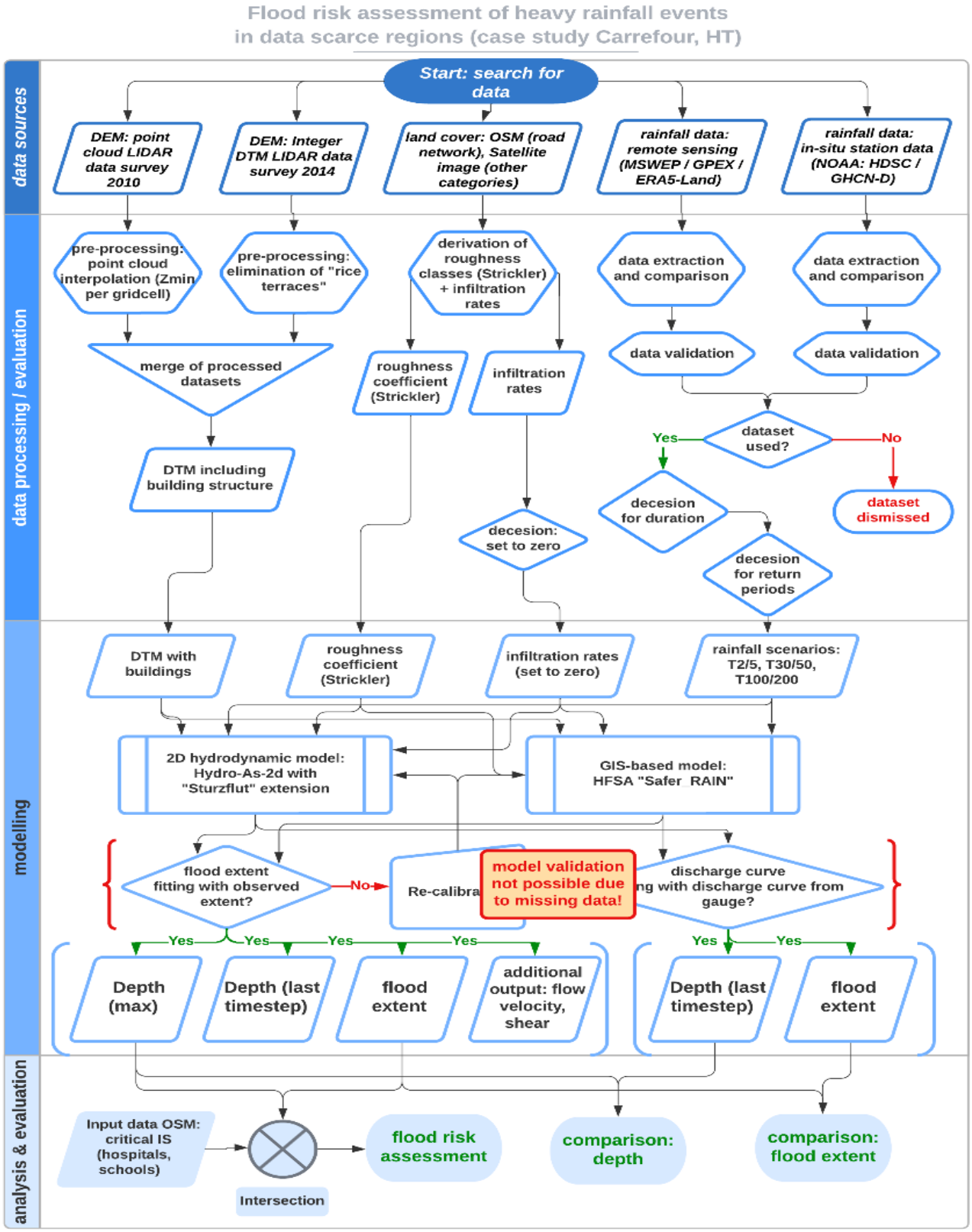


Figure 7: Flow diagram and decision tree for this work

1.2.2 Central research questions

This master thesis has two central research questions, the first being a methodological question and the second being a comparative one:

First, how can precise and spatially high-resolution flood calculations be carried out in areas with scarce input data? This issue is being addressed via the example of the town of Carrefour, Haiti.

Background for this research question is the high vulnerability of many Haitian towns and settlements to (pluvial) floods and the connection between the Salzburg-based Christian-Doppler-laboratory “GEOHUM” (“Earth Observation for Humanitarian Action”), which closely cooperate with the NGO Medicins Sans Frontières (MSF), who run several hospitals in Haiti.

Second, what are the differences in flood hazard calculation between an easy-to-setup merely GIS based approach (“Safer_RAIN”), which can be calculated via a global web platform **and a conventional 2d hydrodynamic flood calculation**, carried out with the industry standard software “HYDRO-AS 2d” combined with the “Sturzflut” extension developed by University of Innsbruck (Klar et al., 2014, see chapter 3.4.5)?

The thematic background for this second research question lies in my 7-year-professional activity in consulting engineer IB Humer GmbH from 2014-2021, where I have been working with Hydro-As-2D / Sturzflut regularly and also collaborated with gecosistema srl, an Italian SME, in several EU-funded research projects (“SWITCH-ON”, "SaferPLACES"). Gecosistema and the University of Bologna together developed a new GIS-based hierarchical filling-and-spilling approach for the modelling of pluvial floods, called “Safer_RAIN” (Samela et al., 2020 , see chapter 3.4.6.)

The scientific background is a DEM-based hierarchical filling-and-spilling algorithm, where the accumulating rainfall is routed via a D-8 algorithm (see Figure 64 in chapter 3.4.6) into the next sink until it´s fully covered with water. Then, accumulating water is routed into the next downstream sink or generally lower laying areas (Samela et al. 2020).

The advantage of this approach is an easy and very fast delineation of flood prone areas in case of heavy precipitation events, where the algorithm delivers good results (preferably rather flat areas). The applicability and accuracy of this approach shall be examined in comparison to a standard hydrodynamic flood calculation, carried out with the commercial hydrodynamic-numerical software “Hydro-As 2d”[18].

18 https://www.hydrotec.de/software/hydro-as-2d/ (access on 2022-10-02)

2. Study area and data

2.1. Carrefour: A coastal densely populated city

The area of interest is the municipal area of Carrefour, located directly west of the Haitian capital Port-au-Prince. Carrefour can be considered as a suburban city, resembling a patchwork of 13 villages. The population of Carrefour was 511,000 inhabitants in the 2015 census, making it the second-most populous city in the country[19]. Its administrative area has the size of 165 km², stretching from north to south in a length of ~20 km and ~10 km in east-west direction. The northern, coastal part is densely populated, whereas the southern upcountry parts are rural with a low population density. So although the population density of 3,096 inhabitants per km² is already high for the total municipal area, one has to keep in mind that the actual urban area, where more than 95% of the total population are living, amounts to no more than 30 km², which leads to a “felt” population density of around 16,000 inhabitants per km² (Vienna, Austria: 4,700 inh./km²).

Figure 8: View of Carrefour (eastern part) down from the hills (https://ayibopost.com/wp-content/uploads/2020/05/carrefour-2000x1125c.jpg, access on 2022-10-05)

19 Source: Institut Haitien de statistique et d’informatique (IHSI), https://ihsi.gouv.ht/ Demographic detail report from year 2015 is not available on the site, but an archived version can be found under https://web.archive.org/web/20151106110552/http://www.ihsi.ht/pdf/projection/Estimat_PopTotal_18ans_Menag2015.pdf (access on 2022-10-05)

Many parts of Carrefour´s today´s agglomeration are deprived areas and informal settlements, (bidonvilles) with a very high population density. They originate from the rural exodus of the past 30 years and out of natural disasters like the catastrophic earthquake of 2010.

Concerning the terrain morphology, the central parts of the urban area is situated on an alluvial fan of the Rivière Froide, which produces a slight but steady gradient form South to North (see Figure 9). The hills adjacent to the urban area reach altitudes of 300 to 600 metres. This special terrain morphology will play an important role in the flood calculation via the filling-and-spilling algorithm “Safer_RAIN”, more to this in chapter 4.2.

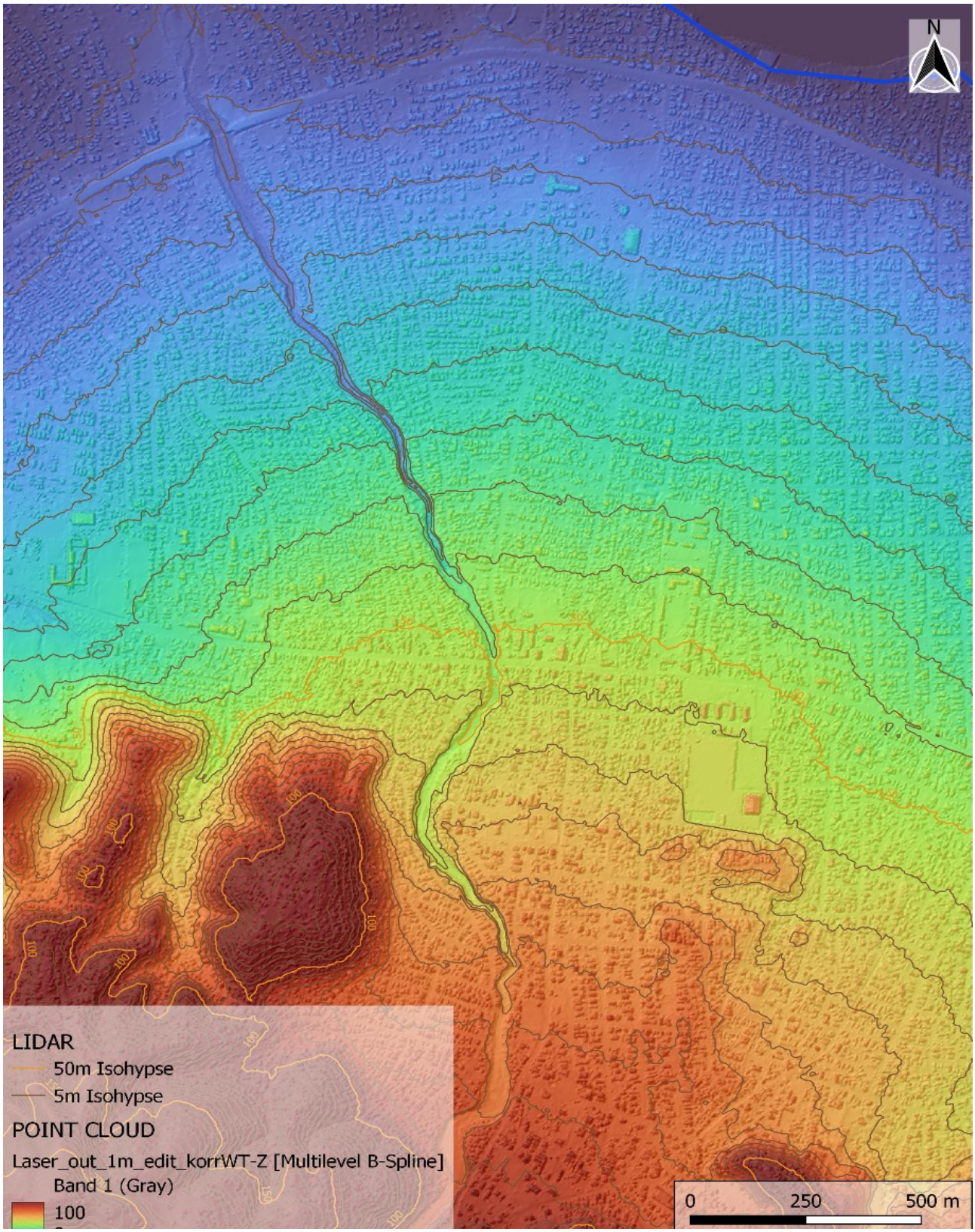
The Rivière Froide itself is deeply cut into the relief of the fan with an average slope height of according to the LIDAR dataset of 5-10 meters (Figure 87), which makes inundations out of the river not impossible, but rather unlikely. This has also been confirmed by the MSF staff, which I have been in contact throughout the whole period of this work.

So the bigger problems and inundation threats lie in the small *ravines* (gullies) as well as spontaneous overland flow coming down from the surrounding hills in case of a heavy precipitation event. In contrast to the main river Rivière Froide, which has a continuous channel flow throughout the year, the smaller ravines are leading water only after rainfall events and lie dry the rest of the year.

Due to these facts and due to the limited calculation capacity of the used software, I decided to limit my calculations to the urban areas of Carrefour plus the surrounding hills, from the western border to the neighbouring commune of Gressier until the border to Haiti´s capital Port-au-Prince, and in north-south direction from the Gulf of Gonave until the ridge (watershed) of the rising hills in the south (red outline of the AOI in Figure 10). The valley of Rivière Froide has been cut south of the city districts of Dufour / Vieux Caille.

The resulting area of interest has a size of 43 km², from which about 50% belongs to the poorly inhabited back country, which is important to include in the flood calculations until the watershed. The whole catchment of the Rivière Froide river (yellow outline in Figure 10) has a size of 62 km², with by far the largest part of it (more than 90%) outside of urban Carrefour and the upper parts belonging to the municipalities of Petionville and Kenscoff.

The river has a broad river bed with wide gravel bars in the middle section, which are able to retard and flatten the discharge peak in case of a heavy rain event covering the whole catchment. Also, the shape of the catchment with the “bottleneck” at the knee of the river, where it turns its flow direction from west to north with hardly any drainage from the left river side, does not speak for a sharply rising and falling discharge peak.


N
LIDAR
50m Isohypse
5m Isohypse
POINT CLOUD
Laser_out_1m_edit_korrWT-Z [Multilevel B-Spline]
Band 1 (Gray)
100
0
250
500 m

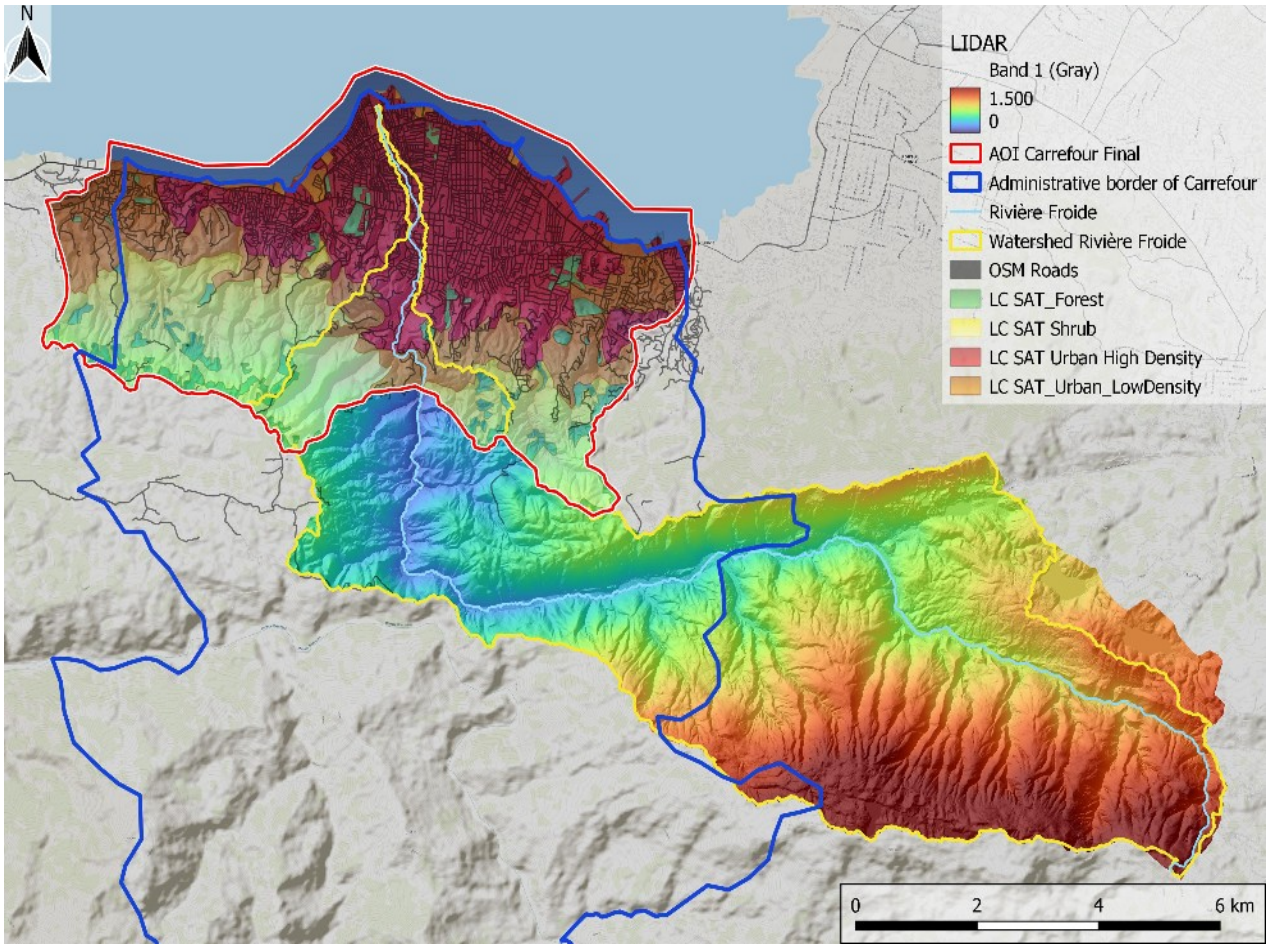


Figure 10: Area of interest (red), Catchment of Rivière Froide (yellow) and city borders of Carrefour (blue)

2.2. Terrain model (LIDAR)

Surprisingly enough, the most crucial input data for all flood calculations, a digital terrain model in high resolution, was not the most difficult ingredient to obtain. This is due to the fact that there have been carried out two LIDAR missions in Haiti:

The first one took place in the year 2010 in response to the January 12th magnitude 7.0 Haiti earthquake and covered the highly impacted areas of Port-au-Prince, Delmas, Carrefour and Leogane[20], and the second in the years 2014-2016, this time covering the whole state of Haiti[21]. Both of them have been published on the leading Open Data portal for LIDAR data, OpenTopography.org, but as usual, the devil is in the details, which will be illustrated in the following paragraphs.

20 World Bank - ImageCat Inc. - RIT Haiti Earthquake LiDAR dataset. Distributed by OpenTopography. https://doi.org/10.5069/G96Q1V50. Dataset access on 2021-12-05

21 HaitiData, The World Bank (2021). Haiti Digital Terrain Model 2014 - 2016. Distributed by OpenTopography. https://doi.org/10.5069/G9GX48R8. Dataset access on 2021-11-08

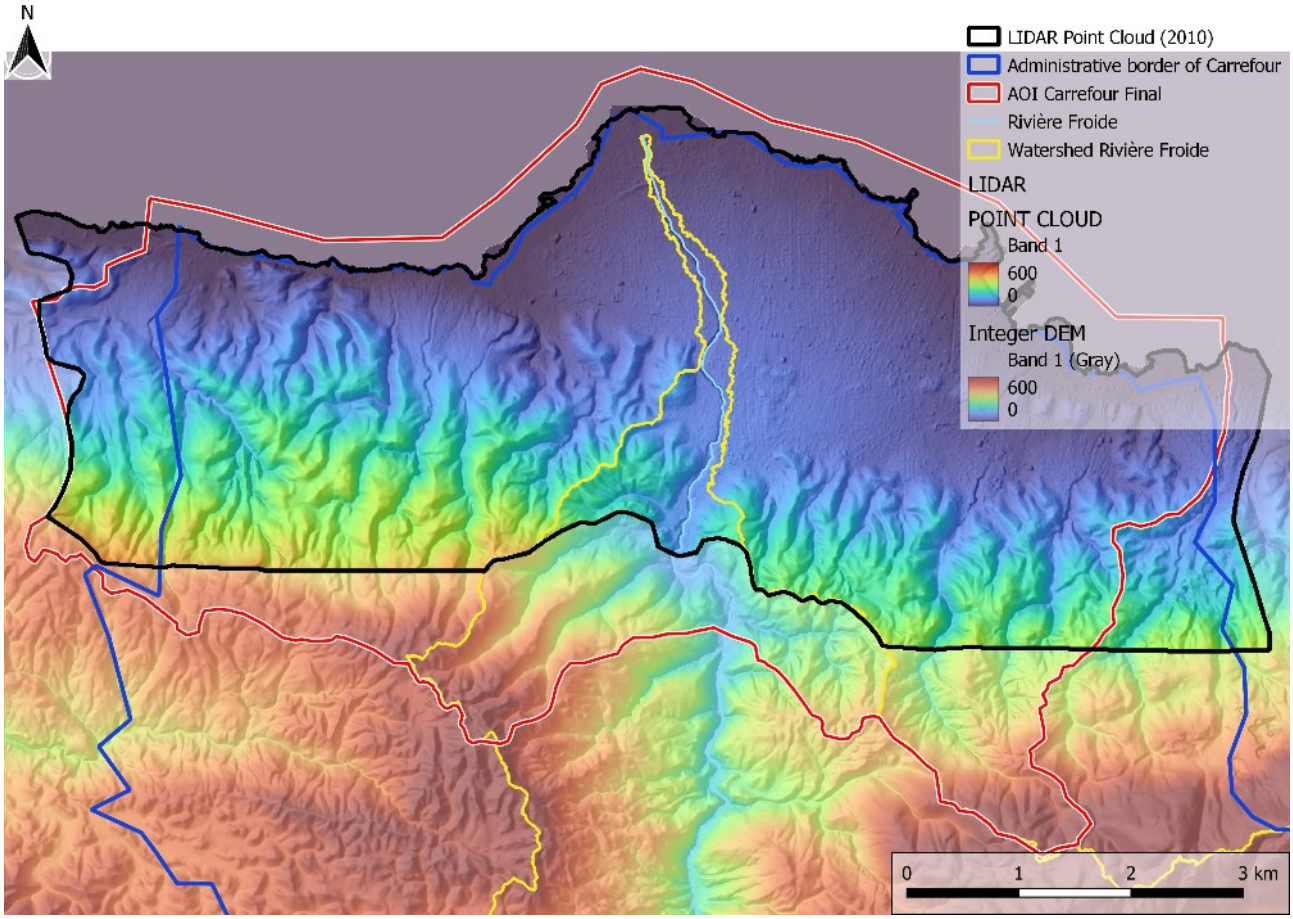


Figure 11: Extent of the Point Cloud LIDAR Data (black) in the AOI (red)

2.2.1 Point-cloud data: Deriving a DTM with buildings

For the 2010 dataset, the original point-cloud data is available, which has proven to be very useful, as one can derive the outlines of the buildings from it although the point density is rather low with 3.4 points per m² in average[22]. Via this process, two issues could be solved at once: First, we have a high-resolution digital terrain model, and second, the buildings with 16 m² and more are integrated in it. The only disadvantages are that the data is rather old (considering the dynamic of formal and informal settlement expansion in the urban areas) and that it does not cover the whole area, as can be seen in Figure 11.

The whole process in obtaining the DTM can be described as follows:

(1) **Choosing the right population:** The LIDAR point cloud is only poorly classified; out of a population of 2,847,342,691 points, only 5,938 (that is 0.0002 %!) are indeed classified as buildings, and not more than 11.6% are classified as ground points (last echo), whereas 87% are unclassified[23]. To get useful results, you can´t filter out only ground points plus buildings, which would be the easiest way otherwise, but rather consider all points and use a suitable interpolation method for converting the point cloud data into a continuous raster surface.

22 https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.072010.32618.1 (access on 2021-12-05)
23 https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.072010.32618.1 (access on 2021-12-05)

(2) **Selecting the best suitable interpolation algorithm:** Given the facts above, we need to trick a bit on the interpolation algorithm. If we choose a standard algorithm like IDW, the buildings and streets are represented the best way. But we also have a lot of trees and other vegetation in our output then, which is especially obstructive in the river- and channel beds as well as the valley grounds. Here, we would have countless “dams” consisting of vegetation which interferes with a consistent drainage behaviour. So I chose “minimum” instead, the cell values of raster cell X is represented by the minimum z value of all LIDAR points per raster cell.

In this manner, we do get some bias in our data (underestimation of z values per raster cell), streets and riverbeds are generally over-represented in the model (broader than they are), whereas buildings are under-represented (smaller than they are) due to the Zmin interpolation algorithm. But this is the only way to keep vegetation to a large extent out of our terrain data.

(3) **Selecting the best suitable interpolation resolution**: Now that we have a suitable interpolation method (min z), we have to find the best compromise for calculation capacity, building accuracy and a minimum of disturbing vegetation in our digital terrain model. This was indeed a trial-and-error technique, as illustrated in Figure 12 – I first went for a 1m resolution, but the amount of cells to calculate in the flood scenarios is 4 times higher than with a 2m resolution, and our AOI amounts to 43 km² anyway, which means 43 million cells with 1m resolution, but only 10.75 million cells with a 2m resolution, which indeed makes a difference for all calculations.

But even more important is the minimisation of vegetation, and even with the Zmin interpolation, we still had a lot of tree canopy in the resulting DTM, which is of course bad for all flood calculations. A **2 meter resolution with Zmin** as interpolation method turned out to be the **best compromise** between still capturing the basic outline of the building substance with a minimum of included tree canopy.

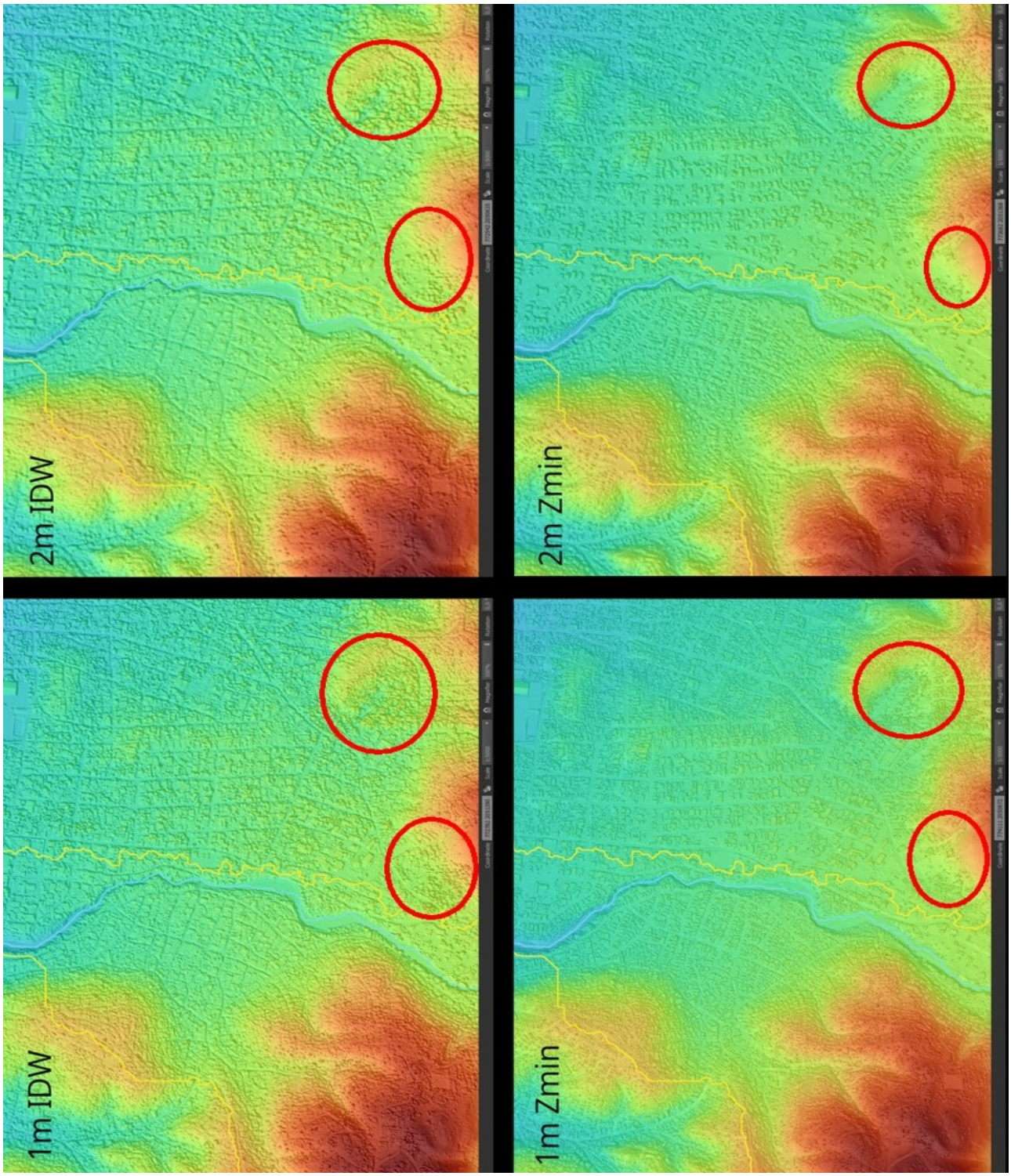


Figure 12: LIDAR: Comparison of different resolutions and interpolation techniques (red circles: tree canopy, to be minimized)

2.2.2 Integer DTM: Reducing the “rice terraces”

The 2016 dataset covers all of Haiti, but has a severe disadvantages for serving as terrain model for flood calculations: the data type of the DTM (point cloud data is not accessible) available on OpenTopography is **integer**, which leads to some kind of “Chinese rice terrace landscape” and produces senseless results for flood calculation if it is used as it is, as illustrated in Figure 14. The integer data type maybe a good way to save storage size compared to the float32 data type as it has **no decimals**, but this leads to the paradox situation that although the DTM is consistent and covering the whole area, for **this task**, its highly likely to produce error-prone results in flood extent and water depth, **unless** it is extensively preprocessed (see Chapter 3.4.1). The terrain height of all objects is expressed in whole meters, all decimal values are cut – and in fact, for a continuous surface like a DTM, this procedure is very problematic.

To be able to use the DTM nevertheless, it has been preprocessed in the following steps(see Figure 13):

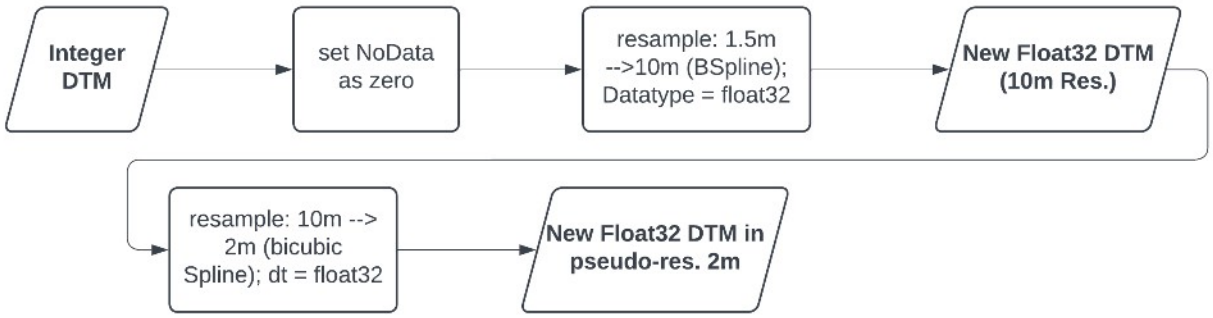


Figure 13: Preprocessing of the Integer DTM

Via this “downsampling” to a lower resolution and afterwards “upsampling” again to a higher resolution (2m to be consistent with the chosen point cloud output), the terrace landscape can be blurred and smoothened a bit, so that the water in the GIS- and 2d model is able to flow downstream.

To our advantage, the more important urban areas, as well as the flatter areas near the coast, are all covered by the 2010 point-cloud data, so we need the integer DTM only for some steep areas in the back country, until the watershed is reached. But this amounts for less than 20% of the total AOI (see Figure 11)

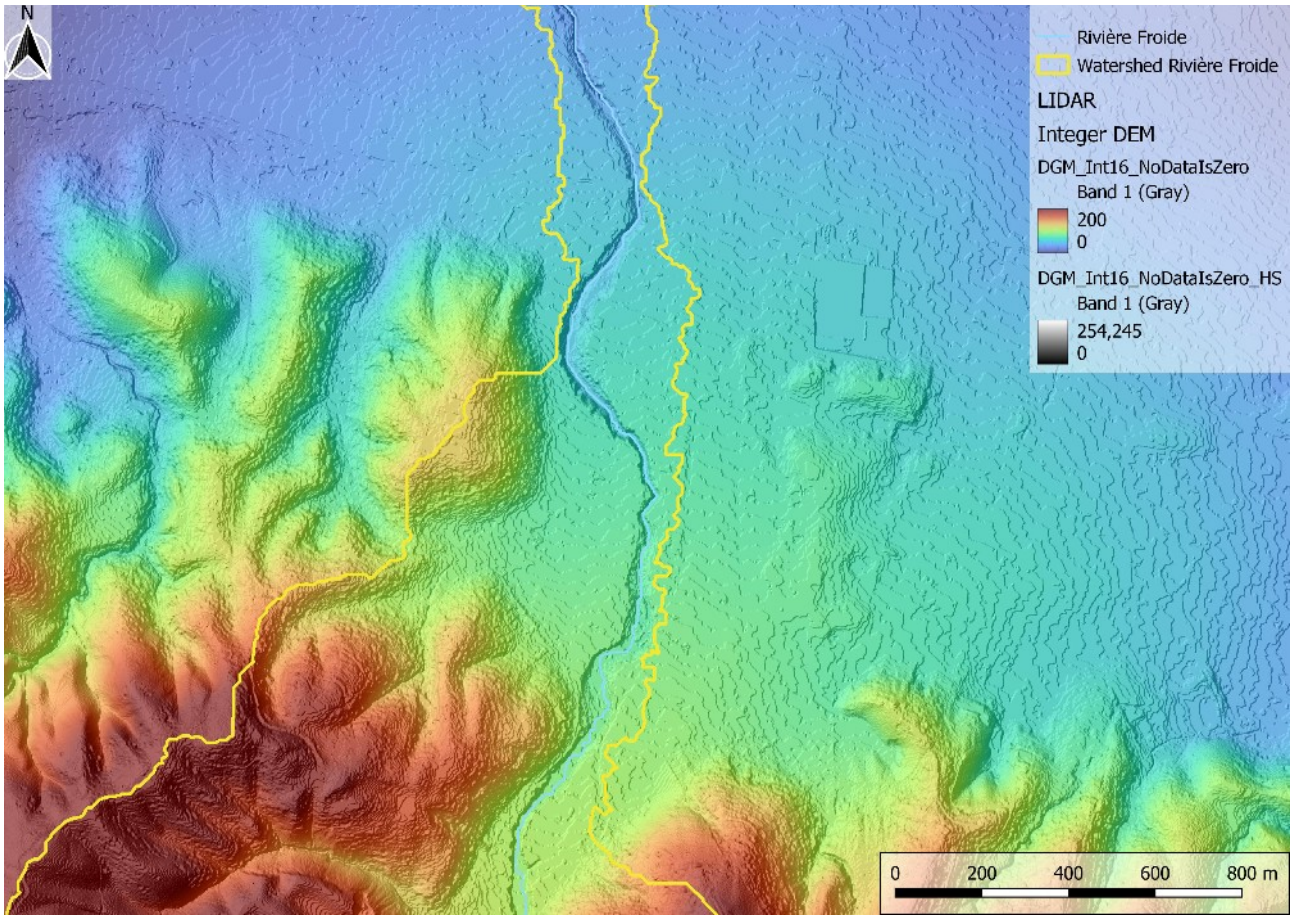


Figure 14: Hill shade of the Integer DTM from HaitiData.org in Carrefour

2.3. Rainfall data: Sources and data formats

Obtaining rainfall data was one of the toughest tasks for this work and hence one of the “core areas” of the widespread data scarcity. Whilst there are numerous global gridded precipitation datasets available by now, **ground truth in form of rain gauge station data was not available for this work**. Also, not every global precipitation dataset is suitable for the estimation of precipitation extremes. Figure 16 on the following page gives an overview on some available datasets (the table dates from the paper on the MSWEP dataset published in 2017, so ERA5-Land and GPEX are not included in it, as they are of more recent origin).

The selection of dataset described in the following chapters is based upon a preselection of the most important global precipitation datasets at the time of elaborating this thesis and is non-exhaustive, as one could easily write not only a master thesis, but a whole dissertation on this topic alone. It concentrates on recent datasets which are eligible for the derivation of precipitation extremes plus two relevant local data sources (UHM / Haitian Government, NOAA / US National Oceanic and Atmospheric Administration for the neighbouring island of Puerto Rico).

It is important to note that while **ground station based precipitation** data like the NOAA datasets are commonly distributed in **ASCII data format** (TXT / CSV), all **satellite based precipitation** (CMORPH, MSWEP, ERA5-Land) **or extreme value datasets** (GPEX) are available only either in

GRIB (GRIdded Binary / General Regularly-distributed Information in Binary form) or in **NetCDF** (network Common Data Form) data format. Both are **multidimensional binary formats** with time and i.e. altitude levels as third and fourth dimension (see Figure 15). **QGIS** can read both GRIB and NetCDF via the mesh module for data visualization, while the scripting languages **Python** and **R** can access the data directly and calculate and export exactly what the user needs.

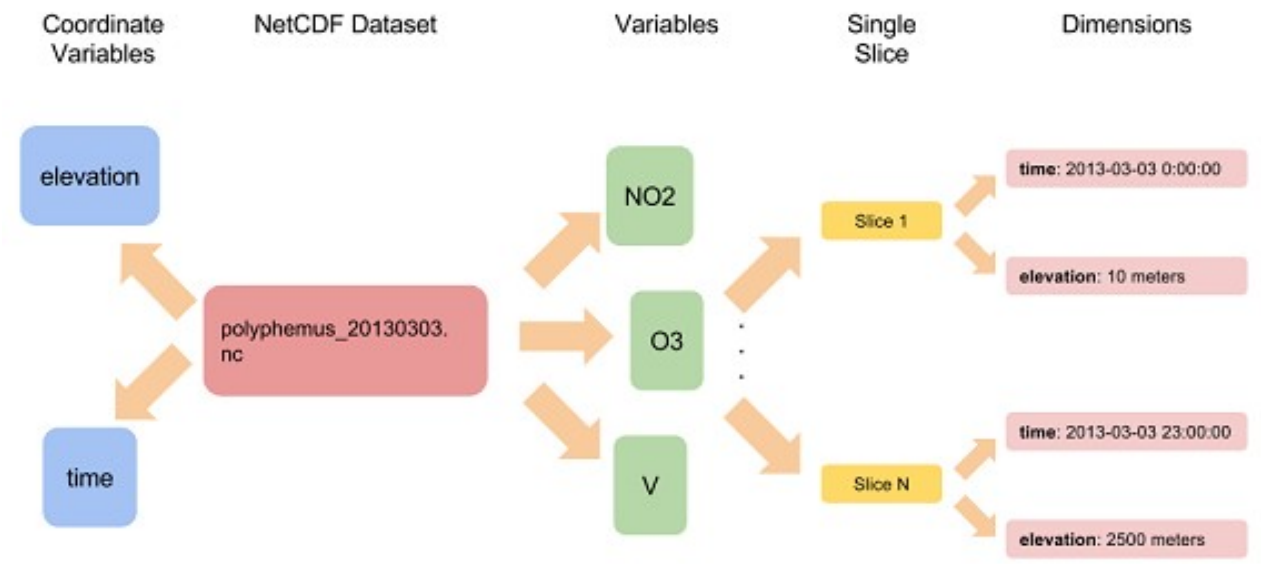


Figure 15: internal structure of a netCDF file - similar for GRIB files (https://docs.geoserver.geo-solutions.it/edu/en/_images/exampleNetCDF.png, access on 2023-01-04)

GRIB is the WMO (World Meteorological Organisation) standard binary format for the exchange and distribution of gridded datasets and exists in multiple editions: While GRIB edition 1 (GRIdded Binary) is only still in use for aviation weather forecast at this day, GRIB Edition 2 (General Regularly-distributed Information in Binary Form) is in broad use as the standard format for numerical weather prediction model output. Both editions of GRIB are designed for the timely dissemination of large volumes of gridded data.[24]

Essential abilities of the GRIB format are:

- The ability to store multiple variables in a single file, such as temperature, precipitation, humidity, and wind speed
- The ability to store data at multiple levels of the atmosphere, such as surface-level data and data at different heights above the surface (height as dimension)
- The ability to store data at multiple forecast times, allowing users to view how weather conditions are expected to change over time (time as dimension)

24 https://www.eumetsat.int/formats (access on 2023-01-04)

NetCDF is "*a machine-independent, self-describing, binary data format standard for exchanging scientific data, including climate and weather data. It is designed for platform independent, self-describing data storage, with the intent that users can understand the data without the need for external resources and access it efficiently in its entirety or in portions.*"[25]

NetCDF was developed by the Unidata Program Center at the University Corporation for Atmospheric Research (UCAR) and is widely used in the atmospheric and oceanic sciences, as well as other fields such as hydrology, ecology, and geology. NetCDF files consist of a header and a data section, with the header containing metadata about the data, such as variable names, units, dimensions, and attributes, and the data section containing the actual numerical data (see Figure 15)

Some of the key features of the NetCDF format include[26]:

- Support for multidimensional arrays, allowing data to be organized into a variety of dimensions, such as time, latitude, longitude, and depth
- Support for metadata, including variable names, units, and other attributes, which can be used to describe the data and make it easier to interpret
- Support for compression, which can be used to reduce the size of large datasets and make them easier to store and transfer
- Support for parallel I/O, which allows multiple processes to read and write data to a NetCDF file simultaneously, improving performance for large-scale data analysis

But not only in data storage format, ground-based gauge data and raster-based sensor data gained via satellite remote sensing do differ. There is also a big difference in the way, the data is gained:

- **Rain gauge data** is punctual data and therefore, hydrologists speak of "**point precipitation**".
- In contrary, **satellite gained remote sensing datasets**, who measure the accumulated precipitation for a certain spatial unit (grid cell) during a certain time period (time step), calculate **average precipitation amounts valid for exactly one grid cell**.
    - For example, the **ERA5-Land parameter "total precipitation in m"** is defined as the "*accumulated liquid and frozen water, comprising rain and snow, that falls to the Earth's surface. It is the sum of large-scale precipitation (…) at spatial scales of the grid box or larger and convective precipitation (...) at spatial scales smaller than the grid box.*"[27]

Finally, an important note can be found in the description of the ERA5-Land "total precipitation" parameter:

25 https://www.eumetsat.int/formats (access on 2023-01-04)
26 https://www.unidata.ucar.edu/software/netcdf/ (access on 2023-01-04)
27 https://apps.ecmwf.int/codes/grib/param-db?id=228 (access on 2023-01-04)

> *"Care should be taken when comparing model parameters with observations, because observations are often local to a particular point in space and time, rather than representing averages over a model grid box."*[28]

This leads us to the concept of the **area reduction factor (ARF)**. In hydrology, the ARF is a key parameter in the design for hydrological extremes. Simplified, the ARF takes into account the fact that the **intensity of a precipitation event** is **indirect proportional** to the **size of affected area** and the **duration of the precipitation event** (amongst other factors, like the return period, seasonality and prevailing type of precipitation, see Chapter 3.2.2): The bigger the area, and the shorter the duration, the higher the ARF applied on point precipitation data.

The remote sensing datasets which measure precipitation via radar, summarize the measured precipitation to the area of the spatial resolution, which automatically levels and averages the precipitation amount for the grid cell X and duration D. If the dataset has a spatial resolution of 0.1*0.1 degrees, the corresponding grid cell has a length of approximately 11.1*10.5 kilometres at 18.5°N, which accounts for an area of 116.25 square kilometres.

According to technical literature (see Breinl et al., 2020, p. 680, Svensson and Jones, 2010, p. 5, Langousis, 2005, p. 12) for a duration of D= 1 hour and an area of ~100 km², the ARF can be assumed with a value between 0.6 and 0.7- which means you have to multiply rain gauge data by ~0.65 to be able to compare the data to grid cell data valid for the whole area. Or, in the other direction: multiply the remote sensing dataset values with this resolution with 1.55, to be able to compare it with in-situ point precipitation data from a rain gauge. More on this in Chapter 3.2.2.

28 https://apps.ecmwf.int/codes/grib/param-db?id=228 (access on 2023-01-04)

Name	Full name and details	Data source(s)	Spatial resolution	Spatial coverage	Temporal resolution	Temporal coverage	Reference(s)
Unified	Climate Prediction Center (CPC) Unified	Gauge	0.5°[a]	Global	Daily	1979–present	Xie et al. (2007), Chen et al. (2008)
	Climatic Research Unit (CRU) Time-Series (TS)	Gauge	0.5°	Global	Monthly	1901–2014	Harris et al. (2013)
	Global Precipitation Climatology Centre (GPCC) Full Data Reanalysis and First Guess	Gauge	0.5°[b]	Global	Monthly	1901–present	Schneider et al. (2014)
/L	PRECipitation REConstruction over Land (PREC/L)	Gauge	0.5°	Global	Monthly	1948–present	Chen et al. (2002)
	University of Delaware (UDEL)	Gauge	0.5°	Global	Monthly	1901–2014	Matsuura and Willmott (2009)
PS	Climate Hazards group Infrared Precipitation with Stations (CHIRPS)	Gauge, satellite	0.05°	50° N–50° S	Daily	1981–present	Funk et al. (2015a)
RPH	CPC MORPHing technique (CMORPH)	Gauge, satellite	0.25°	60° N–60° S	Daily	1998–present	Joyce et al. (2004)
-1DD	Global Precipitation Climatology Project (GPCP) 1-Degree Daily (1DD) Combination	Gauge, satellite	1°	Global	Daily	1996–2015	Huffman et al. (2001)
P-MVK	Global Satellite Mapping of Precipitation (GSMaP) Moving Vector with Kalman (MVK)	Gauge, satellite	0.1°	60° N–60° S	Hourly	2000–present	Iguchi et al. (2009)
G	Integrated Multi-satellitE Retrievals for GPM (IMERG)	Gauge, satellite	0.1°	60° N–60° S	30 min	2014–present	Huffman et al. (2014)
ANN-CDR	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) Climate Data Record (CDR)	Gauge, satellite	0.25°	60° N–60° S	6 hourly	1983–2012	Ashouri et al. (2015)
3B42	TRMM Multi-satellite Precipitation Analysis (TMPA) 3B42	Gauge, satellite	0.25°	50° N–50° S	3 hourly	1998–present	Huffman et al. (2007)
A-Land	Modern Era Retrospective-Analysis for Research and Applications (MERRA)-Land	Gauge, reanalysis	0.5° × 0.67°	Global	Hourly	1979–present	Reichle et al. (2011)
	Princeton global meteorological Forcing Dataset	Gauge, reanalysis	0.25°	Global	3 hourly	1948–2012	Sheffield et al. (2006)
I	WATCH Forcing Data ERA-Interim (WFDEI)	Gauge, reanalysis	0.25°	Global	3 hourly	1979–2014	Weedon et al. (2014)
-CFSR	National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR)	Reanalysis	0.3125°	Global	Hourly	1979–2010	Saha et al. (2010)
nterim	European Centre for Medium-range Weather Forecasts ReAnalysis Interim (ERA-Interim)	Reanalysis	0.25°[c]	Global	3 hourly	1979–2014	Dee et al. (2011)
5	Japanese 55-year ReAnalysis (JRA-55)	Reanalysis	1.25°	Global	3 hourly	1959–present	Kobayashi et al. (2015)
fPS	Cooperative Institute for Climate Studies (CICS) High-Resolution Optimally Interpolated Microwave Precipitation from Satellites (CHOMPS)	Satellite	0.25°	60° N–60° S	Daily	1998–2007	Joseph et al. (2009)
AIN-ASCAT	Based on Advanced Scatterometer (ASCAT) data (Brocca et al., 2016)	Satellite	0.5°	Global	Daily	2007–2015	Brocca et al. (2014)
	CPC Merged Analysis of Precipitation (CMAP)	Gauge, satellite, reanalysis	2.5°	Global	5 day	1979–present	Xie and Arkin (1996, 1997)
EP	Multi-Source Weighted-Ensemble Precipitation (MSWEP)	Gauge, satellite, reanalysis	0.25°	Global	3 hourly	1979–2015	This study

olution for the coterminous USA. [b] 1° spatial resolution for 2014–present. [c] ~ 80 km effective spatial resolution (i.e., the resolution of the employed atmospheric model).

2.3.1 Haitian Government: UHM / CNIGS

The official Government institution for collecting rainfall data in Haiti is the Unité Hydro-Météorologique d'Haïti (UHM)[29]. It is stated on the web portal, that the institution is leading an own network of automated rain gauge stations throughout the country, also some papers refer to it. Unfortunately, the link to it leads nowhere, and I did not receive any response to my inquiries for data.

The Centre national de l'information géo-spatiale (CNIGS)[30] offers 5 datasets combined with precipitation on the HaitiData.org portal:

- dataset “pluie 3-5 octobre 2016”: 24h precipitation sums of Hurricane Matthew hitting the west of the island[31], see Figure 17, Figure 21 and Figure 25)
- datasets “Annual Precipitations (Min/Max/Average) 1998 2010 NATHAT TRMM 34B2”: Annual minimum / maximum / average precipitation sum between 1998 and 2010[32]
- dataset “Haiti Meteorology Rain Stations NATHAT CNM Point”: Location of Haitian rain gauges, but with no data combined.[33]

Unfortunately, this does not lead us far neither for a model validation, nor for a calibration (as we don´t have discharge sums), nor for being able to estimate a typical return period of i.e. 5, 30 or 100 years. For this, we need long-time observation data in a smaller time interval, not monthly, yearly or daily, although daily is better than nothing. But heavy precipitation events causing spontaneous overland flow are typically short-term events, lasting for one or two hours, maybe three, but seldom more.

It would make sense to calculate a “hurricane scenario” on the data provided, after a cross-check of the precipitation sums with other data sources, but that would go beyond the scope of this work. For defining a certain return period, the data is not very helpful, as we neither have certainty on the data accuracy, nor is it possible to classify the event in relation to other precipitation events.

A problem of the “pluie 3-5 octobre 2016” dataset is the fact that the precipitation amounts seem to be interpolated between the particular community sections. The rain gauge station network in Haiti is not very dense, so the interpolation between gauge data alone contains major uncertainties in data accuracy. Moreover, the single value for each administrative area (on community section level) leads to a systematic bias compared to the actual rainfall distribution, as we have – again, like in the integer DTM – a “rice terrace effect” in the data, divided into irregular areas at that. So for sure, this data has to be handled with the utmost care.

29 https://www.meteo-haiti.gouv.ht/index.html, access on 2023-02-07
30 http://www.cnigs.ht/, access on 2023-02-07
31 https://haitidata.org/layers/geonode_data:geonode:pluie3_5_octobre2016, access on 2023-02-07
32 https://haitidata.org/search/?limit=5&offset=0&title__icontains=precip, access on 2023-02-07
33 https://haitidata.org/layers/geonode_data:geonode:hti_meteorology_rainstations_nathat_cnm_point, access on 2023-02-07

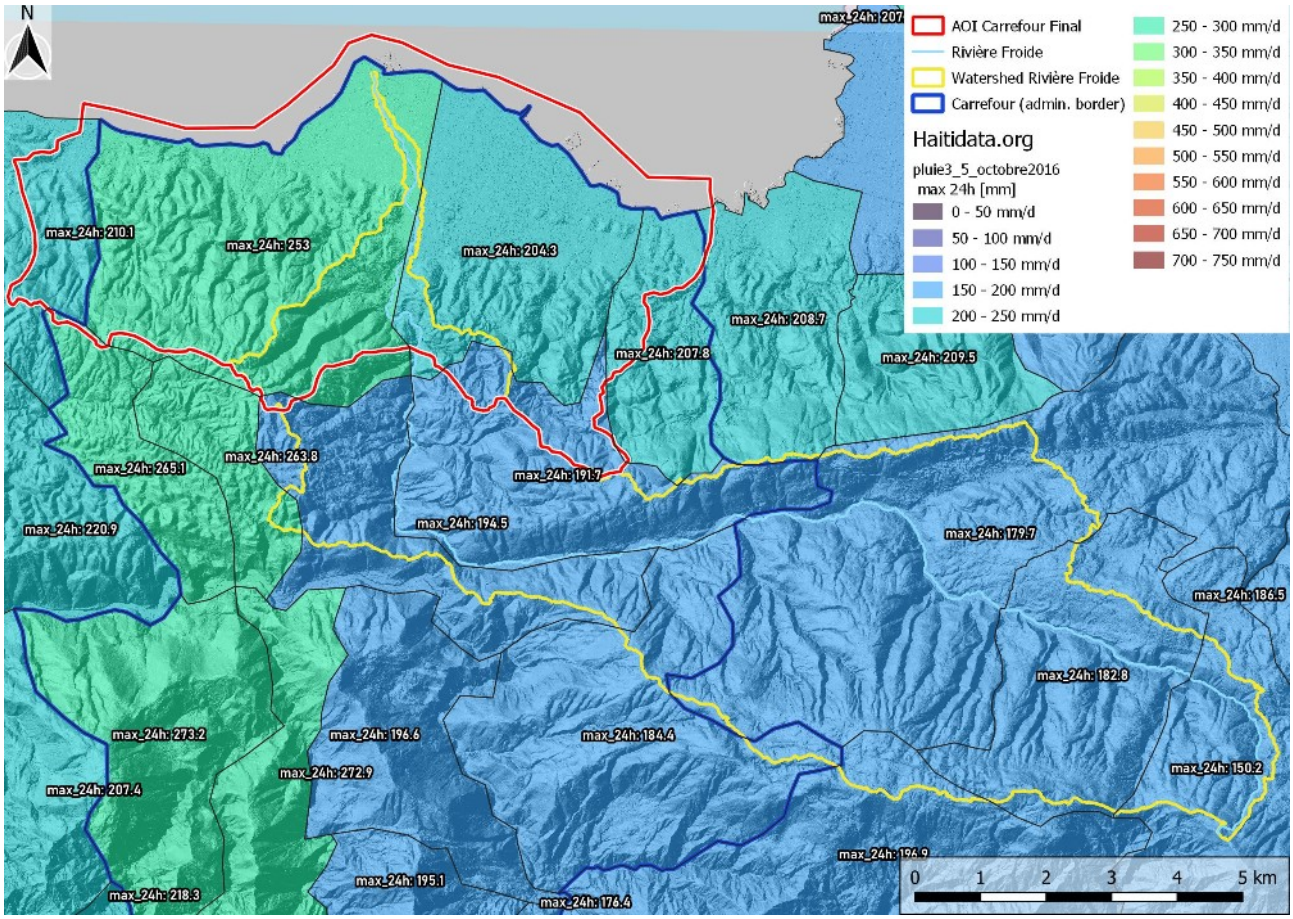


Figure 17: Rainfall Data Hurricane Matthew 2016: Maximum precipitation in 24h (data source: https://haitidata.org/layers/geonode_data:geonode:pluie3_5_octobre2016 ; access on 2022-10-10)

2.3.2 CMORPH (Climate Prediction Center Morphing Technique)

The CMORPH (Climate Prediction Center Morphing Technique) dataset ranges from 60°N to 60° S and has a spatial resolution of 0.25 degrees (Joyce et al. 2004) and a temporal resolution of 0.5 hours. The dataset had been developed by the Climate Prediction Center (CPC) of the US National Oceanic and Atmospheric Administration (NOAA) and is based on passive microwave and infrared data (Joyce et al. 2004). It provides information on both the amount and distribution of precipitation, and is widely used for a variety of purposes, including the development of climate models, the assessment of drought conditions, and the study of extreme precipitation events.

Roughly, there are two available products: First, the raw microwave-only data, which are based on the corresponding satellite measurements only (see Figure 18), and second, the bias-corrected product, which includes in situ based sources (Figure 19). Bias in the raw CMORPH is removed through Probability Density Function (PDF) matching against the CPC daily gauge analysis over land and through adjustment against the pentad GPCP merged analysis over ocean (Xie et al. 2017).

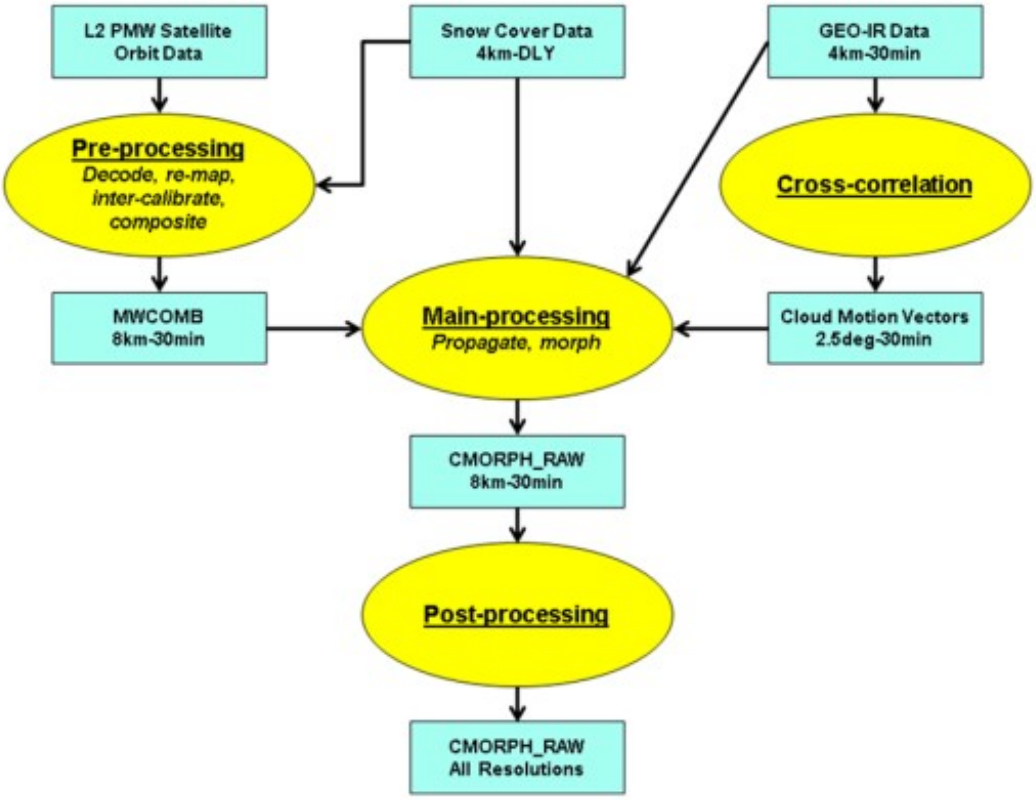


Figure 18: Flowchart for the CMORPH RAW satellite precipitation estimates (Xie et al., 2017, p. 1620)

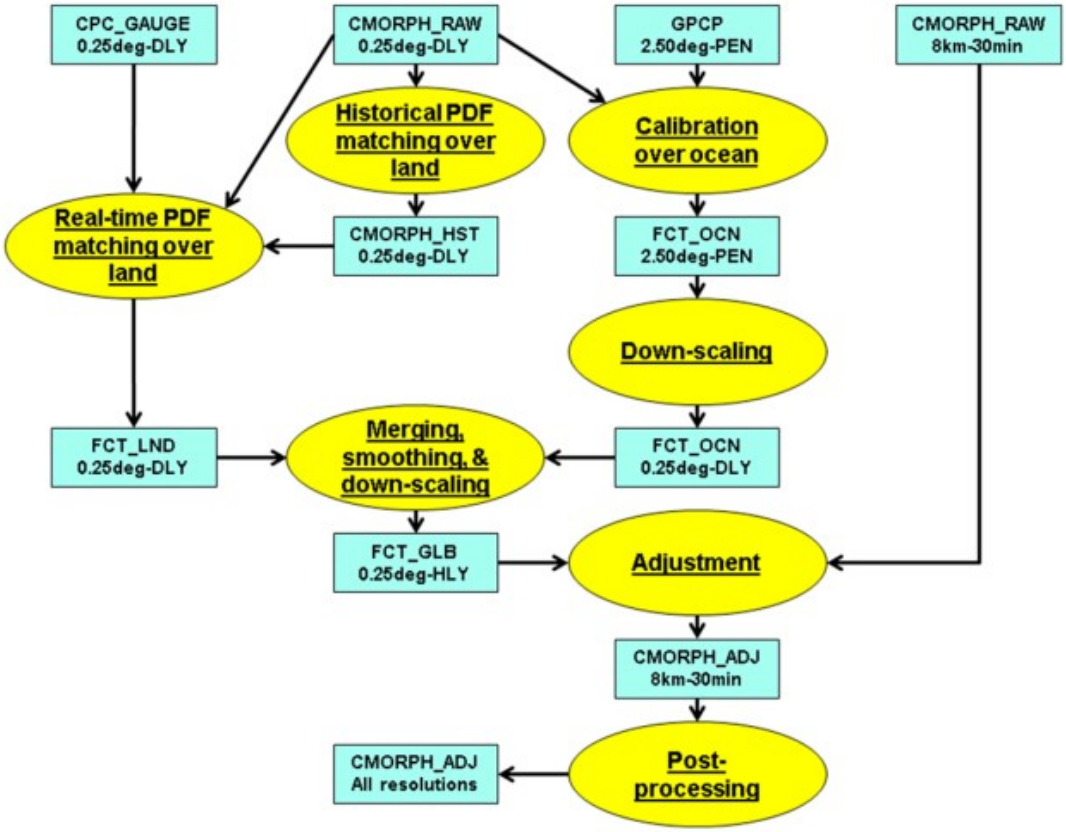


Figure 19: Flowchart for the CMORPH bias correction procedure (Xie et al., 2017, p. 1626)

The data can be downloaded either in a special binary format at the NOAA FTP server[34] or in netCDF or even ASCII format (for smaller datasets, i.e. a narrow region or just a few time steps) at the Climate Data Library of Columbia University[35].

CMORPH does **not** deliver a rain amount X for a certain return period Y itself, this must be done manually via extreme value statistics (i.e. GEV, Gumbel, etc.) with the downloaded datasets; of course, the more data, the better. To give a hint whether the data would be suitable in this region for a statistical evaluation, I downloaded the corresponding dataset (daily mean in mm/h) for the timespan of October 2nd to 6th 2016, where Hurricane Matthew crossed the western peninsula of Haiti from South to North.

Though our AOI of Carrefour was not in the centre of the hurricane track, at least according to (the only freely available) dataset of Haitian authorities on HaitiData.org, there still should have taken place substantial precipitation with a 24h maximum between 180 and 250 mm on Oct 4th 2016 – which corresponds to an average rainfall intensity of 7.5 to 10.5 mm/h (see Figure 21). CMORPH data shows rainfall intensities between 3.5 and 4.7 mm/h for our AOI (see Figure 20), so by factor 2 smaller rainfall intensities than the UHM / CNIGS ground based data.

Keeping in mind that, the area reduction factor (ARF, see Chapter 3.2.2) for an area of around 730 km² (corresponding grid cell in CMORPH – 0.25*0.25 degrees, which are equivalent to ~26*28 km at the latitude of 18.5 degrees North[36]) and a duration of 24 hours can be assumed with a value of 0.9 according to different literature sources (see Breinl et al., 2020, p. 680, Svensson and Jones, 2010, p. 5, Langousis, 2005, p. 12), the intensities in CMORPH are still considerably lower than those of Haitian Government authorities, which presumably refer to interpolated rain gauge data.

Concerning the rainfall pattern as such, the pattern of CMORPH is quite similar and consistent with the UHM / CNIGS picture – the region with the highest amounts of rainfall lies in the Departement du Sud, around the city of Les Cayes. This corresponds to reports on floodlist.com, where the Les Cayes region was reported to have been hit hardest by Hurricane Matthew[37].

34 ftp.cpc.ncep.noaa.gov (access on 2023-02-10)
35 http://iridl.ldeo.columbia.edu/ (access on 2023-02-10)
36 The factor, by which the degrees of longitude (starting from 40,075 km earth circumference / 360 degress = 111.1 km at the equator) have to be mulitiplied with, are equivalent to the cosinus of the corresponding latitude. Therefore applies: degrees longitude at 18.5° North (AOI) = 111.1 * cos (18.5) = 105.36 km; 0.25° lon at 18.5°N = 26.34 km
37 https://floodlist.com/america/hurricane-matthew-causes-deaths-haiti-dominican-republic (access on 2022-10-07)

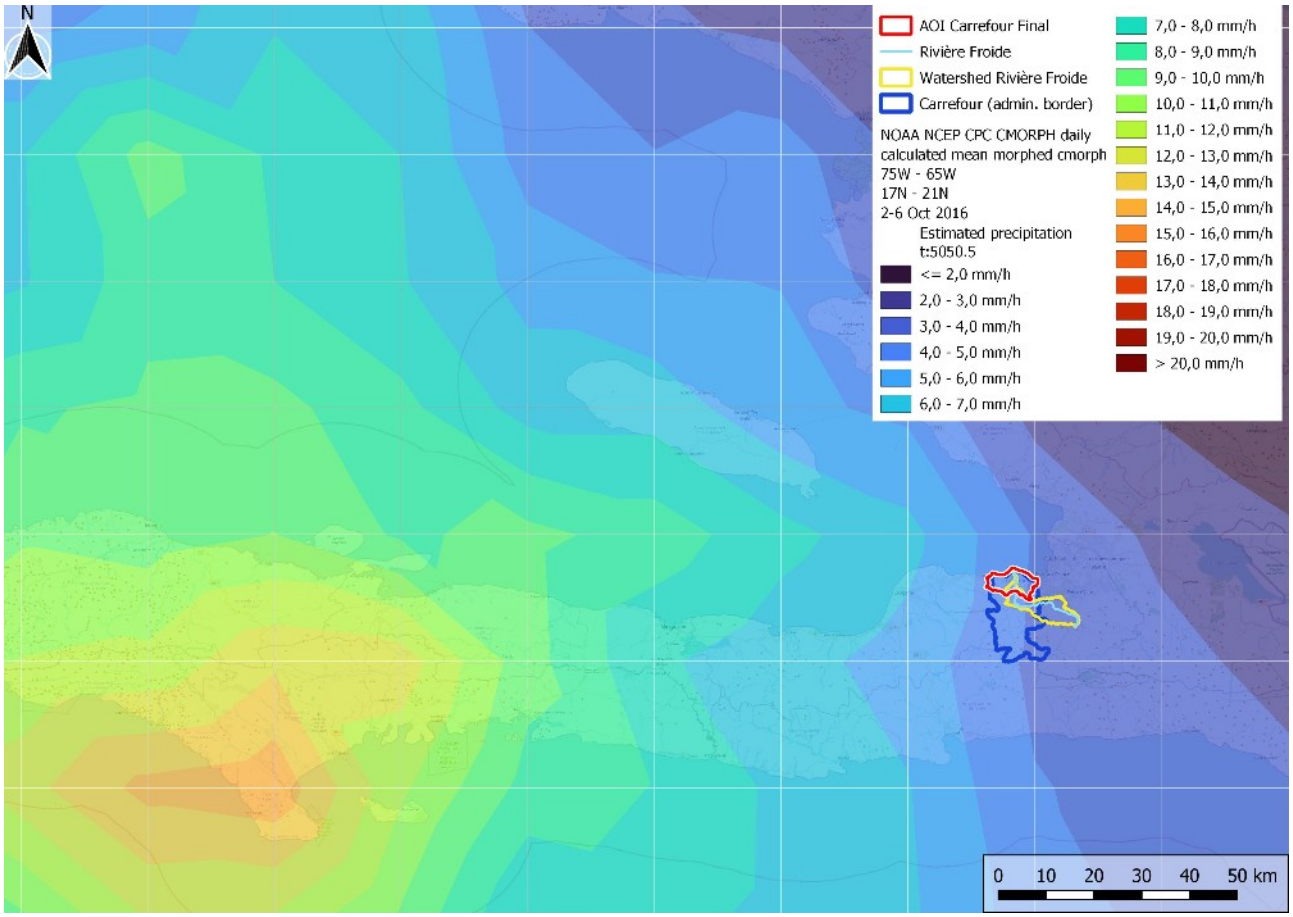


Figure 20: NOAA NCEP CPC CMORPH daily calculated mean in mm/h for 4 Oct 2016 (data source: http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.CPC/.CMORPH/.daily_calculated/.mean/.morphed/.cmorph/, access on 2023-02-14)

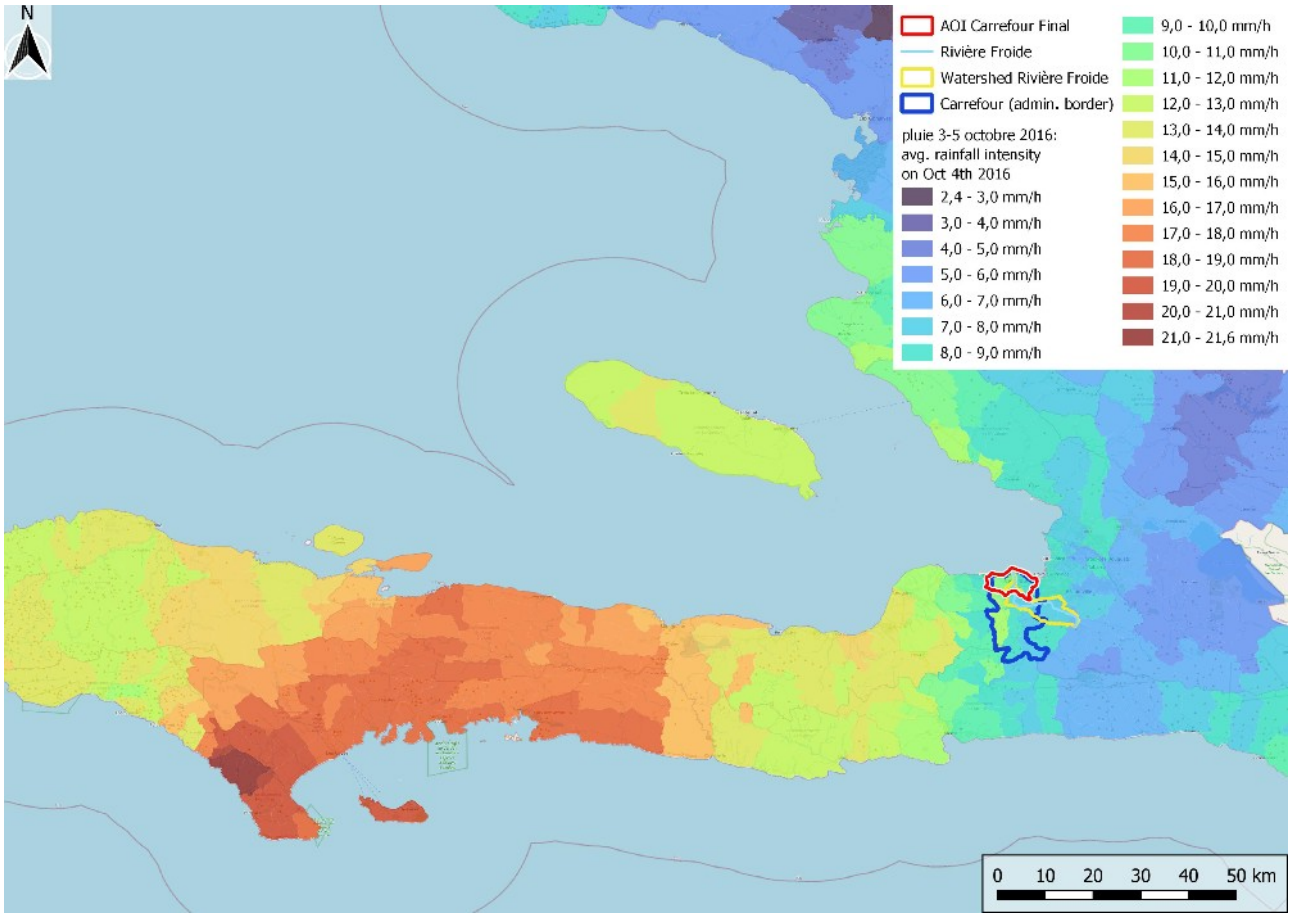


Figure 21: Average rainfall intensity in mm/h on Oct 4th (Daily precip. / 24) according to the CNIGS dataset

2.3.3 MSWEP (Multi-Source Weighted-Ensemble Precipitation)

The MSWEP (Multi-Source Weighted-Ensemble Precipitation) dataset is "*the first fully global precipitation dataset with a 0.1° resolution derived by optimally merging a range of gauge, satellite, and reanalysis estimates*" (Beck et al., 2019). It is a global, historic precipitation dataset (1979–until present) with a 3-hourly temporal and 0.1° spatial resolution and "c*ombines gauge-, satellite-, and reanalysis-based data to provide reliable precipitation estimates over the entire globe*" (Beck et al. 2019).

The data is free to use for non-commercial purposes and can be downloaded from a Google Drive via registration at http://www.gloh2o.org/mswep/. To examine the data, I downloaded the data for the timespan 3rd to 5th October 2016 (Hurricane Matthew) for our AOI in Carrefour and the neighbouring provinces.

The pattern for Oct 4th (daily precipitation sum) is again similar, with very high precipitation sums up to 750 mm for 24 hours in the Western peninsula mountain region inland of coastal city Les Cayes (see Figure 24)– here in this region, the precipitation sums are **considerably** higher than in CMORPH (around twice the values of CMORPH, where the highest precipitation sum for Oct 4th 2016 amounts for ~350mm / 24h).

But this delta in precipitation amounts does **not** persist until our AOI in Carrefour resp. the catchment of Rivière Froide: Here, CMORPH has values between 85 and 120 mm for 24h. MSWEP has similar values here, between 85 and 105 mm for the whole day, compared to 180 – 250 mm according to the Haitian dataset (see Figure 25). The resolution of MSWEP is higher, and it includes more data products and ground datasets than CMORPH, which is the older product.

MSWEP has a grid cell size of ~115 km² (0.1*0.1 degrees, which are equivalent to ~10.5*11.1 km at the latitude of 18.5 degrees North) in these latitudes, which accounts for an area reduction factor (see Chapter 3.2.2) of 0.9 – 0.95 for the duration of 24h (Langousis, 2005), so not much of the delta in absolute sums between MSWEP and UHM/CNIGS in our AOI Carrefour can be explained with that. But as long as we don´t know the ground truth in our area (the origin of the UHM / CNIGS data for Hurricane Matthew based on the smallest administrative units is unclear), it´s only an assumption which product is better in reproducing accurate rainfall sums. For this reason, only the differences between the remote sensing products and the Haitian dataset can be shown here, without being able to define the most accurate dataset.

Looking on the pattern of the highest precipitation amounts in Figure 24 seems – taking into account the hurricane storm path and the topography of the island (highest rainfall sum in the mountains) – quite plausible, in any case more than the rainfall pattern of the CNIGS data (see chapter 2.3.1).

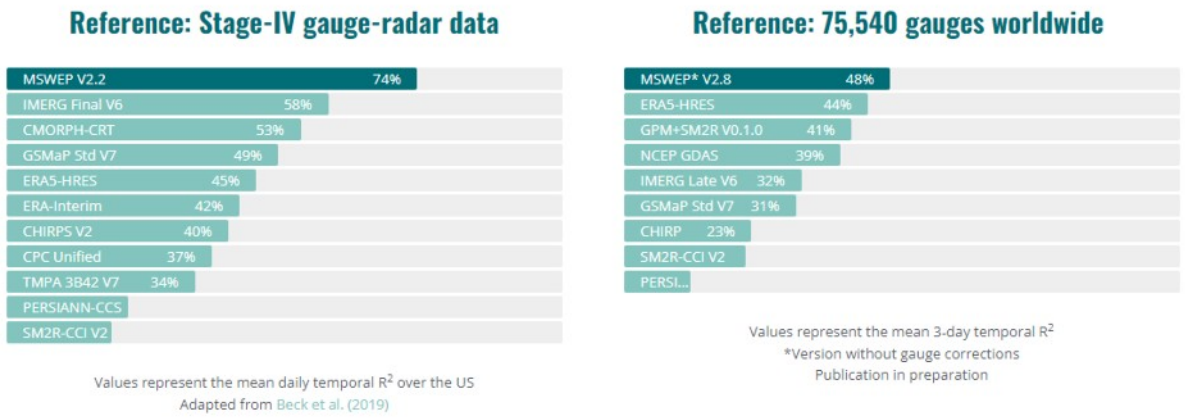


Figure 22: Performance of MSWEP according to their producers (http://www.gloh2o.org/mswep/, access on 2023-02-12)

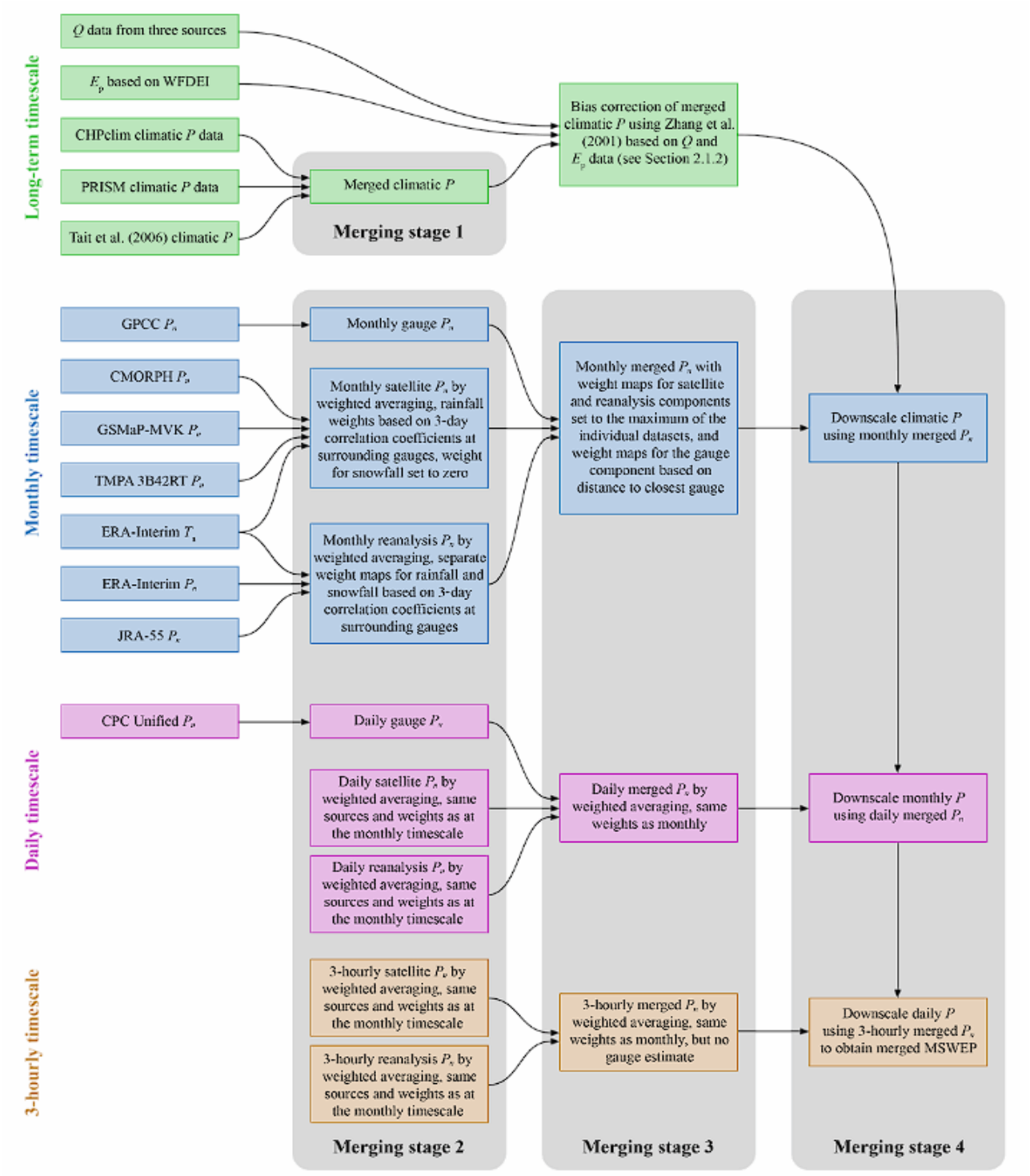


Figure 23: Flowchart of the main processing steps to produce MSWEP V2 (source: BECK et al. 2019)

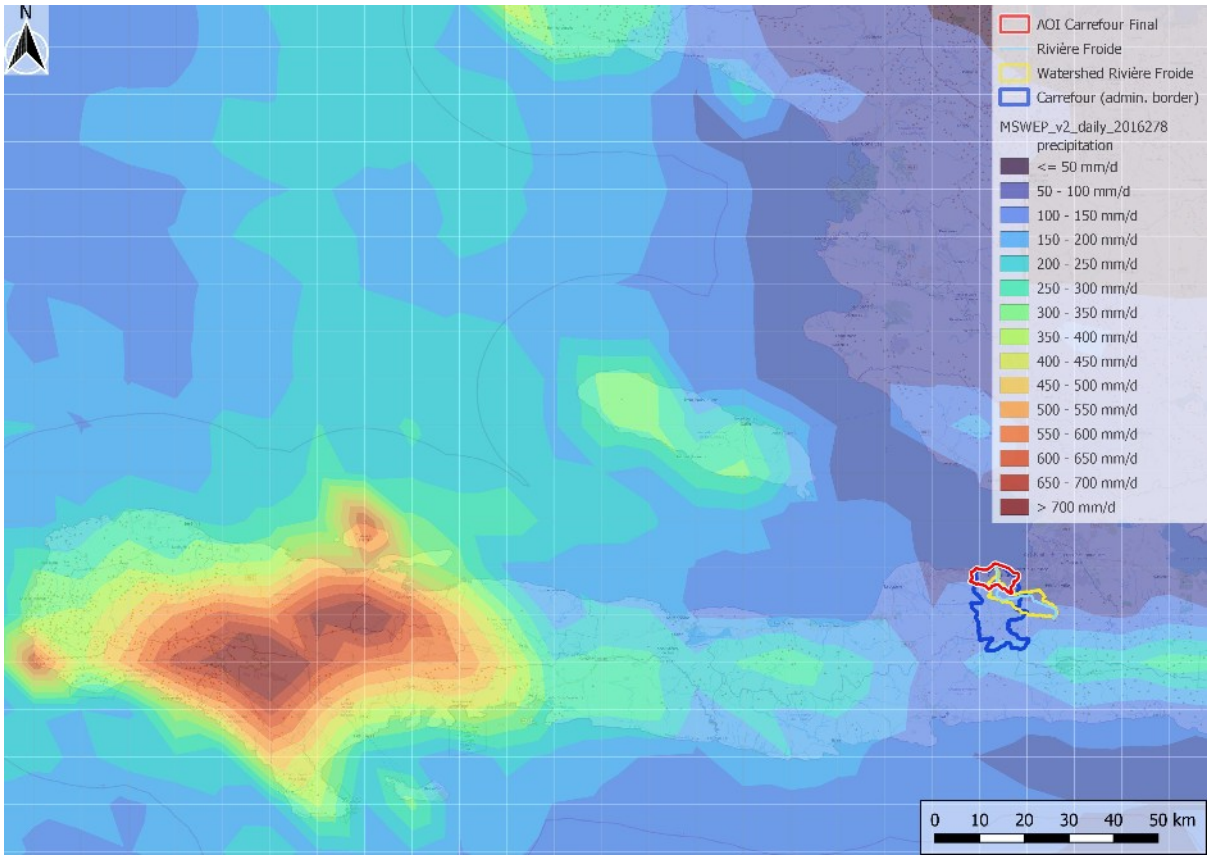


Figure 24: Gloh2o MSWEP v2 daily precipitation for 4 Oct 2016

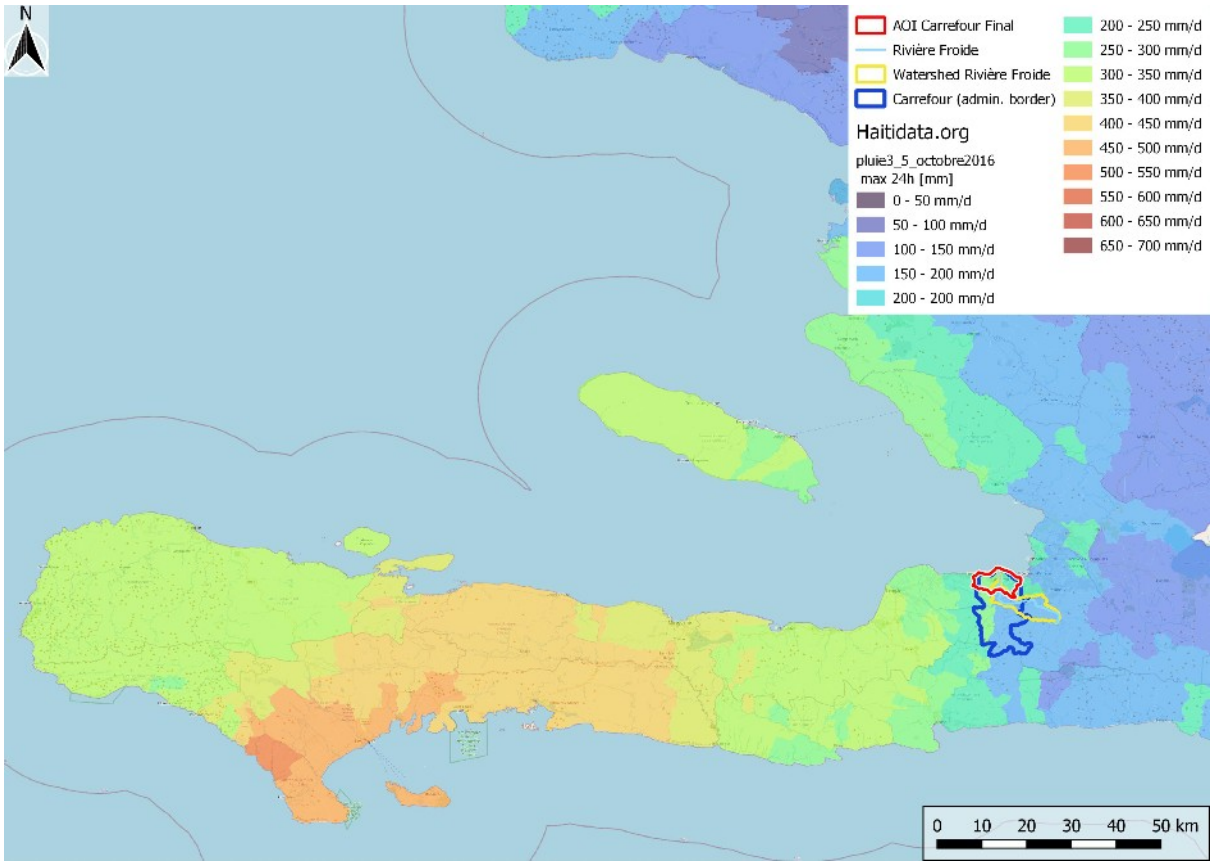


Figure 25: Daily precipitation on Oct 4th according to CNIGS / HaitiData.org

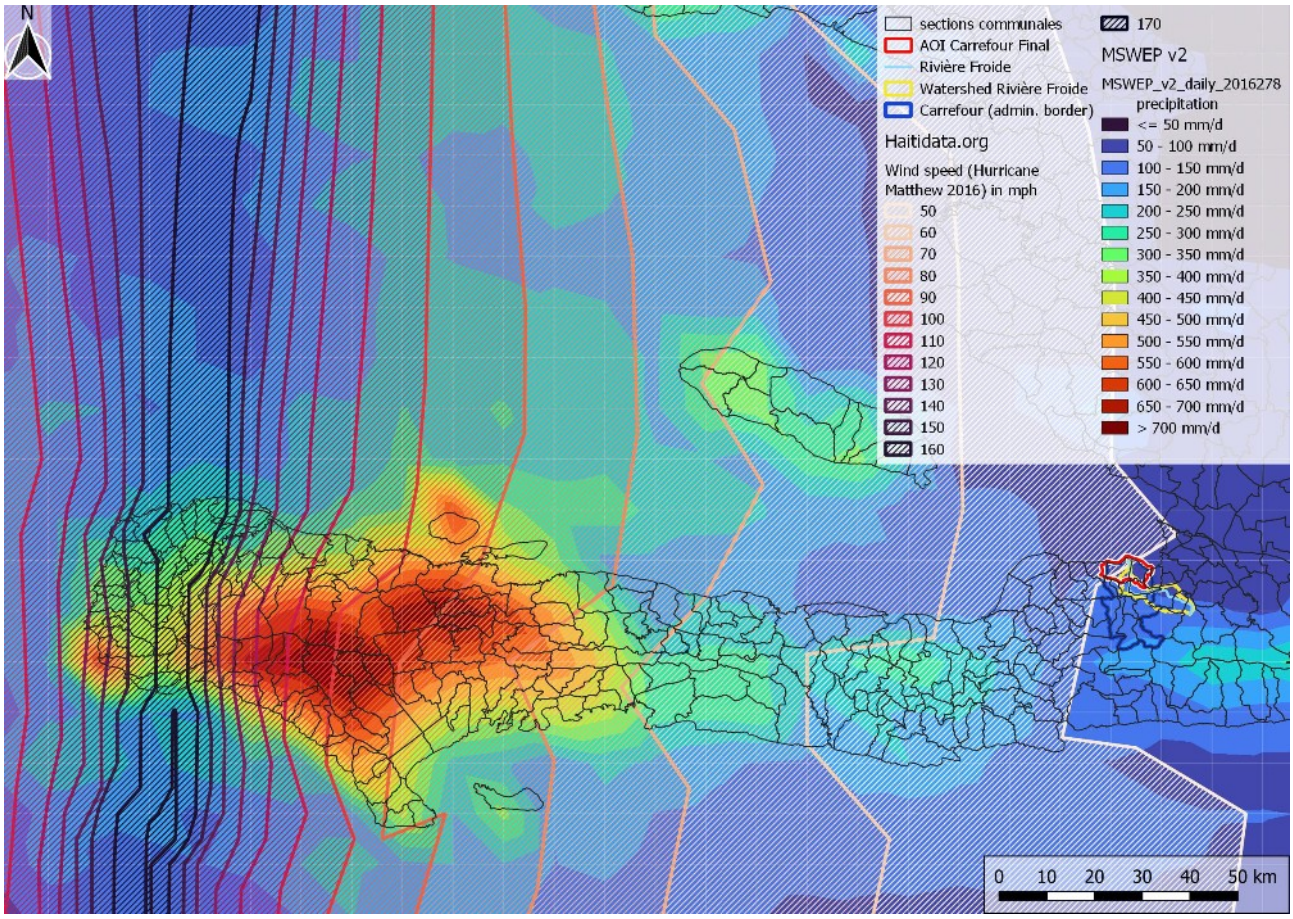


Figure 26: Storm path of Hurricane Matthew in 2016 combined with daily precipitation sum on Oct 4th 2016 according to MSWEP v2

2.3.4 GPEX (Global Precipitation Extremes)

The GPEX (Global Precipitation Extremes) dataset, published as research data by **TU Delft**[38] (Gruendemann et al., 2020) is a global collection of data related to extreme precipitation events. The dataset provides information on extreme precipitation events at both global and regional scales, and is used by researchers, water resource managers, and government agencies to inform decision-making related to extreme precipitation and its impacts.

It differs from the two other datasets mentioned before (MSWEP & CMORPH) in the fact that GPEX explicitly provides information on the intensity, duration and frequency (IDF) of extreme precipitation events and is used for a variety of purposes, including the assessment of flood risk, the design of water resource infrastructure, and the development of climate models.

GPEX is, like the other datasets, a hybrid dataset based on satellite observations, rain gauge measurements and numerical weather prediction models. GPEX is strongly based upon the MSWEP dataset, so it has the same spatial resolution of 0.1 arc degrees. All IDF curve information is based on the precipitation amounts calculated by MSWEP (38 years of data availability, so n=38). It works with 3 different extreme values distributions, namely the widespread Generalized Extreme Value (GEV) distribution, the Peak-Over-Threshold (POT) method and the Metastatistical Extreme

38 https://data.4tu.nl/articles/dataset/GPEX_Global_Precipitation_EXtremes/12764429 (access on 2023-02-15)

Value (MEV) distribution (Gründemann et al., 2021) While the first two methods only use the largest events to estimate precipitation extremes, the MEV distribution *"includes all precipitation events, leads to smoother spatial patterns of local extremes"* (Gründemann et al. 2021).

The data is processed and quality-controlled to ensure its accuracy and consistency, and is available for the following return periods and durations:

- **return periods:** 2, 5, 10, 20, 39, 50, 100, 200, 500, 1000 years
- **durations:** 3, 6, 12, 24, 48, 72, 120 and 240 hours

If we take a closer look into the AOI Carrefour / the Rivière Froide catchment, the GEV distribution shows 24h precipitation amounts of 350mm in the coastal region and almost 500mm in the Upper part of the Rivière Froide Catchment for the 100 year return period (see Figure 27). According to the GPEX GEV data, the precipitation sum of Hurricane Matthew on Oct 4th 2016 (basis 24h rainfall sum) in our AOI was not more than a T2 to T5 event (2 to 5 years return period, see Figure 28).

This could actually reflect reality well, as the actual hurricane storm path was about 200km away in the very west of the Tiburon Peninsula, in the Grad Anse Department (see Figure 26). The rain event will indeed have been of minor importance to the Carrefour area, also, there are no reports of floods and damages for this area concerning Hurricane Matthew.

The paper of (Gruendemann et al., 2020) states that "t*he traditional Generalized Extreme Value (GEV) distribution and Peak-Over-Threshold (POT) methods, which only use the largest events to estimate precipitation extremes, are not spatially coherent. The recently developed Metastatistical Extreme Value (MEV) distribution, that includes all precipitation events, leads to smoother spatial patterns of local extremes"*.

I examined the MEV distribution as well, and for the Hurricane Matthew event, it showed for the Carrefour AOI similar (although **continuously lower**) values for:

- the lower return periods (i.e. for T5)
- for the raster cells in low-lying areas, which corresponds to the cells with lower absolute precipitation amounts

whereas for the other combinations, so higher return periods, i.e. T100, and mountainous raster cells with high absolute precipitation amounts (and thus also high rainfall intensity values in the IDF curves), the delta between GEV and MEV distribution is approximately the ratio 1:0.6 (see Figure 29 in contrast to Figure 27).

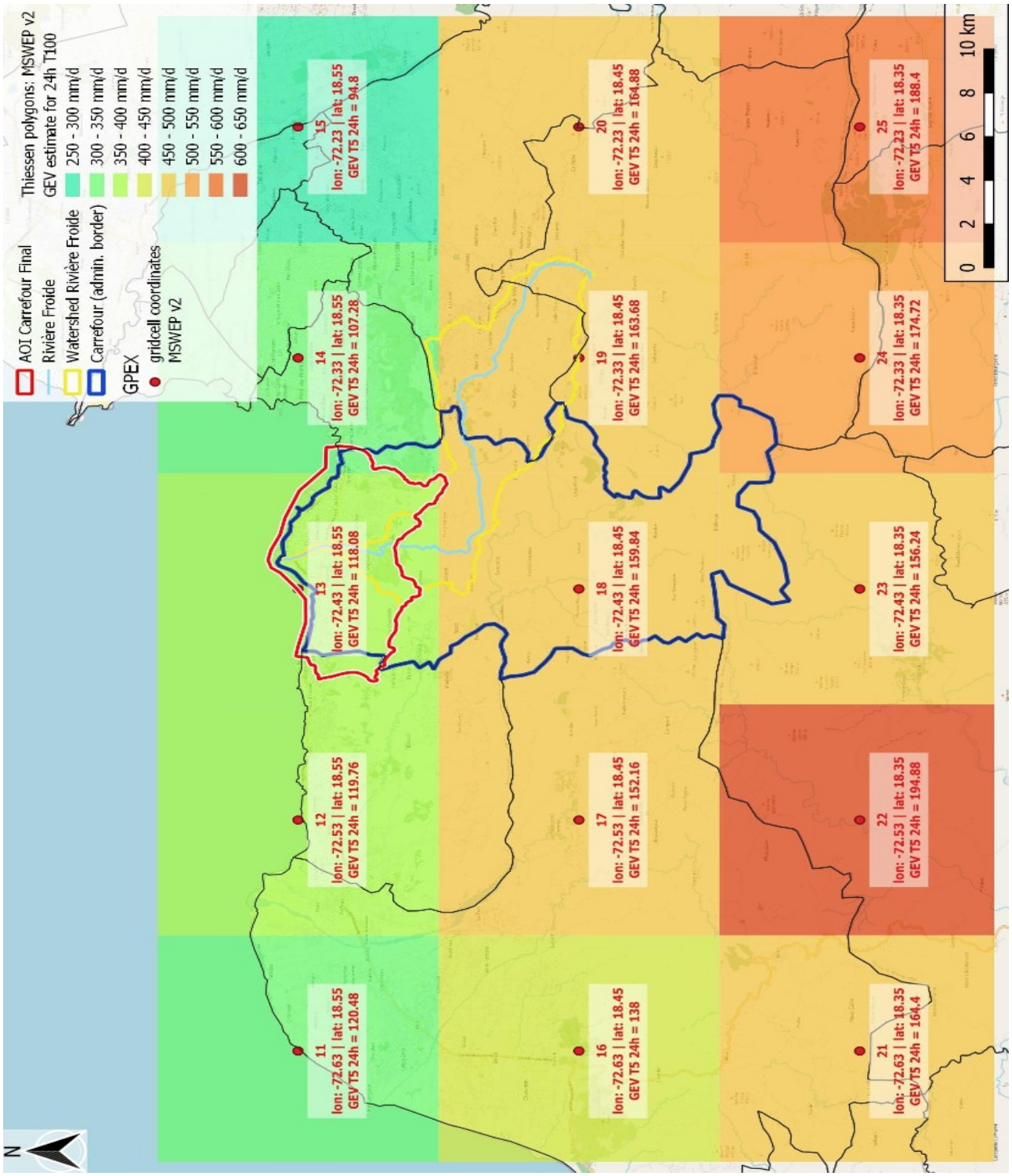


Figure 27: GPEX GEV estimates for a 24h / T100 rain event (data source: Gruendemann et al., 2020)

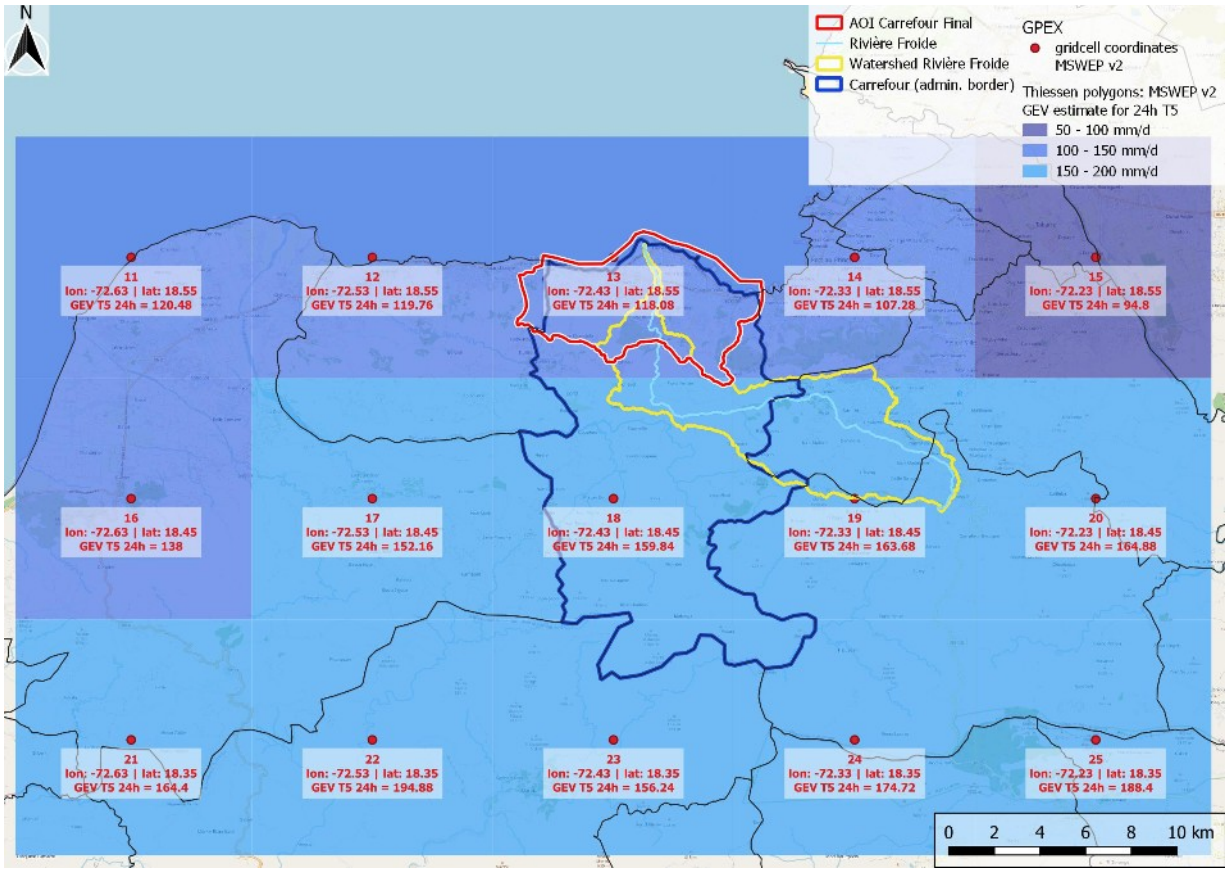


Figure 28: GPEX GEV estimates for a 24h / T5 rain event (data source: Gruendemann et al., 2020)

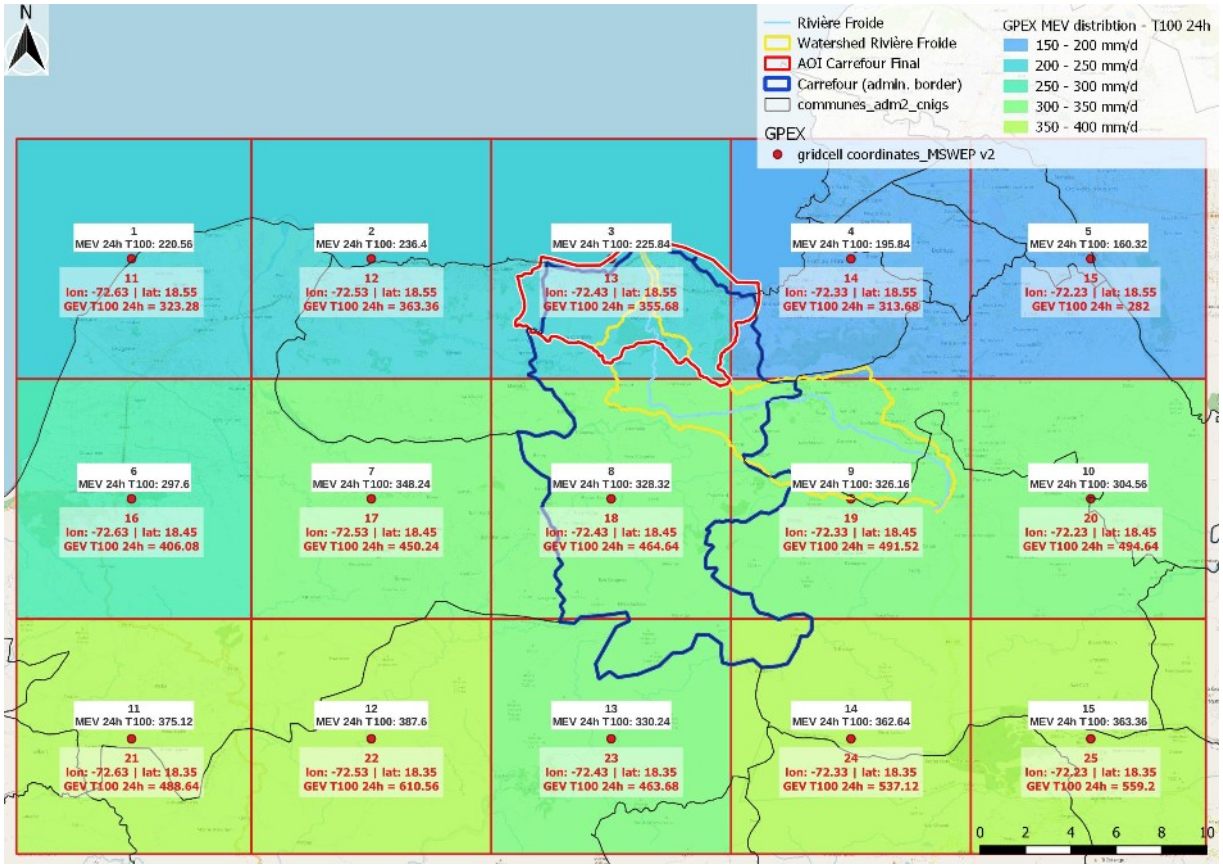


Figure 29: GPEX MEV estimates for a 24h / T100 rain event (data source: Gruendemann et al., 2020)

Via libraries and packages in the **scripting languages Python and R**, it is also possible to **access the data in the netCDF files directly**, without visualization in a GIS client, and to evaluate and further process the data included. In this manner, with help of the “extRemes” library for R, I was able to derive IDF curves (see Chapter 3.2.3 for more detailed information on this) from the GPEX values (GEV and MEV distribution) and compare them against each other as well as to the manually fitted GEV quantiles in R from the observed annual maxima of the corresponding duration (see Figure 30 and Appendix 6.1 Extraction of IDF Curves from GPEX dataset).

The results confirm the first impression illustrated above, the MEV distribution (taking into account all precipitation events instead of the annual maxima only) shows consistently lower values for all durations and return periods, but especially for the higher return periods. This makes sense when regarding the methodology of the MEV distribution, which takes into account all precipitation events of a certain region and time period, unlike the GEV or Gumbel distribution, which take into account only the annual maxima. But it is not possible to tell, if the derived values for i.e. a T100 event with duration 24 hours are more accurate with this approach, with the (few) information we have.

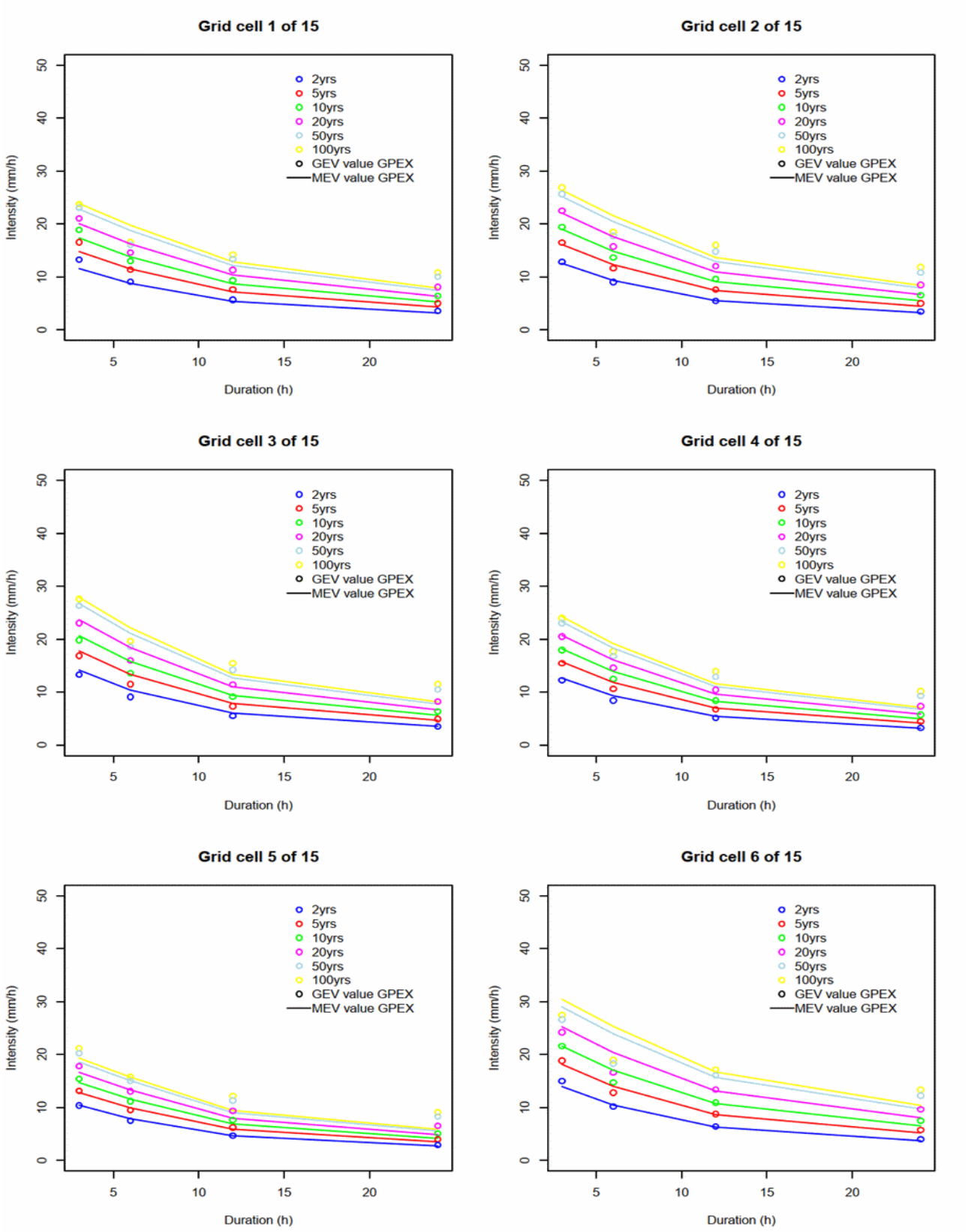


Figure 30: MEV estimates vs. GEV estimates (GPEX dataset) for Durations 3h, 6h, 12h and 24h and return periods T2, T5, T10, T50 and T100 (data source: Gruendemann et al., 2020)

2.3.5 ERA5-Land

The land component of the fifth generation of European ReAnalysis Data, short **ERA5-Land**, is provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) in Reading, UK. The dataset is described as

> *"a (climate) reanalysis dataset providing a consistent view of the evolution of land variables over several decades at an enhanced resolution compared to ERA5. (..) Reanalysis combines model data with observations from across the world into a globally complete and consistent dataset using the laws of physics. Reanalysis produces data that goes several decades back in time, providing an accurate description of the climate of the past."* ("ERA5-Land hourly data from 1981 to present," 2019, p. 5).

It is a high-resolution, multi-variable climate reanalysis dataset that provides information on various atmospheric, land-surface and cryosphere variables over land areas from the year 1950 to the near present. The data covers a wide range of variables including temperature, precipitation, humidity, wind speed and direction, surface pressure, soil moisture and more, providing a comprehensive view of the Earth's land-surface and atmospheric conditions. The data is available for scientific research as well as for operational weather forecasting and climate analysis and is widespread among the scientific community, especially Europe.

ERA5-Land does not assimilate observations directly. Each segment or stream is initialized with meteorological fields from ERA5. Air temperature, humidity, and pressure are corrected using a daily lapse rate derived from ERA5. Subsequently, the land surface model is integrated in 24h cycles providing the evolution of the land surface state and associated water and energy fluxes (Muñoz-Sabater et al., 2021).

Figure 31 hows a diagram of the algorithm used for each 24 h production cycle of the ERA5-Land dataset:

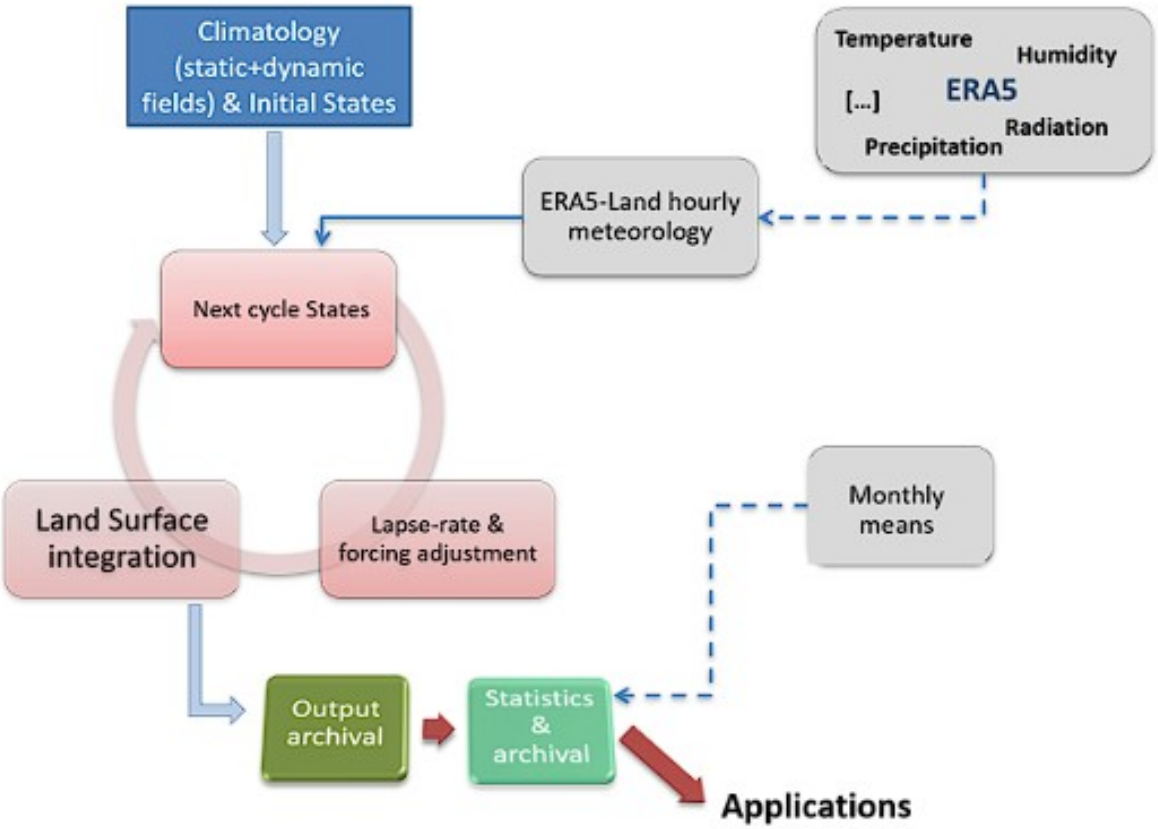


Figure 31: Environment of the ERA5-Land dataset (Muñoz-Sabater et al., 2021, p. 4352)

ERA5-Land produces a total of 50 variables describing the water and energy cycles over land. A main advantage of ERA5-Land compared to ERA5 and the preceding model ERA-Interim is the horizontal resolution, which is enhanced globally to 0.1 degrees (latitude: 11.1 km, longitude: 10.5 km at 18.5°N) compared to 30 km in ERA5 or 80 km in ERA-Interim, whereas the temporal resolution is hourly like in ERA5 (Muñoz-Sabater et al., 2021).

Table 3 gives a good overview on the characteristics of the three ERA datasets:

	ERA-Interim	ERA5	ERA5-Land
Period publicly available*	1979–Aug 2019	1950 onwards	1981 onwards (1950–1980, in 2021)
Spatial resolution	79 km/60 levels	31 km/137 levels	9 km
Land surface model	IFS (+ TESSEL)	IFS (+ CHTESSEL)	CHTESSEL
Model cycle (year)	Cy31r2 (2006)	Cy41r2 (2016)	Cy45r1 (2018)
Output frequency	6-hourly (analyses) 3-hourly (forecasts)	Hourly	Hourly
Uncertainty estimate	None	Based on a 10-member 4D-Var ensemble at 63 km	As for ERA5
Availability behind real time	n/a	2–3 months (final product) 5 d (preliminary product)	2–3 months (final product) 5 d (preliminary product, in 2021)

* Availability at the time of submitting this paper. n/a – not applicable.

Table 3: Key parameters of ERA-Interim, ERA5, and ERA5-Land (Muñoz-Sabater et al., 2021, p. 4356)

The parameter in use for the purpose of this work is called “total precipitation” (see Table 4) and is measuring

> *“Accumulated liquid and frozen water, including rain and snow, that falls to the Earth's surface. It is the sum of large-scale precipitation (that precipitation which is generated by large-scale weather patterns, such as troughs and cold fronts) and convective precipitation (generated by convection which occurs when air at lower levels in the atmosphere is warmer and less dense than the air above, so it rises). (..)* ***This variable is accumulated from the beginning of the forecast time to the end of the forecast step****. The units of precipitation are depth in metres. It is the depth the water would have if it were spread evenly over the grid box.* ***Care should be taken when comparing model variables with observations****, because observations are often local to a particular point in space and time, rather than representing averages over a model grid box and model time step.”* [39]

The exact citation is important in this case, as for me at the first glance, the hourly precipitation values were not self-explaining. Indeed, a closer look at the parameter description[40] had proven to be valuable: The “total precipitation” parameter is cumulative for 24 hours, with the maximum (of the day before) reached at 00:00 UTC. In other words: From 01:00 to 00:00 of the next day, the hourly precipitation sums are cumulative and if you are really interested in hourly rainfall intensities, you first have to deduct the precipitation sum of the time step before:

Dataset	Variables	What it is	Grib units	Netcdf Units	To get the total precipitation for an hour (mm)	To get the total precipitation for a day (mm)
ERA5-Land hourly data	Total precipitation	Accumulations are from 00 UTC to the hour ending at the forecast step [1]	m (of water equivalent)	m (of water equivalent)	$\mathrm{tp}\ [\mathrm{mm}] = \begin{cases} \mathrm{tp}_h\ [\mathrm{m}] \cdot 1000 & h = 01\mathrm{UTC} \\ (\mathrm{tp}_h\ [\mathrm{m}] - \mathrm{tp}_{h-1}\ [\mathrm{m}]) \cdot 1000 & \text{otherwise} \end{cases}$	$\mathrm{tp}\ [\mathrm{mm}] = \mathrm{tp}_{d+1\ 00\mathrm{UTC}}\ [\mathrm{m}] \cdot 1000$ where d is the day for which the average flux is being computed. The time step labelled d+1 00UTC should also be taken because it contains the accumulated flux over the previous 24 hours.

Table 4: Parameter description of ERA5-Land for "total precipitation" (https://confluence.ecmwf.int/pages/viewpage.action?pageId=197702790, access on 2023-03-21)

To explore the dataset and its characteristics, I chose once again the rain event of Hurricane Matthew on Oct 4th 2016 in Haiti. The ERA5-Land dataset does **not** cover the whole area of the Carrefour AOI, as only grid cells which fall entirely onto land mass are covered, whereas grid cells on the coastline or in maritime areas do not have data.

Extrapolating from the neighbouring grid cells in the east and south, ERA5-Land shows precipitation sums between 120 and 150 mm for the whole day[41] in our AOI of Carrefour (see Figure 32) and the Rivière Froide catchment, which is a bit higher than the values of CMORPH (85-120 mm for 24h, Figure 20) and MSWEP (85-105 mm for the whole day, Figure 24), but still lower than the stated 180–250 mm according to the Haitian dataset (see Figure 25). Accounting for the ARF (see Chapter 3.2.2) with a value of ~0.9 for D=24h and A=100 km², the values can be multiplied with 1.11, which leads to precipitation sums of 130-165 mm for the whole day in our AOI – not too far away from the 180-250 mm of the CNIGS/UHM dataset from Haiti.

39 https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.e2161bac?tab=overview (access on 21-03-2023)
40 https://confluence.ecmwf.int/pages/viewpage.action?pageId=197702790 (access on 21-03-2023)
41 Muñoz Sabater, J., (2019): ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (access on 04-01-2023), 10.24381/cds.e2161bac

Looking at the pattern of the 24h precipitation sums (Figure 32), the absolute maximum in the western region of the Tiburon peninsula, near centre of the hurricane track (see Figure 26), has significantly lower values in ERA5-Land (around 350 mm for 24 hours) than in the Haitian dataset (between 500 and 600 mm in 24h, see Figure 25) and in the MSWEP dataset (up to 800 mm in 24h, see Figure 24).

Taking into account the ARF with a value of ~0.9 (see Breinl et al., 2020, p. 680, Svensson and Jones, 2010, p. 5, Langousis, 2005, p. 12) for D=24h and A=100 km², the values can be multiplied with 1.11, which leads to a maximum precipitation sum of ~390 mm for the whole day in the Western Tiburon Highlands – still quite far away from the 500-600 mm of the CNIGS/UHM dataset or the 700-800 mm from MSWEP.

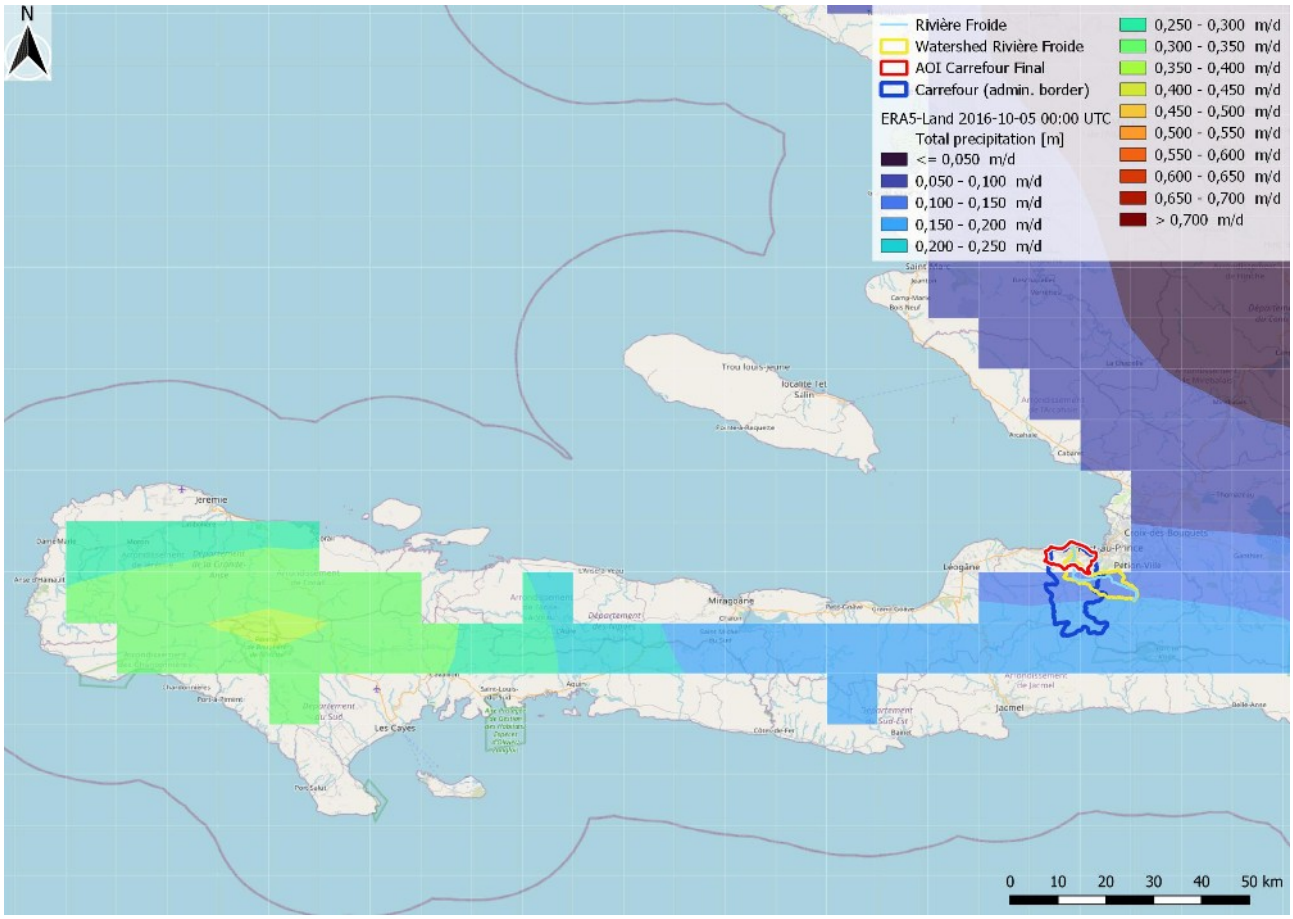


Figure 32: Copernicus ERA5-Land precipitation for 2016-10-05 00:00 UTC

2.3.6 NOAA Global Historical Climatology Network-Daily (GHCN-D)

The National Oceanic and Atmospheric Administration (NOAA) of the US Department of Commerce provides several precipitation datasets that can be useful for various applications. As data of the existing rain gauge network in Haiti is presently not available, the NOAA station network may serve as an alternative to acquire reliable precipitation data from a region nearby to Haiti – and compare it to the global remote sensing datasets mentioned before.

For all US States and territories (therefore including the neighbouring island of Puerto Rico), there are even point precipitation frequency estimates available, and this will play an important part in the model validation in chapter 3.2.3.

The **Global Historical Climatology Network – Daily** (**GHCN-D**, Menne et al. 2012) dataset integrates daily climate observations from approximately 30 different data sources. The current version 3 was released in September 2012 and contains station-based measurements from well over 90,000 land-based stations worldwide, about 2/3 of which are for precipitation measurement only.

> *"GHCN-Daily (..) will function as the official archive for daily data from the Global Climate Observing System (GCOS) Surface Network (GSN) and i**s particularly well suited for monitoring and assessment activities related to the frequency and magnitude of extremes**. The total of 1.4 billion data values includes 250 million values each for maximum and minimum temperatures, 500 million precipitation totals, and 200 million observations each for snowfall and snow depth. Station records, some of which extend back to the 19th century, are updated daily where possible and are usually available one to two days after the date and time of the observation."*[42]

The data can be downloaded at no cost either from the NCEI Climate Data Online Portal[43] or directly from the HTTP Server[44]. For Haiti, there is no data available, but for the neighbouring state of the Dominican Republic. The nearest rain gauge to the AOI in Carrefour is the station DRM00078482 MariaMontezIntl – the airport of Santa Cruz de Barahona, 140km in easterly direction of Carrefour.

For Hurricane Matthew (04-10-2016), both CMORPH and ERA5-Land have precipitation sums between 90 and 100 mm there (see Figure 33), whereas MSWEP has a rainfall sum of 60 mm. The GHCN-D dataset (Menne et al., 2012a) shows for the day after (05-10-2016) the precipitation sum of 263.9 mm (see Table 5), which is at least more than double the value of all remote sensing datasets. **So even when accounting the ARF**, which can be assumed with 0.9-0.95 for 24h precipitation values, see Chapter 3.2.2, **the point precipitation data from this trusted in-situ source is double than triple the values of the available remote sensing datasets**.

	A	B	C	D	E	F	G
1	STATION	NAME	LATITUDE	LONGITUDE	ELEVATION	DATE	PRCP
10891	DRM00078482	MARIA MONTEZ INTERNATIONAL, DR	18.251	-71.120	3.0	2016-10-02	
10892	DRM00078482	MARIA MONTEZ INTERNATIONAL, DR	18.251	-71.120	3.0	2016-10-03	12.4
10893	DRM00078482	MARIA MONTEZ INTERNATIONAL, DR	18.251	-71.120	3.0	2016-10-04	0.0
10894	DRM00078482	MARIA MONTEZ INTERNATIONAL, DR	18.251	-71.120	3.0	2016-10-05	263.9
10895	DRM00078482	MARIA MONTEZ INTERNATIONAL, DR	18.251	-71.120	3.0	2016-10-06	58.7
10896	DRM00078482	MARIA MONTEZ INTERNATIONAL, DR	18.251	-71.120	3.0	2016-10-07	0.5
10897	DRM00078482	MARIA MONTEZ INTERNATIONAL, DR	18.251	-71.120	3.0	2016-10-08	0.0
10898	DRM00078482	MARIA MONTEZ INTERNATIONAL, DR	18.251	-71.120	3.0	2016-10-09	

Table 5: GHCN-D dataset for station DRM00078482 Maria Montez International Airport, DR (data source: https://www.ncei.noaa.gov/pub/data/ghcn/daily/by_station/)

42 https://www.ncei.noaa.gov/pub/data/cdo/documentation/GHCND_documentation.pdf (access on 2023-03-25)
43 https://www.ncei.noaa.gov/cdo-web/search?datasetid=GHCND (access on 2023-03-25)
44 https://www.ncei.noaa.gov/pub/data/ghcn/daily/ (access on 2023-03-25)

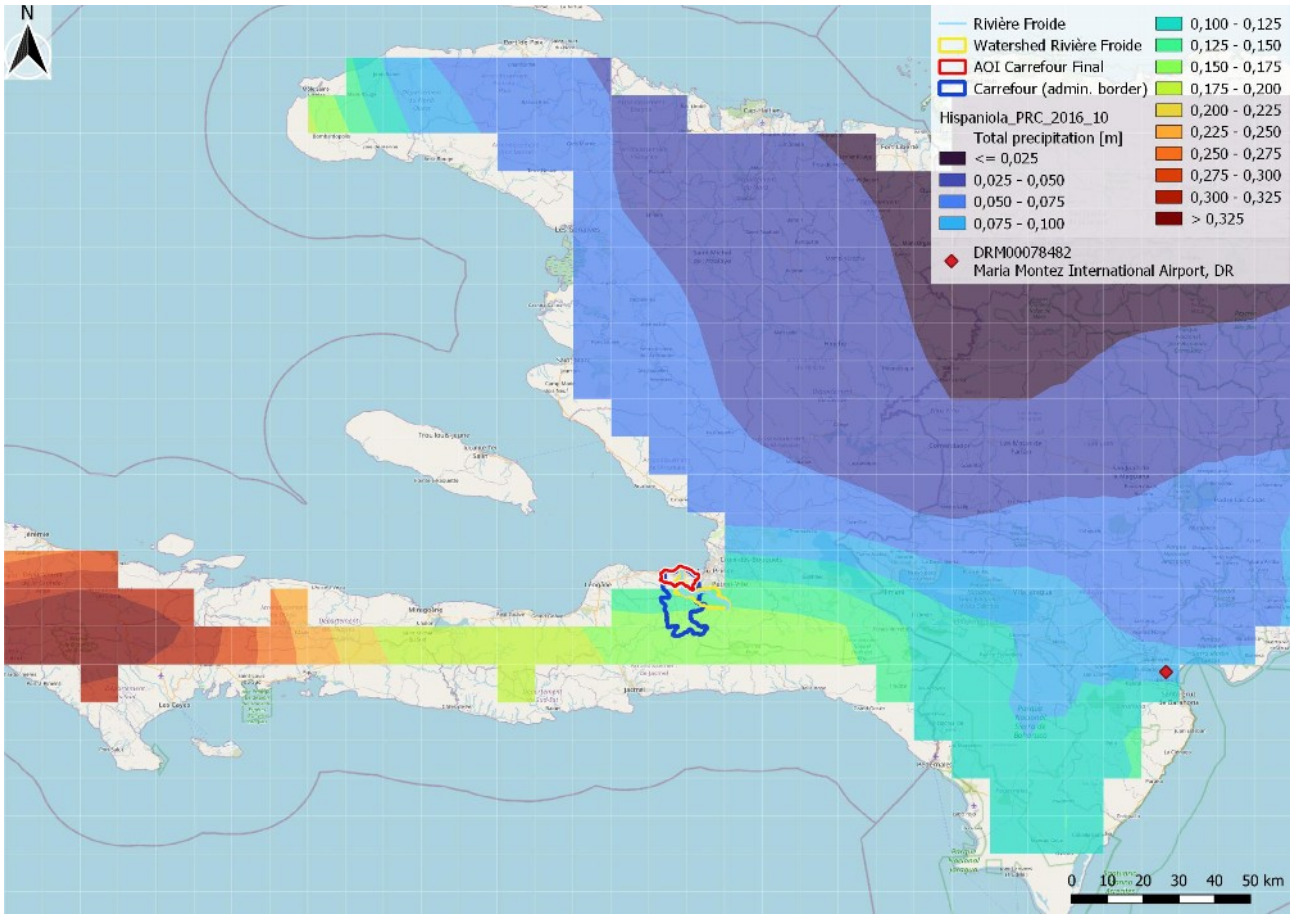


Figure 33: ERA5-Land 24h precipitation in m for 04-10-2016 (red dot = station DRM00078482)

3. Methods and concepts

3.1. Data scarcity and the consequences

As it has been particularised in the last main chapter as well as in the introduction, the issue of input data scarcity is a crucial factor for flood modelling in our area of interest. But the pluvial flood simulation requires knowledge of how much rain is likely to fall within a certain amount of time and over a specific area.

One pragmatic solution in this case could be the ignorance of the missing station-based data and either just rely on the remote sensing datasets or simply run a set of different scenarios for a precipitation event P during time period (duration) D, regardless of a classification to a certain return period T. In fact, this approach can be argued as long as there is any station-based data available in the area, to which the different rainfall scenarios can be compared with – making it an inductive approach to what is potentially going to happen if a rainfall event with precipitation amount P will reoccur in the future.

Nevertheless, there are ways and possibilities to assess a certain return period T to a design event with rainfall intensity P/D, to stick to the variables assigned above. As in-situ station data was not available for Haiti, the chosen approach was to compare the remote sensing datasets with station data and point precipitation frequency charts from the south coast of Puerto Rico, where conditions should be somewhat similar to our area of interest – not too far away (about 600km beeline), with a similar topography (coastal area with quickly rising terrain in the back country).

3.2. Model validation

3.2.1 Remote sensing datasets vs. station-based datasets: ERA5-Land and MSWEP v2 vs. GHCN-D

As already thematically mentioned in the previous chapter 2.3, the data comparison between station based ground-truth data and remote sensing gridded dataset is not as trivial as maybe expected, not only due to different data formats, but also due to different areas of reference (point data vs. area average) and other factors, which lie in the specific methodology of each remote sensing dataset.

To be able to decide which dataset will fulfil the aim of providing the most reliable data source for this work, I take a closer look at the comparison between station based data (on the example of the NOAA daily precipitation dataset GHCN-D) and two remote sensing datasets (namely ERA5-Land and MSWEP v2).

On January 6th 1992, there had been a major pluvial flood event in the Southern and central parts of Puerto Rico, according to a report of the USGS:

> *"Severe flooding affected at least 40 municipalities in Puerto Rico during January 5-6, 1992. A combined weather system, consisting of a cold front and an upper level trough,*

> *produced 24-hour total rainfall of as much as 20 inches on the island's mountainous interior. Rainfall intensities for 1-, 2-, 3-, and 6-hour durations exceeded previously recorded island-wide maximum values." (Torres-Sierra, 1996:1)*

20 inches (around 500 mm) of rain in 24 hours have been measured in the island's interior, at the stations "Toro Negro Forest" and "Cayey 1E". But also on the coastline, at station "Ponce 4E", there had been measured 8.30 inches (210 mm) in 24 hours. The idea is now to juxtapose the precipitation amounts of the USGS report of 1996 with the online data from GHCN-D, the data from ERA5-Land and from MSWEP.

The first check, internal sync between the 1996 USGS report (Figure 34) and the NOAA GHCN-D online data archive[45] could be completed successfully: The stations in question could be easily selected in the online portal and downloaded as a single CSV dataset, which afterwards had been converted into a point shape-file for better visual comparison, with daily precipitation sums as well as the calculated cumulative rainfall sum for both days.

Based on the cumulative rainfall for both days, contour lines with a virtual raster resolution of 0.01° (~1km) and 2 inch equidistance had been created in the GI-Software "SAGA-GIS" via the interpolation algorithm "Multilevel Bspline"[46]. The precipitation pattern created from the contour lines (see Figure 35)looks for the most part quite similar to Figure 34: two maxima in the mountains near the Southern coast and low precipitation sums in the Northern part, especially in the Northwest.

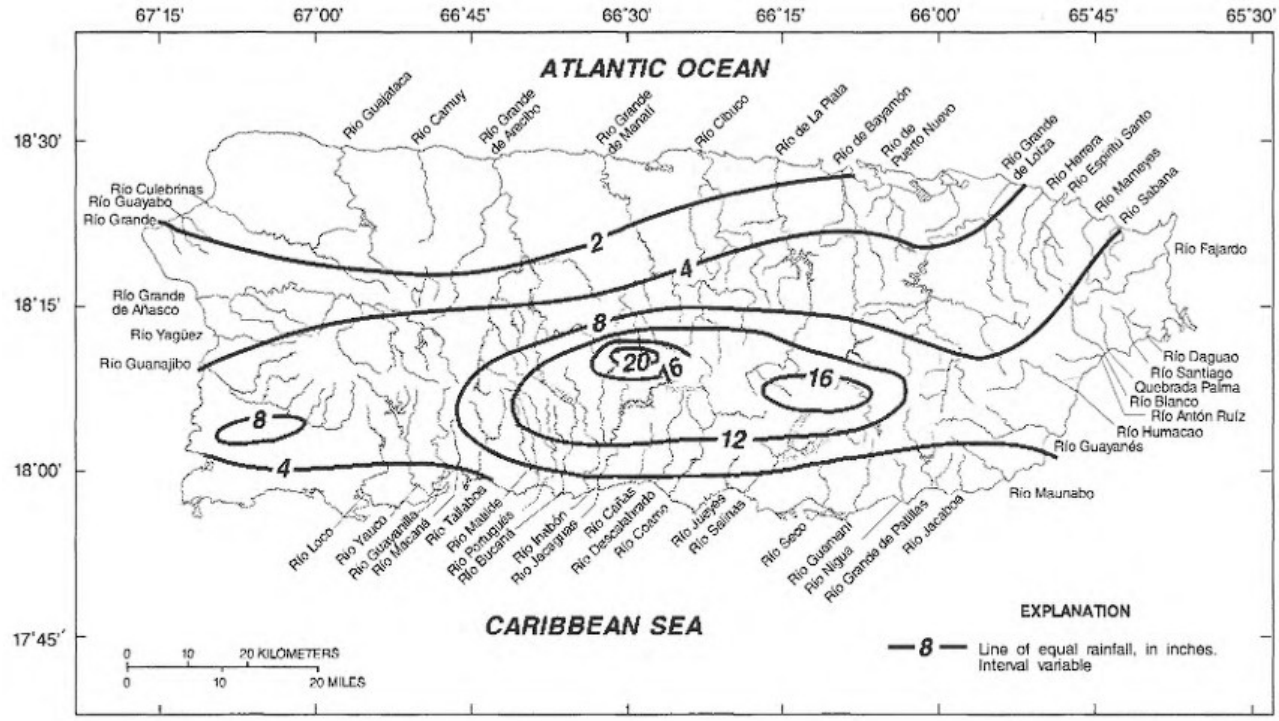


Figure 34: Isolines: Cumulative precip. over PRC Jan 5/6, 1992 (data compiled by NOAA, (Torres-Sierra, 1996, p. 12)

45 https://www.ncei.noaa.gov/maps/daily/ (access on 2023-03-26)
46 https://saga-gis.sourceforge.io/saga_tool_doc/2.2.5/grid_spline_4.html (access on 2023-04-06)

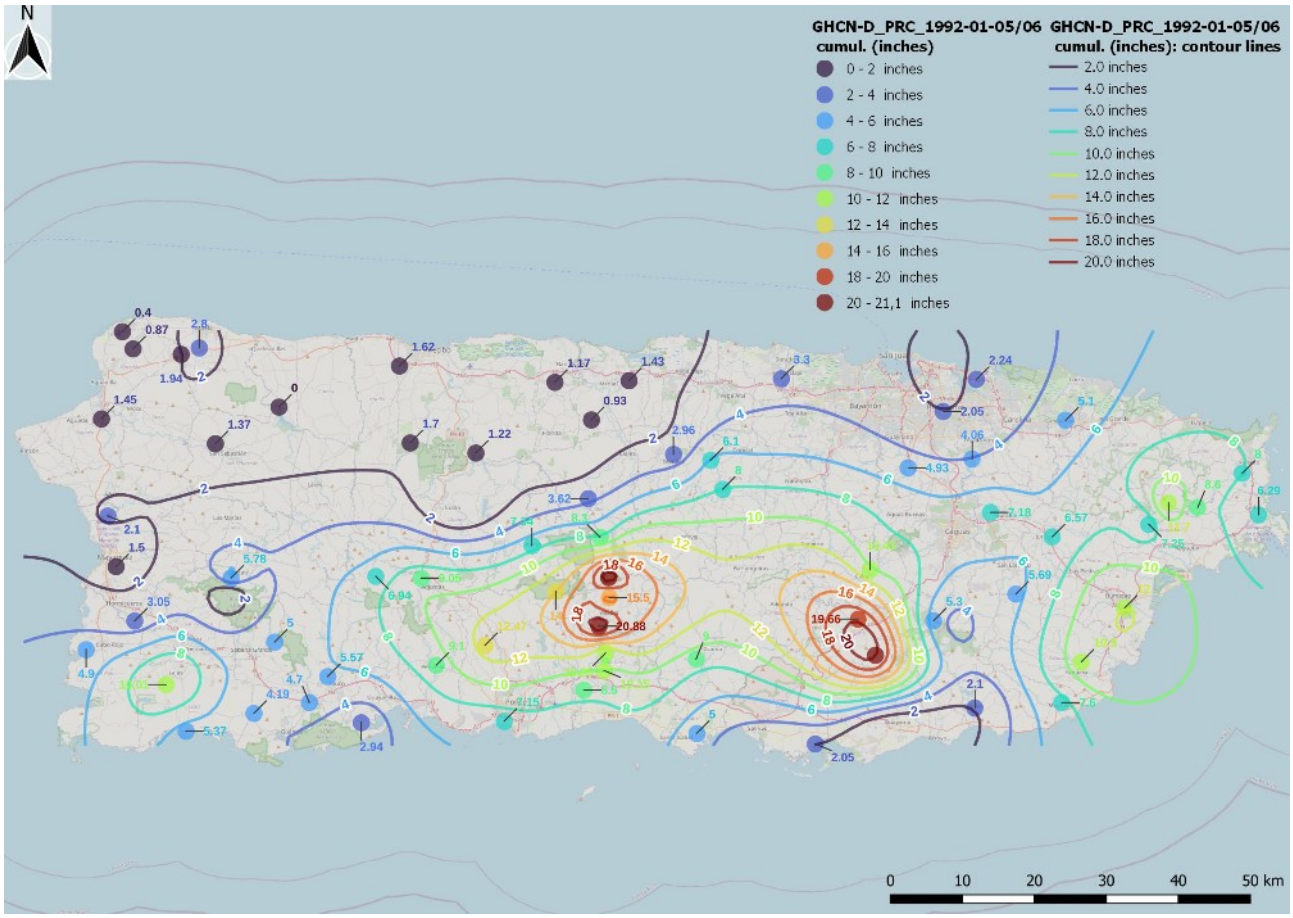


Figure 35: Isolines and station data: Cumulative precip. over PRC Jan 5/6, 1992 (data source: https://www.ncei.noaa.gov/maps/daily/ - access on 2023-03-26)

It must be pointed out that NOAA daily station data seems to measured not from 00:00 to 24:00, but from 08:00 AM day 1 to 08:00 AM day 2. This makes the comparison with satellite based datasets more difficult, as some of the remote sensing precipitation data may fall in the day before / day after compared to the station data and vice versa – but the sum of the 2 days (1992-01-05 and 1992-01-06) can be compared.

The comparison with **ERA5-Land** turned out to be **completely unsuccessful**: The rainfall event simply does not exist as such in ERA5 datasets, neither in ERA5, nor in ERA5-Land (see Figure 37), whatever the reason may be. The maximum 24h sum is only 40 mm in case of ERA5-Land and about the same value in the ERA5-reanalysis dataset, which is **by no means even close to being in the range, where it should be** according to the rain gauge data. Of course, this may be bad luck in this case.

MSWEP v2 (Figure 38) on the other hand shows **both a plausible pattern** which is **well in sync** by and large **with** the interpolated rainfall sums based on **GHCN-D** (Figure 35) respectively the USGS report (Figure 34) **and precipitation amounts which are mostly not too far away from the rain gauge sums**, in the case of station RQC00667292 Ponce 4E even a match: Ponce 4E (some kilometres in NE direction from the city of Ponce) had measured 211 mm of daily rainfall on Jan 6th 1992 and 5 mm of precipitation on Jan 5th, summarizing to 216 mm in total for both days. MSWEP daily precipitation sum on Jan 6th 1992 is 219 mm for the same area.

At the precipitation maxima in the mountains, for the central region around stations RQC00660040 ACEITUNA WATER TREATMENT PLANT (393.7 mm for both days, 100% of it on Jan 6th) and RQC00662336 CERRO MARAVILLA (355.6 mm for both days, thereof 304.8 mm on Jan 6th), MSWEP v2 delivers a (linear interpolated) daily precipitation sum of 275 mm for station ACEITUNA and 180 mm for station CERRO MARAVILLA. Regarding an ARF of 0.9 for a 24h duration and a dataset resolution of ~100 km², the by ARF 0.9 reduced GHCN-D station data (355 mm for ACEITUNA and 275 mm for CERRO MARAVILLA) is **factor 1.3** the MSWEP v2 data in case of station ACEITUNA and **factor 1.5** in case of station CERRO MARAVILLA.

By exporting the vertices of the GRIB mesh in QGIS and comparing them via the "Nearest Neighbour" algorithm to the station data, it is possible to analyse the differences of the GHCN-D station dataset and the MSWEP v2 dataset in a more systematic way:

Figure 39 shows the delta between the daily precipitation sum (Jan 6th 1992) of MSWEP v2 dataset vertices (resolution = 0.1 degrees) and the rainfall amounts measured in situ by the GHCN-D station network – reduced by ARF = 0.9 (100 km² / 24h duration, see Chapter 3.2.2). **In the arithmetic average, the MSWEP dataset precipitation is about 80% of the reduced (ARF 0.9) GHCN-D station data** (see Figure 40) which seems like quite a good value for me.

t

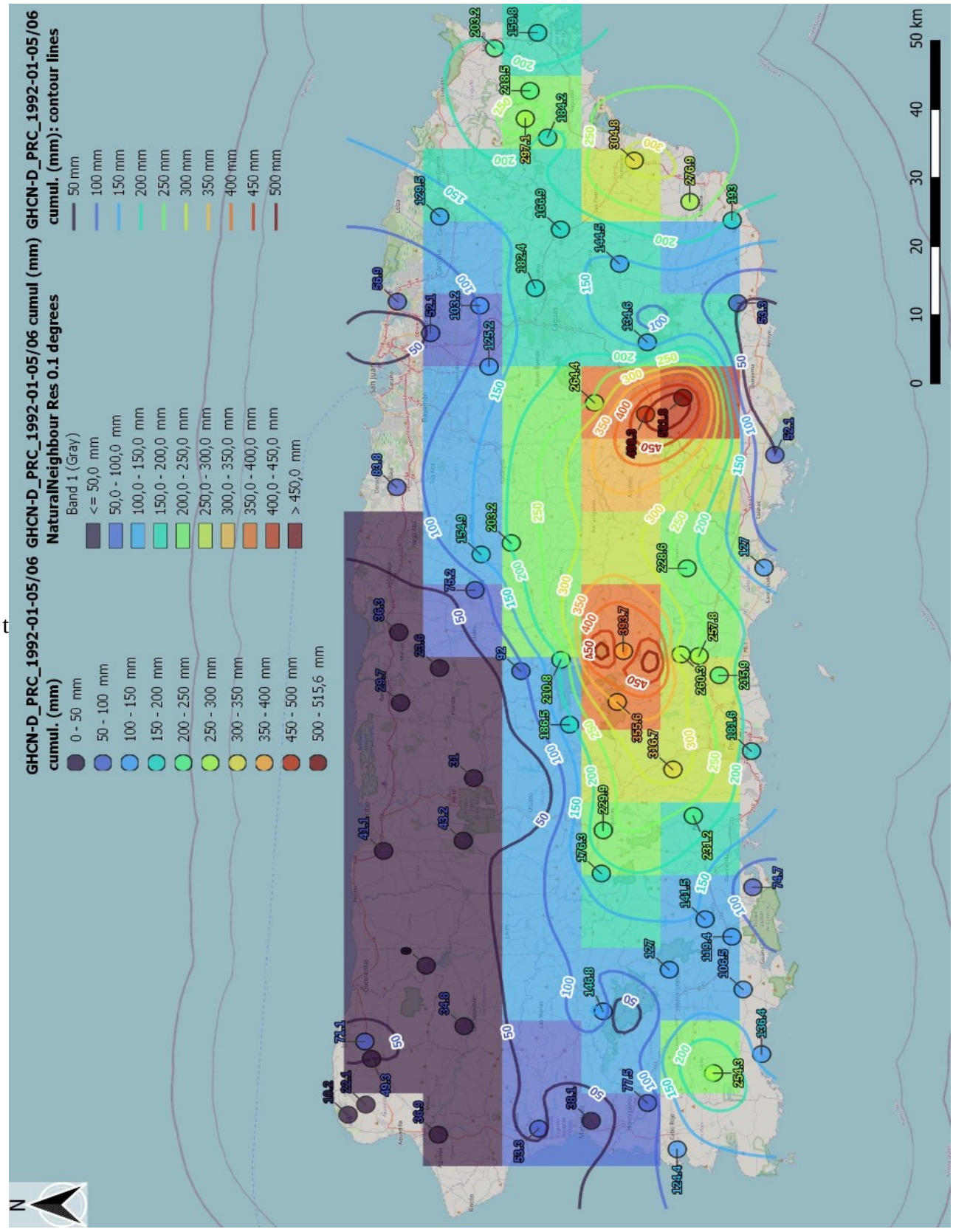


Figure 36: Natural Neighbour interpolation (0.1 degrees), isolines and station data: Cumulative precip. (mm) over PRC Jan 5/6, 1992

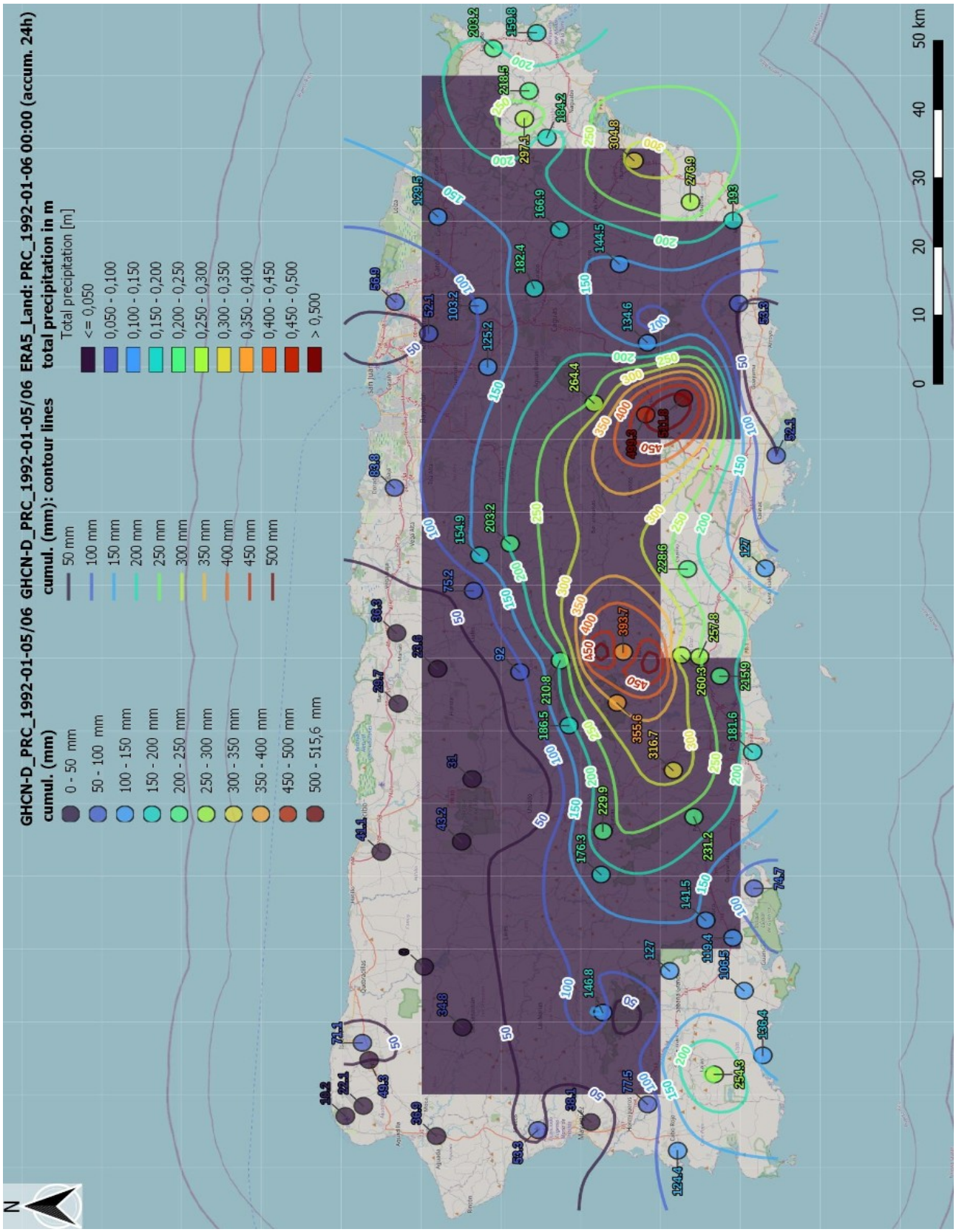


Figure 37: ERA5-Land Total precipitation (m), 0.1 degrees, isolines and station data: Cumulative precip. (mm) over PRC Jan 5/6, 1992

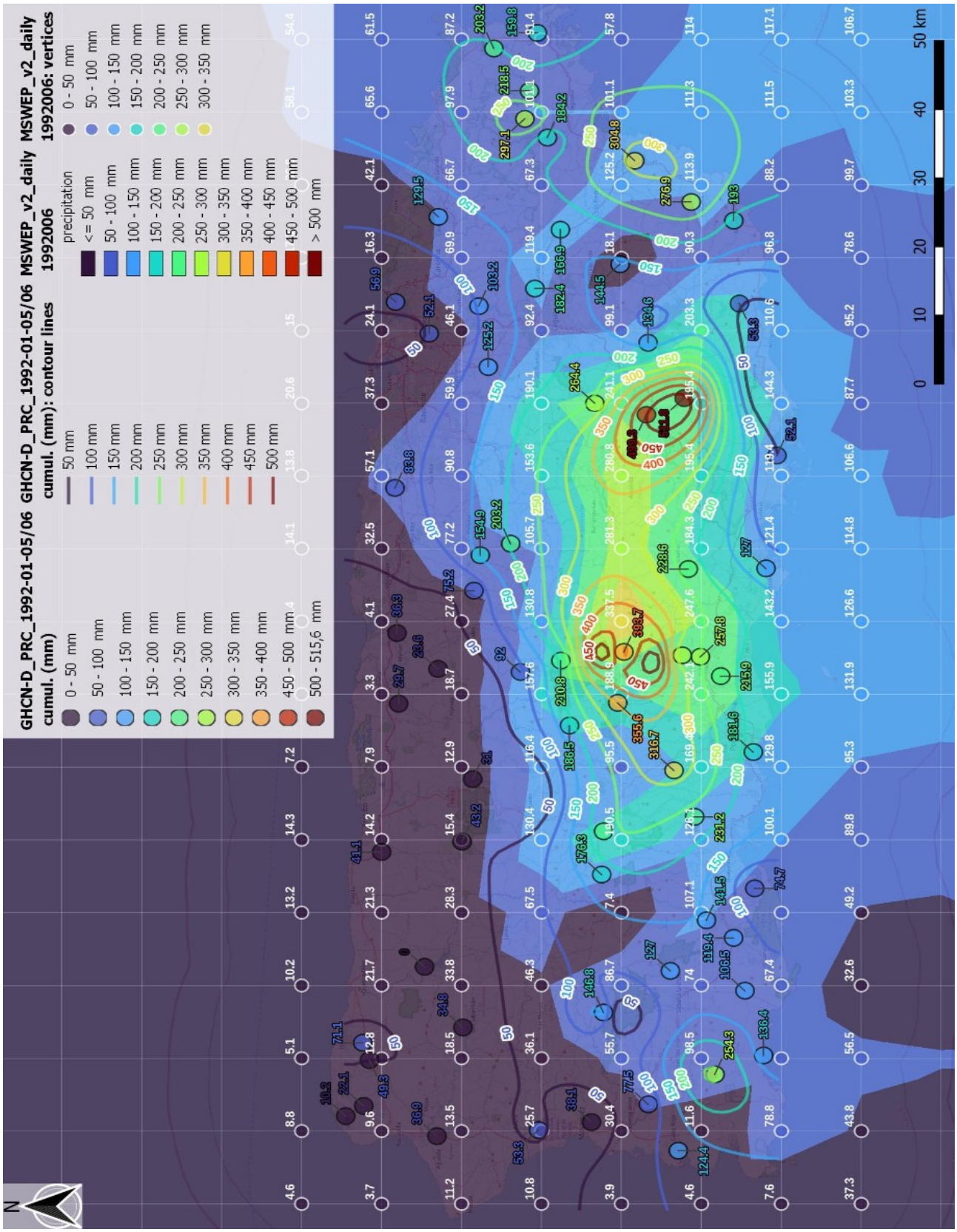


Figure 38: MSWEP v2: Daily precipitation sum in mm on Jan 6 1992 (vertices plus linear interpolation), GHCN-D isolines and station data: Cumulative precip. (mm) over PRC Jan 5/6, 1992

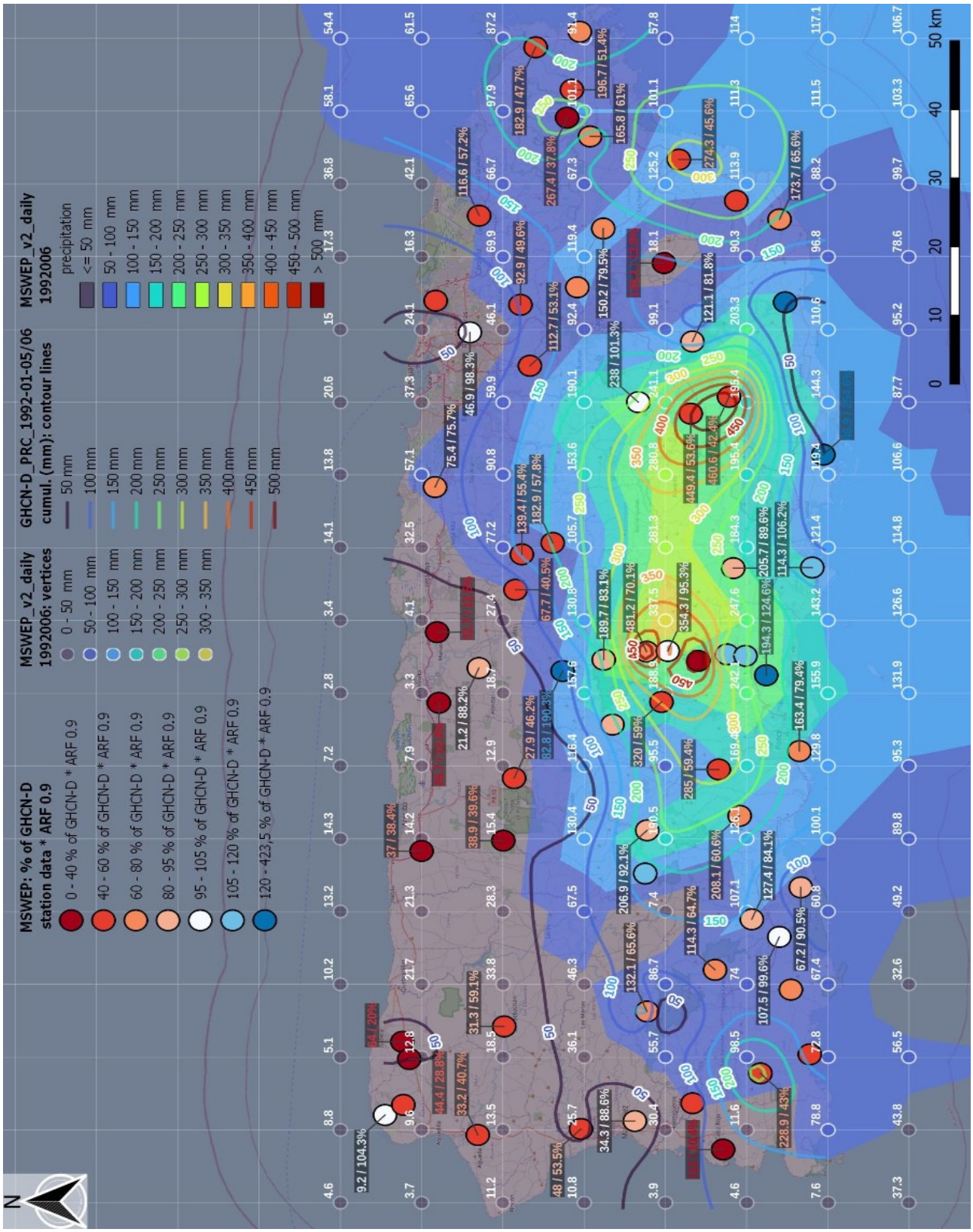


Figure 39: MSWEP v2 versus GHCN-D: Nearest MSWEP v2 vertex to GHCN-D rain gauge (big dots): % of reduced precipitation sum (value MSWEP v2 = x % of station value GHCN-D * 0.9)

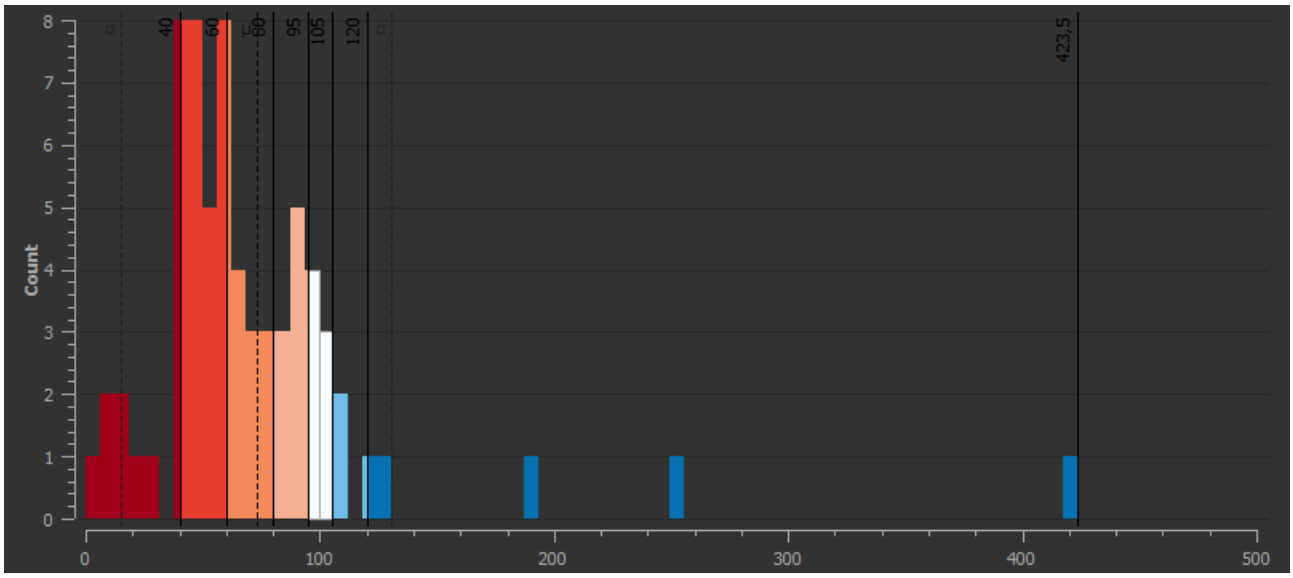


Figure 40: Histogram of the divergences of MSWEP v2 to reduced GHCN-D station data (nearest MSWEP vertex to station) in % of GHCN-D

3.2.2 Point precipitation and area reduction factor (ARF)

The Areal Reduction Factor (ARF) is a key parameter in the design for hydrological extremes. Point rainfalls are only representative for a limited area, and for larger areas, the area-average rainfall amount is likely to be smaller than at the point of the maximum observed depth. The estimation of areal reduction factors is concerned with the relationship between the point precipitation and the area-average rainfall (Svensson and Jones, 2010).

The use of ARFs is convenient as networks of rain gauges with long series, which are needed for accurate rainfall frequency estimation, are generally sparse (not only in Haiti) and hence often do not allow for an appropriate characterisation of the associated spatial rainfall patterns.

Generally spoken, the ARF can be defined as follows:

> *"For a basin of area A, the ARF is the ratio between the area-average rainfall intensity over a duration D with return period T and the point rainfall intensity for the same D and T." (Langousis, 2005)*

In other words: High ARF values show a rather small difference between point precipitation and the area-average rainfall and vice versa. It is important to note that there is no uniform method for the calculation of the ARF.

The ARF has been found to vary with

- the concerned area
- the duration
- the return period

- the dominant terrain type (mountains vs. flat areas)
- the predominant weather type and seasonality

Traditionally, ARF estimates are based on empirical methods (US Weather Bureau method, annual maxima centred method etc.), but more recently, a range of analytical methods have been applied.

Figure 41 and Figure 42 show different empirical approaches of defining the ARF:

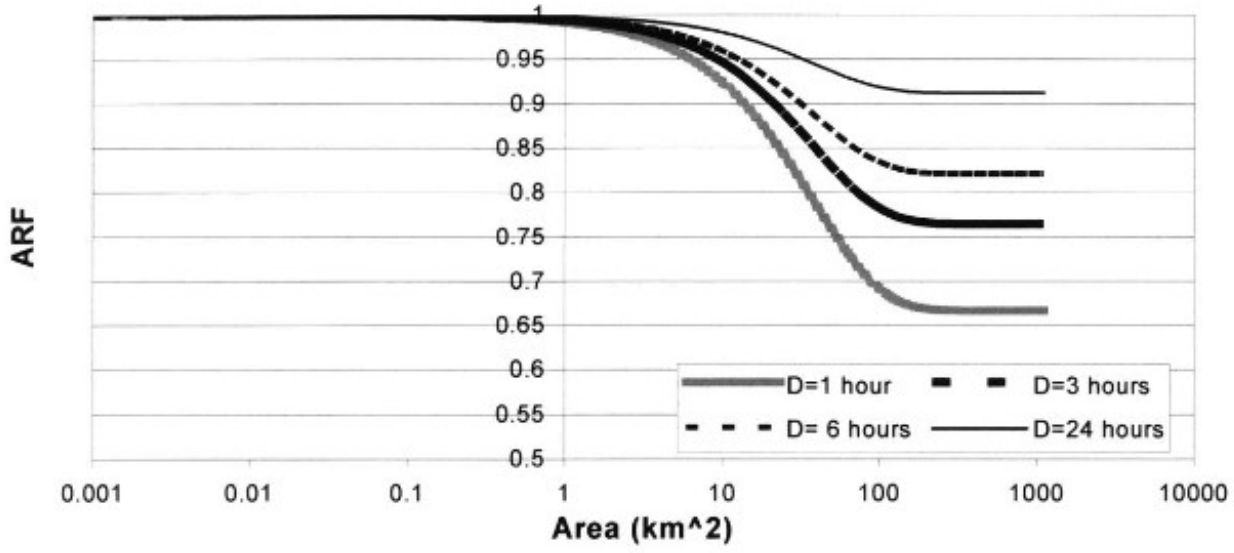


Figure 41: ARF curves (Leclerc and Schaake, “US Weather Bureau method”) for durations of 1, 3, 6, and 24 hours and areas up to 10,000 km² (Langousis, 2005, p. 12)

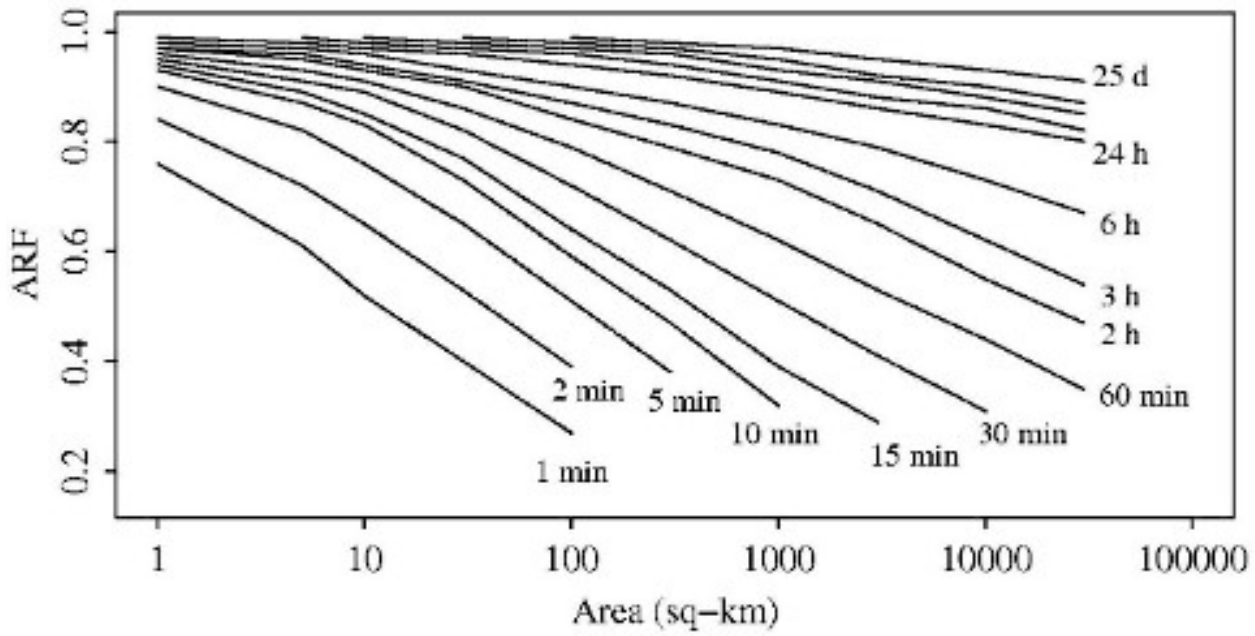


Figure 42: ARFs for precipitation in the UK presented in the Flood Studies Report NERC 1975 (Svensson and Jones, 2010, p. 4)

Concerning the last factor mentioned above, BREINL, MÜLLER-THOMY and BLÖSCHL point out in their paper (2020) that the prevailing precipitation type (strati-form vs. convective) has significant impact on the ARF as well:

> *"Results indicate that the ARFs decay faster with area in regions of **increased convective activity** than in regions dominated by stratiform processes. Low ARF values occur where and when lightning activity (as a proxy for convective activity) is high, but some areas with reduced lightning activity exhibit also rather low ARFs as, in summer, convective rainfall can occur in any part of the country. **ARFs tend to decrease with increasing return period**, possibly because the contribution of convective rainfall is higher. The **results of this study are consistent with similar studies in humid climates** and provide new insights regarding the relationship of ARFs and dominating rainfall processes." (Breinl et al., 2020)*

Via a complex geo-statistical procedure (variogram modelling and block kriging), BREINL et al. show at the example of Austria that the ARF may also be determined from seasonality (more convective precipitation in the summer leads to lower ARF values, which is a higher delta between point precipitation and area-average rainfall) as well as from (micro-) climatic aspects (areas with dominating stratiform rainfall, like the Bregenzer Wald region in Vorarlberg, tend to have higher ARF values – so a lower delta between point precipitation and area-average rainfall – than areas with high lightning activity like parts of Styria (see Figure 43).

But as different and inhomogeneous all these approaches may be, it is possible to draw two essential conclusions which are extremely important for this work:

- For the **duration of 24h and an area of ~100 square kilometres** (approx. the cell size of the remote sensing rainfall datasets MSWEP and ERA5-Land) where the comparison between station based (point precipitation) data in the form of GHCN-D and area-average rainfall in the form of ERA5-Land and MSWEP has been made, all methods of ARF calculation show similar values, between 0.9 and 0.95, regardless of all other parameters like return period and seasonality. **In other words, the ARF is not the crucial parameter to explain the observed difference here**.
- For the **duration of 1h, the situation is different**: For the same area (100 km²), the ARF varies between 0.7 (US Weather bureau method), 0.8 (UK Flood Studies Report) and 0.75-0.90 (Breinl et al., 2020:680, depending on return period, region and seasonality). This means that vice versa, **the precipitation amounts of the hourly ERA5-Land dataset can be multiplied by 1.25 on average to account for the ARF.** But again, also for the 1h duration, the ARF is not able to explain the observed difference between ERA5-Land and in situ station-based data.

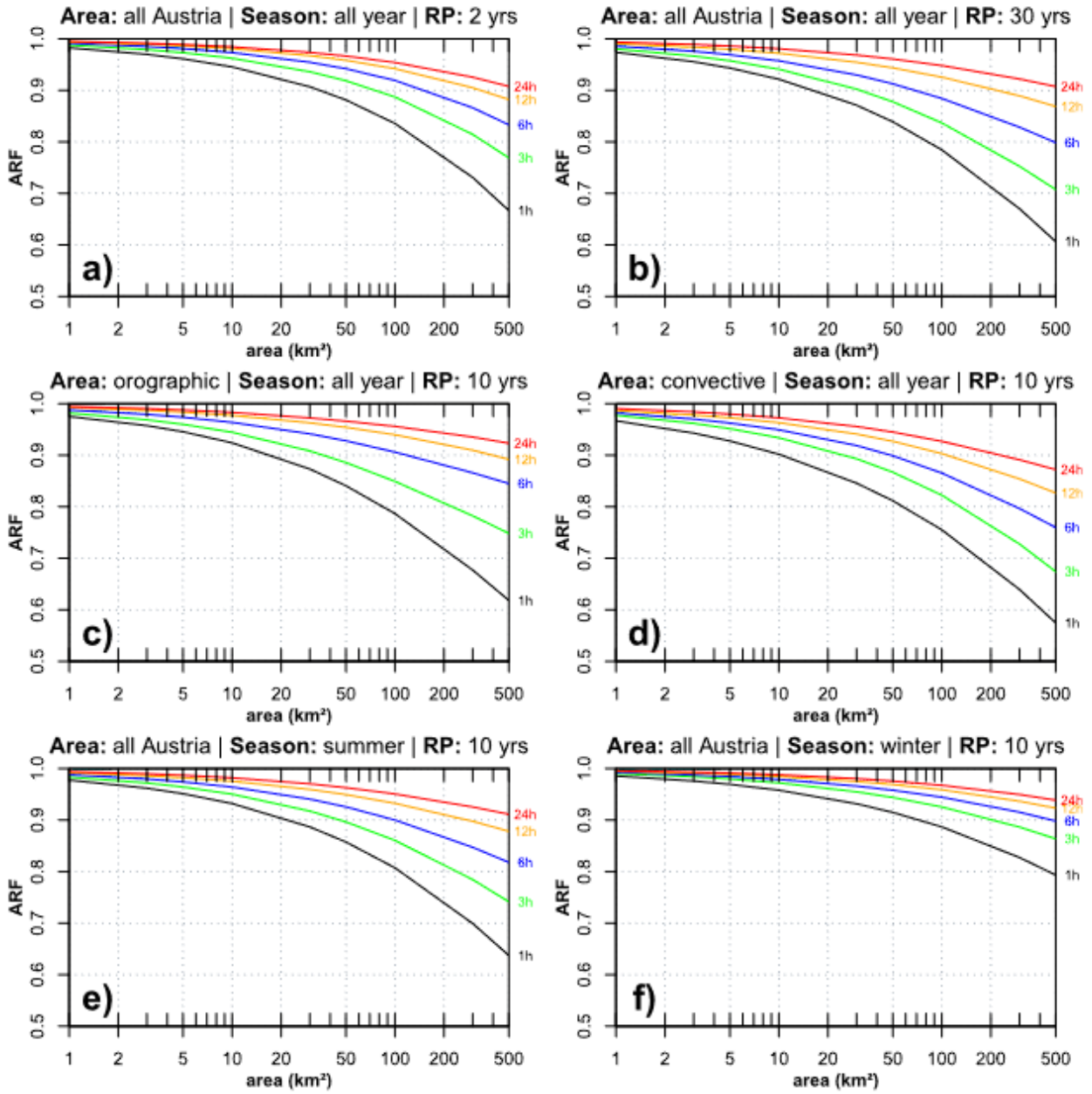


Figure 43: ARFs for different return periods, seasons and regions. Comparisons are shown for (a),(b) two return periods, (c),(d) two regions, and (e),(f) summer and winter (Breinl et al., 2020, p. 680)

3.2.3 Rainfall statistics (IDF Curves)

IDF (Intensity-Duration-Frequency) Curves are an important concept in hydrology, being one of the most commonly used tools in water resources management, flood calculation, civil engineering and water resource planning. They are a graphical representation of the relationship between rainfall intensity I, duration d, and frequency (= return period) T for a specific location: They are typically plotted with rainfall intensity on the y-axis and duration on the x-axis (see Figure 44).

IDF curves are used to quantify the probability of a particular rainfall event occurring at a specific location and are often used in the design and analysis of water resource infrastructure such as storm water systems, drainage systems, and flood protection structures (Koutsoyiannis et al., 1998).

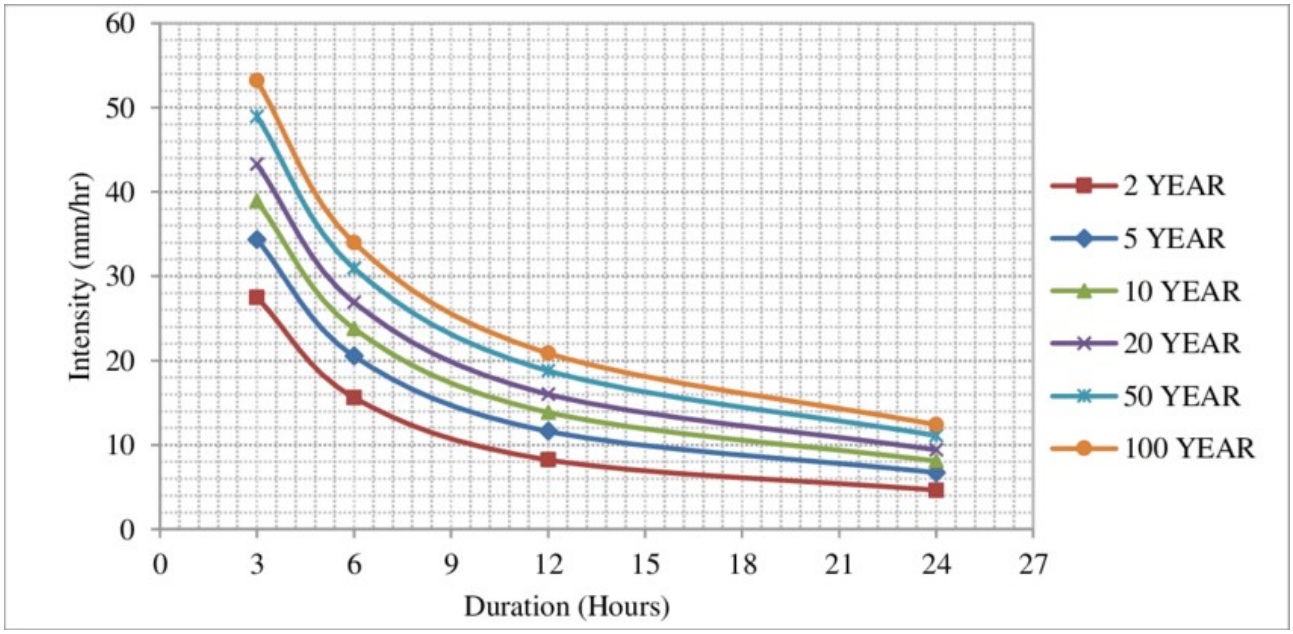


Figure 44: Example for a typical graphical representation of IDF Curves (here: present IDF curve for Bangkok metropolis station using annual maximum series from 3 hour rainfall depth from 1981 – 2010, (Shrestha, 2013, p. 30)

Derivation of IDF curves generally relies on the availability of long-term historical precipitation data and consists in fitting an extreme values distribution (mostly either the Generalized Extreme Value distribution, short GEV, or the Gumbel Distribution) to extreme rainfall values. Rain gauges, providing long records of precipitation data, are traditionally used to estimate IDF curves at the gauge locations. Nevertheless, the representativeness of gauge-derived IDFs decreases moving away from the gauge location and no or very sparse information is available for the many ungauged locations of the Earth. (Marra et al., 2017)

Nevertheless, owing to the sparseness of gauge networks worldwide (especially in developing countries), this approach raises important issues when design applications either require information at local scale or on basis of short to very short rainfall durations, or in both cases.

In the case of this work, both factors play an important role: For the AOI of Carrefour, we have data scarcity

- **in space** (no rain gauge in Carrefour, other existing Haitian gauge data currently not available)
- **in time** (most remote sensing rainfall datasets measure daily or 3 hourly – with the exemption of ERA5-Land, which offers an hourly precipitation value)

As a consequence, we are stuck between a rock and a hard place in case of our AOI: Either we are choosing available but distant rain gauge data (in this case this would be the Dominican Republic or the US Oversea territory of Puerto Rico), or we opt for IDF analysis built on remote sensing

datasets only – which differ significantly in most cases from in-situ data sources, as had been elaborated in chapter 3.2.1 and only one dataset has hourly resolution (ERA5-Land).

In order to find the best compromise for this work, I chose to **compare the (MSWEP-based) GPEX dataset** (which explicitly provides information on IDF of extreme precipitation events) to **IDF curves available from NOAA** for the island of Puerto Rico, of two different grid cells at the South coast around the city of Ponce, which has a very similar topography (quickly rising hills and mountains after the shoreline).

GPEX is strongly based upon the MSWEP dataset, so it has the same spatial resolution of 0.1 arc degrees. All IDF curve information is based on the precipitation amounts calculated by MSWEP (38 years of data availability, so n=38). The GPEX data (Gruendemann et al., 2020) has been extracted via the same R script (see Appendix 6.2) which has been used already for the comparison of the Haitian CNIGS/UHM dataset of the precipitation during Hurricane Matthew in the year 2016, only adapted for a different region.

To get an overview of the GPEX dataset in the Ponce region, I extracted 10 grid cells parallel the coast line (see Figure 45) and compared them a) with the station rainfall sums of the January 1992 rain event (to get an idea about the local precipitation regime in case of a heavy rain event) and b) with the point precipitation frequency estimates[47] (Bonnin et al., 2008) from the NOAA HDSC dataset[48].

The **Hydro-meteorological Design Studies Center (HDSC)** of the US National Weather Service NOAA offers detailed point precipitation frequency estimates from all rain gauges nationwide. They are built of the same database like GHNC-D, but with the difference that hourly and sub-hourly durations (until the 5-min interval) are included as well. This matters in the case of this work, as, like mentioned before, the relevant rainfall duration for AOI is one hour maximum, presumably below that (see Chapter 3.4.2). So IDF statistics for 3 hours and upwards (like GPEX offers) alone will not really help us in this case. They must be extrapolated and compared with another data source, this will be the HDSC point precipitation frequency estimates.

47 https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_pr.html (access on 2023-04-10)
48 https://hdsc.nws.noaa.gov/pub/hdsc/data/pr/ (access on 2023-04-10)

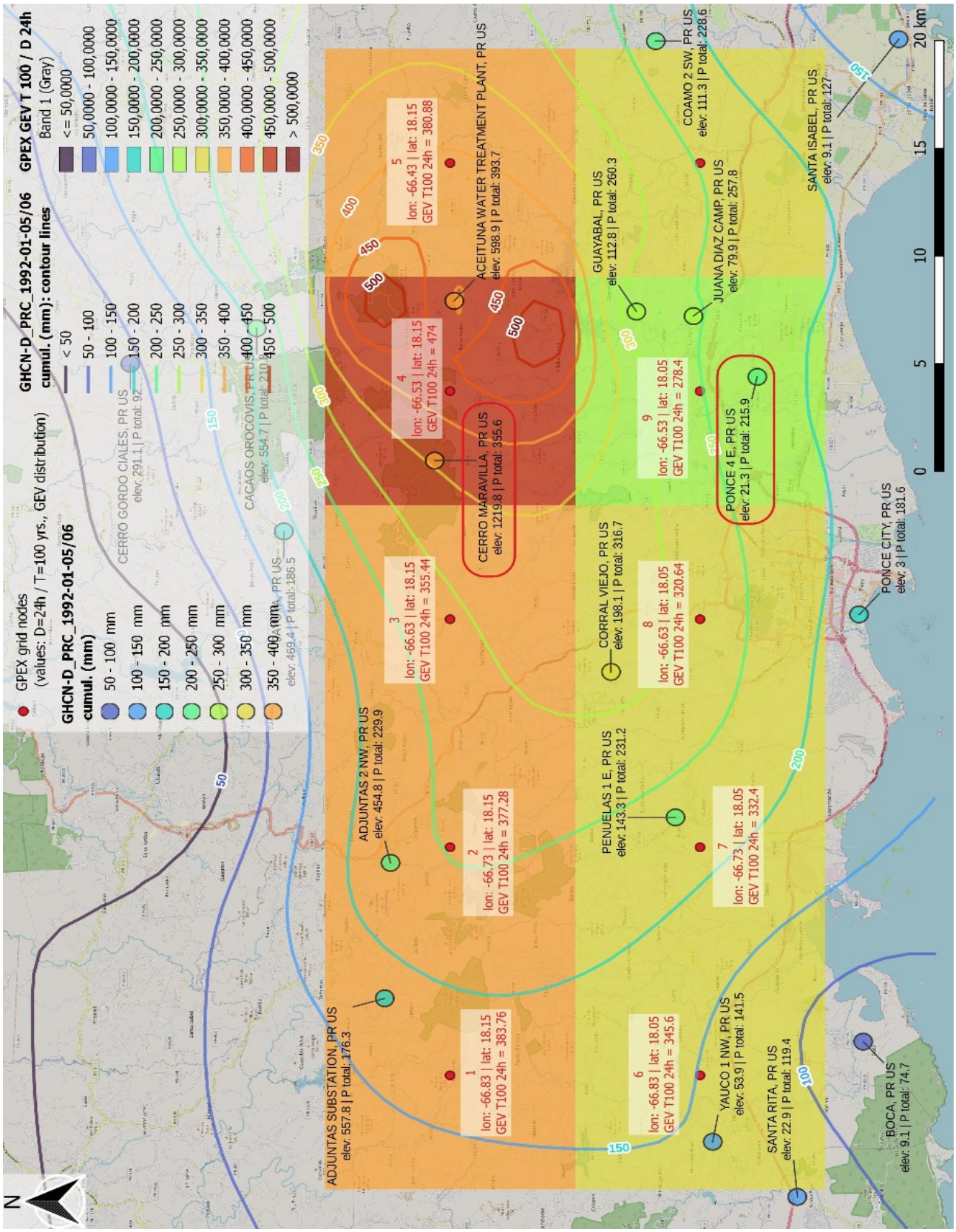


Figure 45: NOAA GHCN-D / HDSC stations (with precip values of 1992-01 rain event), GPEX GEV values for D=24h, T=100yrs (stations in red outline: chosen for comparison GPEX to NOAA / HDSC precipitation frequency estimates).

For the comparison between GPEX and HDSC, I chose the two stations “PONCE 4 E” (21m altitude, 18.0258°N / -66.5253° W) and “CERRO MARAVILLA” (1220m altitude, 18.1547° N / --66.5619° W), which lie in neighbouring GPEX grid cells and from which I believe that they represent the possible lower and upper boundary of the precipitation regime in the Carrefour AOI quite well – “PONCE 4” for the urban area of Carrefour along the coast and “CERRO MARAVILLA” for the uphill region at the watershed (though twice as high here, so orographic rainfall sums will be higher here than at the watershed of the Carrefour AOI).

Furthermore, both stations have time series with sub-daily measurements down to the duration of one hour (only a part of the PRC station network has sub-daily rainfall measurement) and in the 10 extracted grid cells of the GPEX dataset, the two grid cells where the stations are situated, also represent the upper and the lower boundary of all values (see Figure 45), so it will be an interesting and representative evaluation.

Looking at the annual maxima (time series available for the 1h duration for 33 years, 1971-2003)[49], the highest value in the records for the 1h duration is 2.8 inches / 71.1 mm, but the time series ends in the year 2003, so any heavy precipitation event after that is not included.

3.2.3.1 IDF Statistics: GPEX vs. HDSC for station “PONCE 4E” (PRC)

The point precipitation frequency estimates for PONCE 4E are delivered together with the upper and lower boundary of the 90% confidence interval (see Table 6):

PDS-based precipitation frequency estimates with 90% confidence intervals (in millimeters/hour)[1]										
Duration	Average recurrence interval (years)									
	1	2	5	10	25	50	100	200	500	1000
5-min	**112** (107-129)	**148** (136-163)	**177** (162-194)	**198** (180-216)	**223** (202-244)	**241** (218-264)	**258** (232-285)	**275** (247-305)	**297** (264-333)	**314** (278-354)
10-min	**76** (73-88)	**101** (93-111)	**121** (111-133)	**135** (123-148)	**152** (138-166)	**165** (149-181)	**176** (159-194)	**188** (168-209)	**203** (181-228)	**215** (190-242)
15-min	**65** (63-76)	**87** (79-95)	**104** (95-113)	**116** (105-126)	**130** (118-142)	**141** (127-155)	**151** (136-166)	**161** (144-179)	**174** (155-195)	**184** (162-207)
30-min	**52** (50-61)	**69** (63-76)	**83** (76-91)	**93** (84-101)	**104** (95-114)	**113** (102-124)	**121** (109-133)	**129** (115-143)	**139** (124-156)	**147** (130-166)
60-min	**39** (37-45)	**51** (47-57)	**62** (56-67)	**69** (62-75)	**77** (70-85)	**84** (76-92)	**90** (81-99)	**96** (86-106)	**103** (92-116)	**109** (96-123)
2-hr	**23** (23-28)	**32** (29-35)	**40** (36-44)	**45** (41-49)	**52** (46-57)	**56** (50-62)	**61** (54-68)	**66** (58-74)	**72** (63-81)	**77** (66-87)
3-hr	**18** (16-20)	**23** (21-26)	**30** (27-33)	**34** (30-37)	**39** (35-43)	**43** (38-48)	**47** (41-53)	**51** (45-58)	**56** (49-65)	**60** (52-70)
6-hr	**11** (10-13)	**15** (13-17)	**19** (17-22)	**23** (20-26)	**27** (24-31)	**31** (27-35)	**34** (29-40)	**38** (32-44)	**43** (36-51)	**47** (39-56)
12-hr	**7** (6-8)	**9** (8-10)	**12** (10-14)	**14** (12-17)	**18** (15-21)	**21** (17-24)	**24** (20-28)	**27** (22-32)	**31** (25-37)	**35** (28-42)
24-hr	**4** (4-5)	**5** (5-6)	**7** (6-8)	**9** (8-10)	**12** (10-13)	**14** (12-16)	**16** (13-18)	**18** (15-21)	**21** (18-25)	**24** (20-28)

Table 6: NOAA HDSC Precipitation frequency (PF) estimates for station "PONCE 4E" (18.0258° N / 66.5253° W) (data source: https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_pr.html, (access on 2023-04-10)

For the station PONCE 4E (see figures 46, 47, 48, 49), the comparison looks like this:

49 https://hdsc.nws.noaa.gov/pub/hdsc/data/_TimeSeries_stations/PR_66-7292_ams.txt ((access on 2023-04-10)

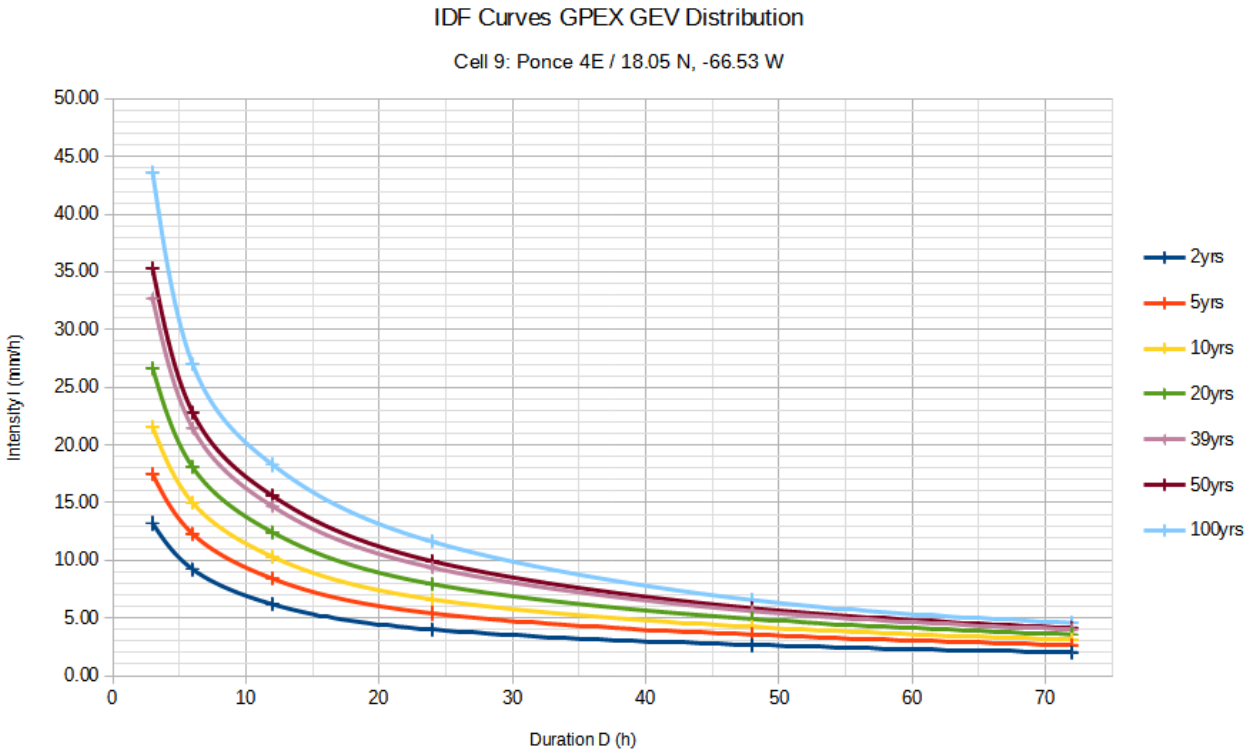


Figure 46: IDF Curves for T2 / T5 / T10 / T20 / T39 / T50 / T100 for GPEX gridcell with centroid at 18.05°N / -66.53°W – time range 3h until 72h (data source:

At the **upper boundary of the durations (72h)**, the GPEX dataset has the curves of the return periods pretty tight to each other between 2.0 (T2) and 4.6 (T100) mm/h, whereas the HDSC station-derived dataset delivers intensities between 2.0 (T2) and 8.0 (T100) mm/h. So the intensities are the same between GPEX and HDSC for T2/72h, but quite different for T100/72h (HDSC with roughly twice the value of GPEX).

For the **24h duration**, GPEX sees the intensities between 4.0 (T2) and 11.6 (T100) mm/h; HDSC station data supplies intensities between 4.0 (T2) and 16.0 (T100) mm/h. **The January 1992 rainfall event** (210 mm at PONCE station in 24h = 8.75 mm/h in average) **would therefore classify as a ~T30 rain event in GPEX, but only as a ~T10 rain event in the HDSC station data from PONCE 4E**.

At the **duration of 3h** (=lower boundary of GPEX dataset), GPEX has values between 13.2 (T2) and 43.6 (T100) mm/h, whereas HDSC delivers values between 24.0 (T2) and 50.0 (T100) mm/h – so unlike at the 72h duration, at the 3h duration, the rainfall intensity per hour is almost twice the value of HDSC in GPEX for the 2-year return period, whereas the values are quite similar for the 100 years return period.

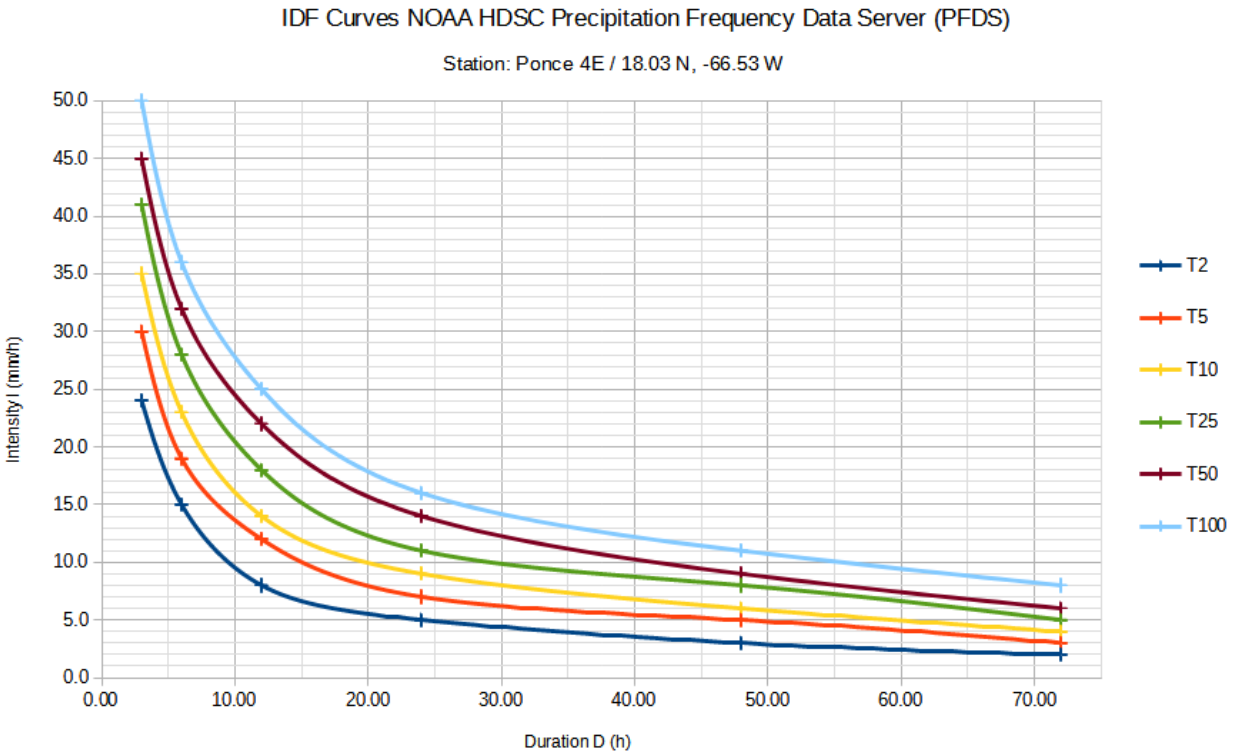


Figure 47: IDF Curves for T2 / T5 / T10 / T25 / T50 / T100 for NOAA HDSC station “PONCE 4E” (18.03°N / -66.53°W) – time range 3h until 72h (data source:

For being able to compare the GPEX and the HDSC dataset even below the officially delivered 3h duration of GPEX, it was necessary to extrapolate the curve of every return period backwards in LibreOffice Calc to get values of the 1 hour duration. All rainfall frequency functions are exponential functions, unfortunately, the function behind the GEV fitted values in the GPEX dataset cannot be extracted, only the values themselves. So I fitted the extrapolation curve in Calc via a “normal” exponential function: At the example of the T100 curve, the extrapolation curve has the function f(x) = 97.431823 * x ^ -0.697255 and a determination coefficient R^2 of 0.994680 (see Figure 48).

Of course, data extrapolations are always risky and fraught with uncertainty, but it is the only method I could think of to at least get a rough impression of the difference between the two datasets at the 1h duration, which will be used for the flood calculation in the AOI Carrefour.

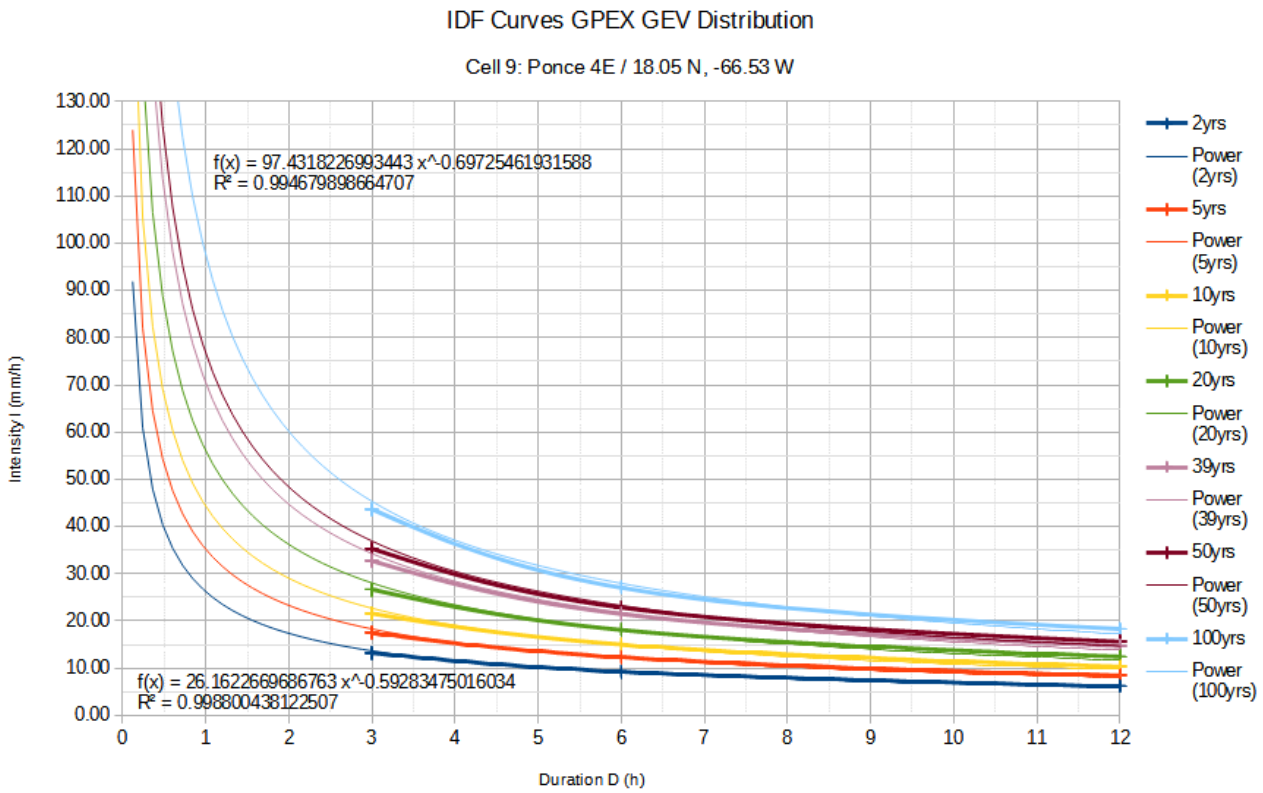


Figure 48: IDF Curves for T2 / T5 / T10 / T20 / T39 / T50 / T100 for GPEX gridcell with centroid at 18.05°N / -66.53°W – time range 0.25h until 12h (data source: Gruendemann et al., 2020; values between 0.25h and 3h are extrapolated via exponential functions)

Taking into account these uncertainties, for the **duration of 1 hour**, the **extrapolated GPEX** dataset delivers rainfall intensities between 25 mm/h for the T2 return period (**35 mm/h for T5**) and **97.5 mm/h for the T100** return period for **D=1h**, whereas **HDSC** delivers rainfall intensities between 56 mm/h for T2 (**65 mm/h for T5**) and **91 mm/h for T100.** So for the 100-year return period, the delta between the two datasets would not be significant (in both cases between 90 and 100 mm/h), whereas for the lower return periods T2 and T5, GPEX values are only about half the values of HDSC.

For **T50**, the **extrapolated GPEX** dataset returns a rainfall intensity of **78 mm/h** for the **1h duration**, whereas **HDSC** returns an intensity of **85 mm/h** for **D=1h.**

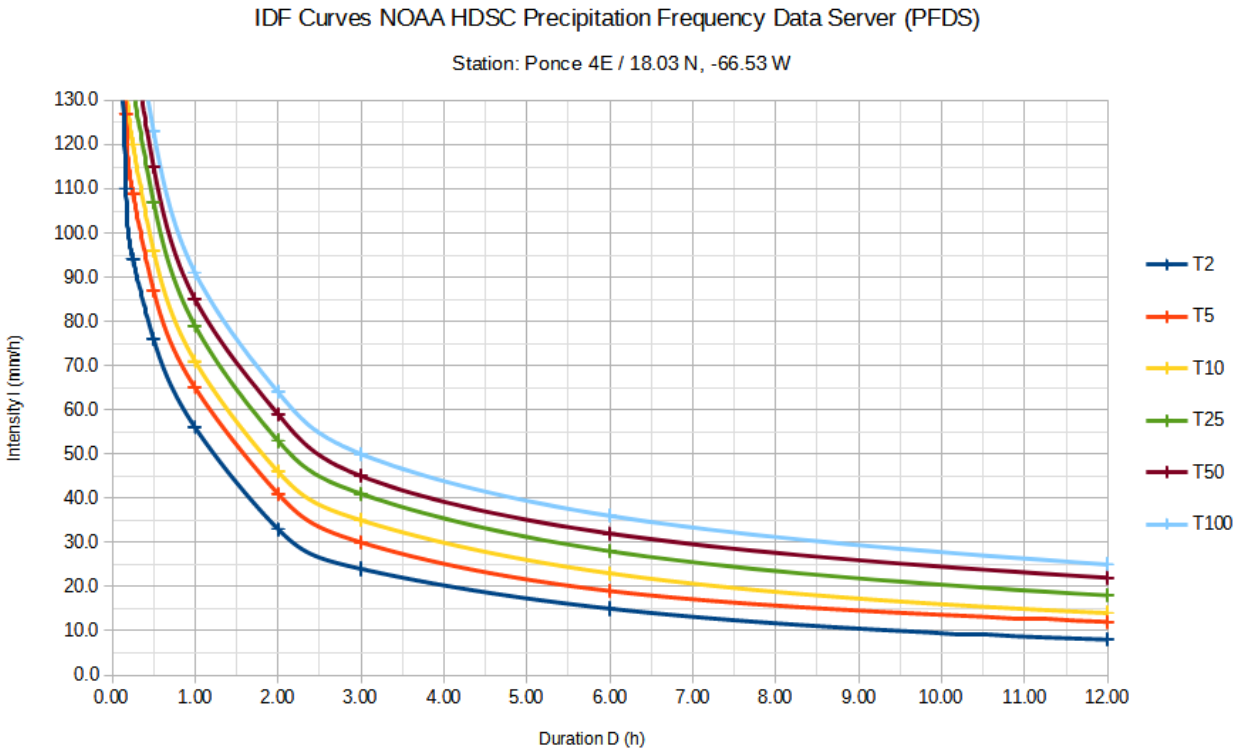


Figure 49: IDF Curves for T2 / T5 / T10 / T25 / T50 / T100 for NOAA HDSC station “PONCE 4E” (18.03°N / -66.53°W) – time range 0.25h until 12h (data source: Bonnin et al., 2008)

3.2.3.2 IDF Statistics: GPEX vs. HDSC for station “CERRO MARAVILLA” (PRC)

The point precipitation frequency estimates for CERRO MARAVILLA are delivered together with the upper and lower boundary of the 90% confidence interval (see Table 7):

PDS-based precipitation frequency estimates with 90% confidence intervals (in millimeters/hour)[1]										
Duration	Average recurrence interval (years)									
	1	2	5	10	25	50	100	200	500	1000
5-min	184 (172-194)	220 (208-235)	241 (228-258)	265 (248-284)	299 (278-322)	327 (301-355)	356 (325-388)	386 (349-424)	429 (382-475)	462 (407-515)
10-min	126 (118-133)	150 (142-160)	165 (156-176)	181 (170-194)	204 (190-220)	223 (206-242)	243 (222-265)	264 (238-290)	293 (261-325)	315 (278-352)
15-min	107 (101-114)	129 (122-137)	141 (133-151)	155 (145-166)	175 (163-188)	191 (176-207)	208 (190-227)	226 (204-248)	251 (223-278)	270 (238-301)
30-min	86 (81-91)	103 (98-110)	113 (107-121)	124 (116-133)	140 (130-151)	153 (141-166)	167 (152-182)	181 (163-199)	201 (179-222)	216 (190-241)
60-min	64 (60-68)	76 (72-81)	84 (79-90)	92 (86-99)	104 (97-112)	113 (105-123)	124 (113-135)	134 (121-147)	149 (133-165)	160 (141-179)
2-hr	47 (41-47)	54 (51-58)	62 (58-68)	71 (66-78)	85 (77-93)	96 (87-105)	108 (96-119)	120 (106-134)	137 (119-155)	151 (129-172)
3-hr	31 (28-33)	38 (35-41)	44 (41-49)	52 (47-58)	64 (57-71)	73 (64-83)	84 (72-95)	95 (81-109)	111 (93-129)	124 (102-146)
6-hr	17 (15-19)	22 (19-24)	27 (24-31)	33 (29-38)	42 (36-49)	50 (42-58)	58 (48-69)	67 (54-80)	81 (63-98)	92 (71-113)
12-hr	9 (8-11)	12 (10-14)	17 (14-20)	21 (17-26)	28 (22-34)	34 (27-42)	41 (31-52)	49 (37-62)	61 (44-80)	72 (50-95)
24-hr	6 (5-7)	7 (6-9)	10 (9-12)	13 (11-16)	18 (15-22)	23 (18-27)	28 (22-33)	34 (26-40)	43 (33-52)	51 (38-62)

Table 7: NOAA HDSC Precipitation frequency (PF) estimates for station "CERRO MARAVILLA" (18.1547° N / -66.5619° W) (data source: https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_pr.html, (access on 2023-04-10)

For the station CERRO MARAVILLA (see figures 50, 51, 52, 53), the comparison looks like this:

Also for grid cell Nr. 4 (see Figure 45), which covers the area of CERRO MARAVILLA station, at the upper boundary of the durations (72h), the GPEX dataset has the curves of the return periods pretty tight to each other between 2.5 (T2) and 7.7 (T100) mm/h, whereas the HDSC station-derived dataset delivers intensities between 3.0 (T2) and 12.0 (T100) mm/h. **So again, the intensities are almost the same between GPEX and HDSC for T2/72h, but quite different for T100/72h (HDSC with almost twice the value of GPEX)**.

For the 24h duration, GPEX sees the intensities between 5.1 (T2) and 19.8 (T100) mm/h; HDSC station data supplies intensities between 6.0 (T2) and 28.0 (T100) mm/h. **The January 1992 rainfall event** (305 mm at CERRO MARAVILLA station in 24h = 12.7 mm/h in average) **would therefore classify as a ~T25 rain event in GPEX, but again, only as a ~T10 rain event in the HDSC station data from PONCE 4E.**

At the duration of 3h (=lower boundary of GPEX dataset), GPEX has values between 16.6 (T2) and 63.4 (T100) mm/h, whereas HDSC delivers values between 38.0 (T2) and 84.0 (T100) mm/h – so this time, at the 3h duration, the rainfall intensity per hour in the HDSC station dataset is more than double the value of GPEX for the 2-year return period, whereas the values are quite more similar for the 100 years return period (HDSC value = GPEX value * 1.33).

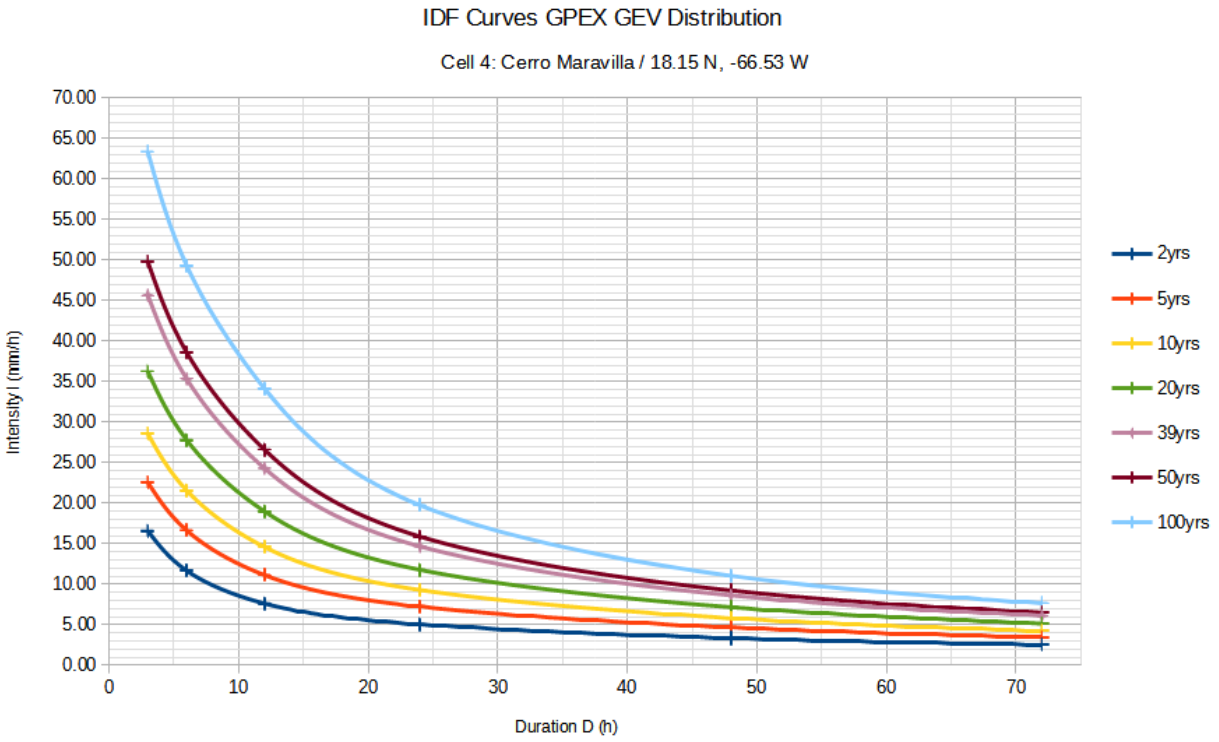


Figure 50: IDF Curves for T2 / T5 / T10 / T20 / T39 / T50 / T100 for GPEX grid cell with centroid at 18.15°N / 66.53°W – time range 3h until 72h. Note the different scale on y-axis compared to Figure 46. (data source: Gruendemann et al., 2020)

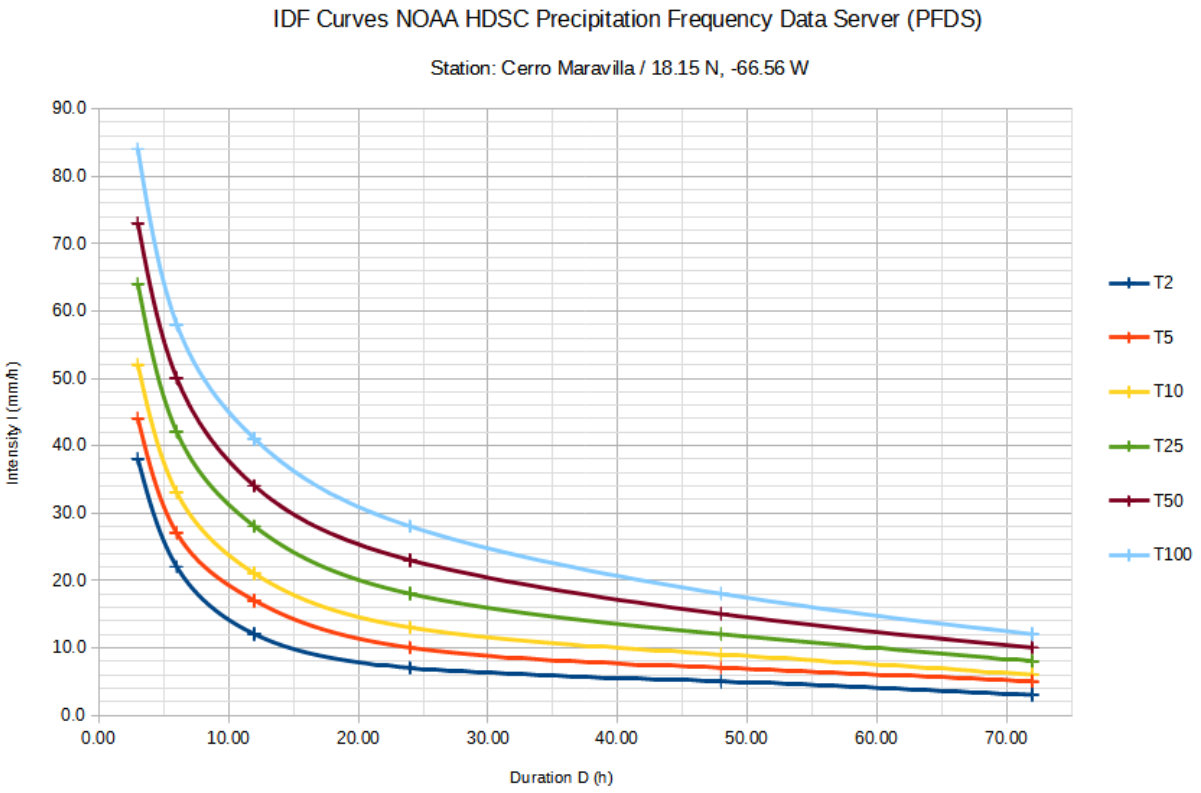


Figure 51: IDF Curves for T2 / T5 / T10 / T25 / T50 / T100 for NOAA HDSC station “CERRO MARAVILLA” (18.15°N / -66.56°W) – time range 3h until 72h (data source: Bonnin et al., 2008)

The extrapolation of the GPEX value in LibreOffice Calc did not fit that good in case of the Cerro Maravilla grid cell (see Figure 52) - but again, taking into account these uncertainties, for the **duration of 1 hour**, the **extrapolated GPEX** dataset delivers rainfall intensities between **32 mm/h for the T2 return period** (**45 mm/h for T5**), almost **120 mm/h for T50** and a value in the **range between 150 and 160 mm/h for the T100** return period for **D=1h**, whereas **HDSC** delivers rainfall intensities between **76 mm/h for T2** (**84 mm/h for T5**) and **124 mm/h for T100.** So this time, for the 100-year return period, the delta between the two datasets would be significant (155 mm/h in GPEX vs. 124 mm/h in HDSC). For the return periods of T2 and T5, the delta is even higher: GPEX values are again only about half the values of HDSC.

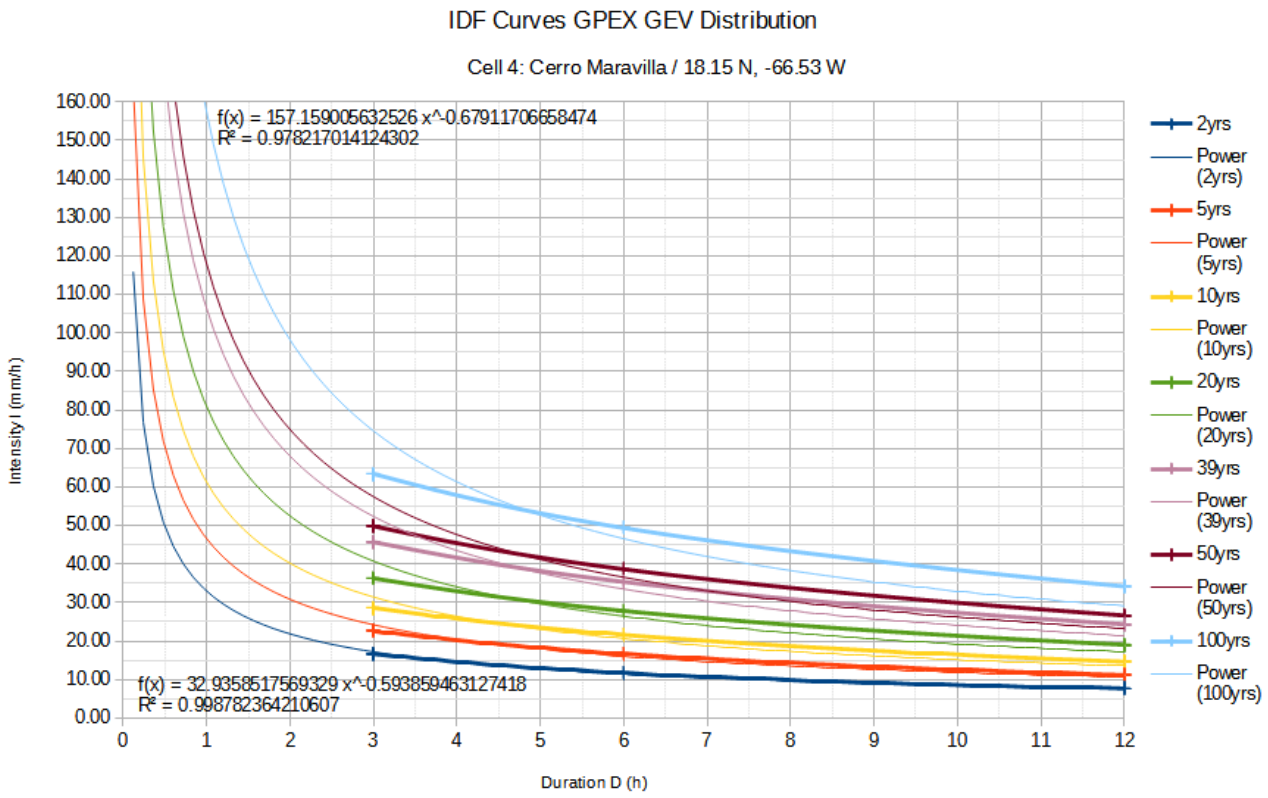


Figure 52: IDF Curves for T2 / T5 / T10 / T20 / T39 / T50 / T100 for GPEX gridcell with centroid at 18.15°N / -66.53°W – time range 0.25h until 12h (data source: Gruendemann et al., 2020; values between 0.25h and 3h are extrapolated via exponential functions)

For **T50**, the **extrapolated GPEX** dataset returns a rainfall intensity of **120 mm/h** for the **1h duration**, whereas **HDSC** returns an intensity of – and this is interesting – **113 mm/h** for **D=1h.** In Figure 53, one can clearly see the "bump" between the durations 1 hour and 2 hours, here, in the higher return periods, the curve progression is a bit unusual and suggests the suspicion that for the T50 and the T100 return period, either the 1h value is too low or the 2h value is too high.

Looking at the annual maxima (time series available for the 1h duration for 29 years)[50], the highest value in the records for the 1h duration is 3.4 inches / 86.3 mm, but the time series ends in the year 2003, so any heavy precipitation event after that is not included.

Summarizing the results of the two stations PONCE 4E and CERRO MARAVILLA, these are the key conclusions:

- All GPEX dataset values for the duration 3h have to be multiplied by a factor of 1.25 (inverse of 0.8) to account for the ARF of ~0.8 (~100 km² grid cell area, see Chapter 3.2.2) → GPEX * 1.25 = HDSC. The same is valid for the extrapolated 1h duration (ARF of ~0.7) values, only with the factor 1.4 (inverse of 0.7) → GPEX * 1.4 = HDSC
- For the lower boundary of the GPEX dataset durations (3h), the lower return periods (T2-T10), regarding an ARF of 0.8 for direct comparison to the station derived HDSC dataset, have a delta of ~ factor 1.5, the higher return periods (T50, T100) by factor ~1.2

50 https://hdsc.nws.noaa.gov/pub/hdsc/data/_TimeSeries_stations/PR_66-2336_ams.txt (access on 2023-01-11)

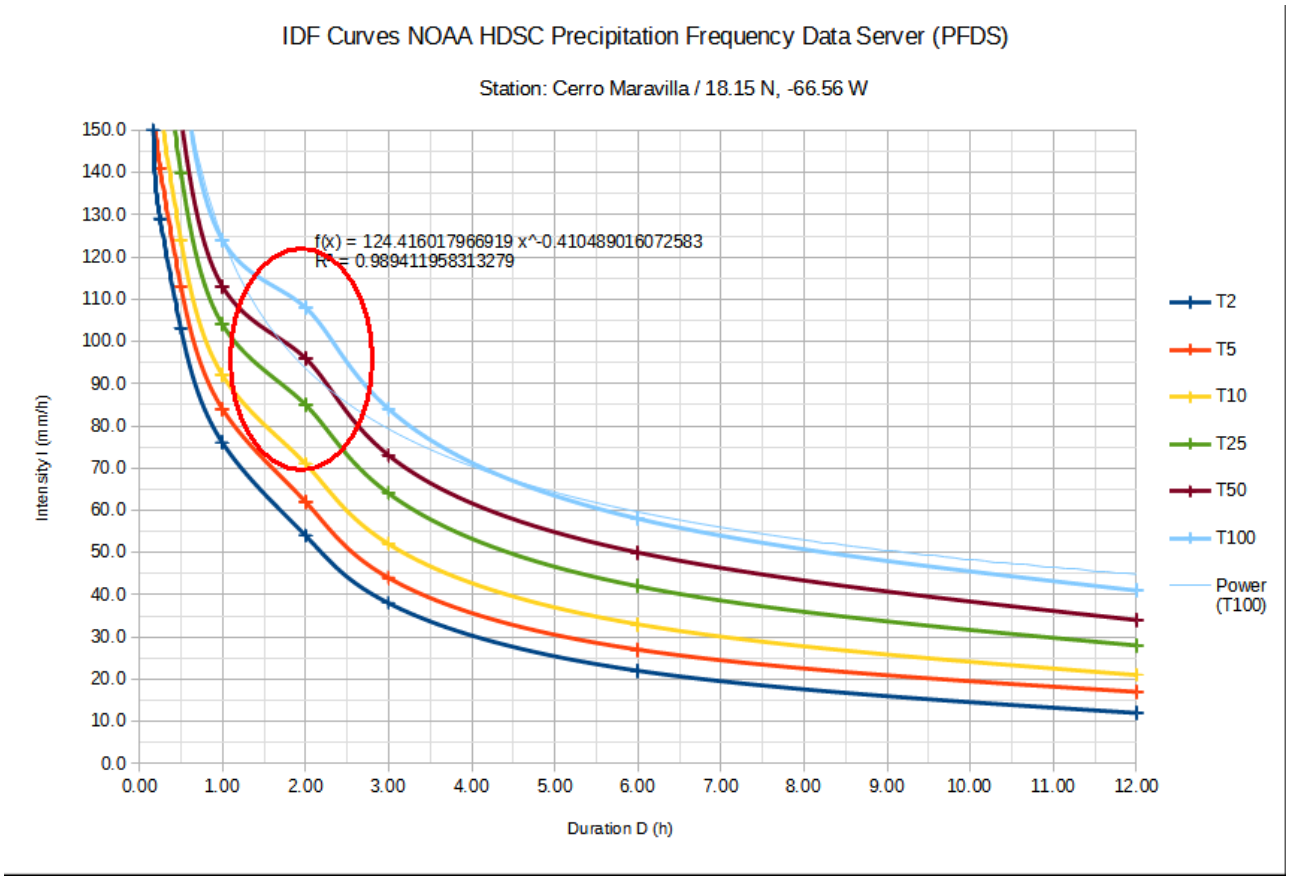


Figure 53: IDF Curves for T2 / T5 / T10 / T25 / T50 / T100 for NOAA HDSC station "CERRO MARAVILLA" (18.15°N / -66.56°W) – time range 0.25h until 12h (data source: Bonnin et al., 2008). Note the "bump" at D=2h!

- Taking into the account the extrapolation until the duration of 1h, for PONCE 4E, the T100 intensity does not differ a lot between GPEX and HDSC – regarding an ARF of 0.7, GPEX (97.5 mm/h) would be even considerably higher than HDSC (91 mm/h + 0.7 = 64 mm/h). For the mountain region, with example CERRO MARAVILLA, the extrapolated GPEX values (~150 mm/h) is already higher than the HDSC data (125 mm/h), including an ARF of 0.7, the difference is increasing to factor 1.7.
- As the **extrapolation of the IDF curves of the GPEX** dataset can **not** be seen as a **reliable method for deriving rainfall intensities** for design precipitation and therefore has been done illustration reasons only, the **decision for the used rainfall intensities** in the design precipitation for the **flood simulation in Chapter 4** will be made upon the **HDSC** point precipitation frequency estimates of the two stations **Ponce 4 E and Cerro Maravilla** and the rainfall evaluation of the Damien rain gauge in Port-au-Prince (see following chapter 3.2.3.3) carried out by (Heimhuber et al., 2015).

3.2.3.3 IDF Statistics: station "Damien", Port-au-Prince, Haiti

In the year 2015, Heimhuber et al., 2015 have carried out a study on flood risk assessment for the informal settlement of "Onaville" in the cul-de-sac plain north of Port-au-Prince. Therefore, (then

available) rain gauge data from the station of Damien, basing on 68 years of observation (1927-2002) has been evaluated with the resulting IDF Curve in Figure 54:

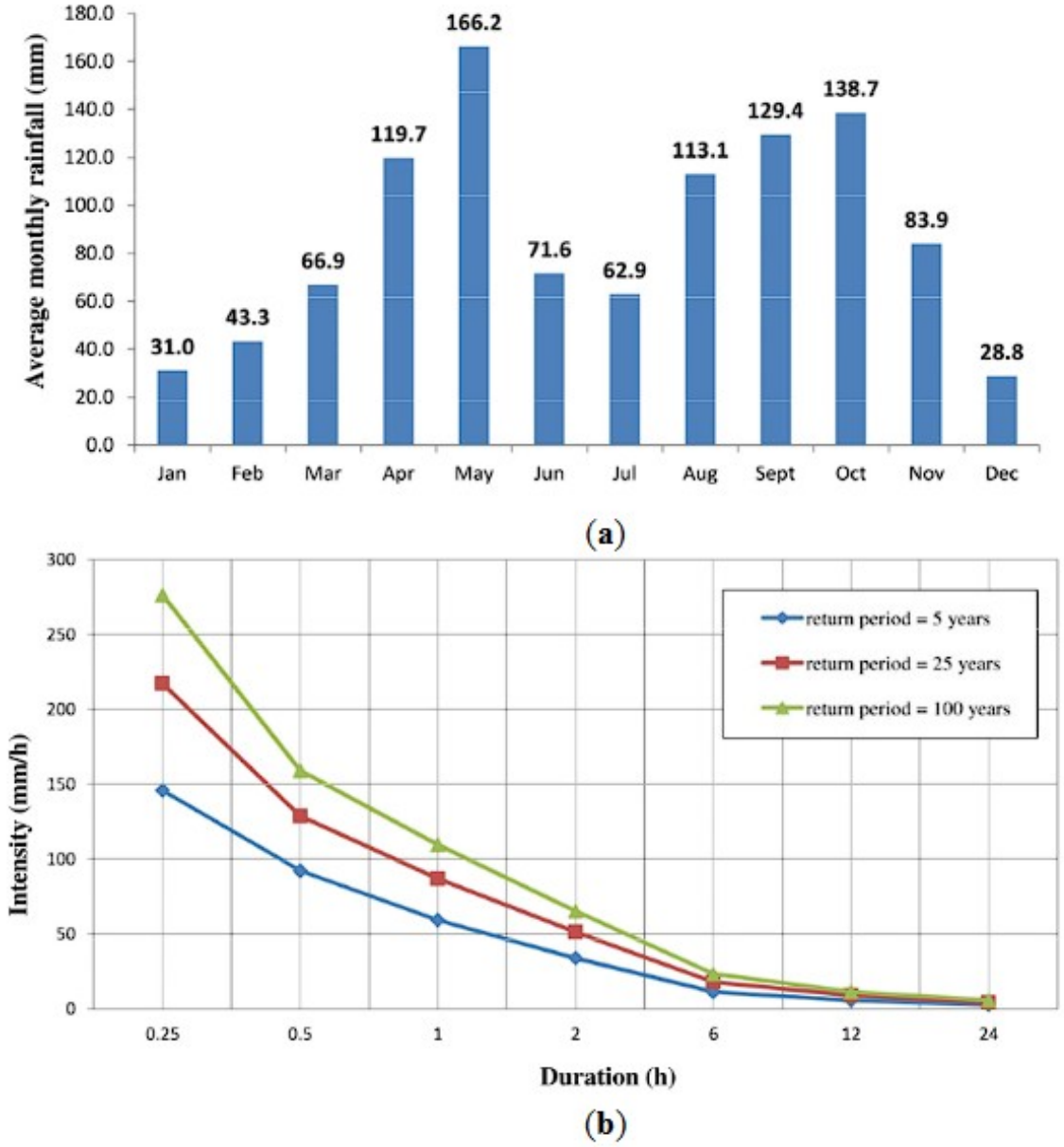


Figure 54: (a) Average monthly rainfall at Damien rain gauge (n=68, 1927-2002) and (b) IDF-curves for Damien station

The rainfall intensities for the duration of 1h in Figure 54 (~110 mm/h for the T100 return period, ~85 mm/h for T25, ~55 mm/h for T5) are similar to the intensities of the two stations Ponce 4E and Cerro Maravilla at the Puertorican South coast, as well as to the extrapolated GPEX data in Ponce.

3.3. Model calibration

For pluvial flood model calibration, the gold standard would be the availability of three essential parameters:

- Rain gauge data with a temporal resolution of at least 15 min or higher
- At least one discharge gauge at a downstream location with a temporal resolution of at least 1 hour or higher OR adequate satellite images / ground photographs, from which the peak of the flood extent can be estimated
- Either the existence of acknowledged design precipitation data or the availability of long-term rainfall data of at least one gauge in or near the area of interest, with the same temporal resolution like stated above (15 min or higher), to build IDF curves (see chapter 3.2.3)

With all three of these, you are able to a) calibrate and furthermore also validate your flood model and b) with the help of extreme value statistics, classify previous precipitation events for a certain probability of recurrence, commonly known as return period.

In the case of Haiti, I had none of these three parameters at my disposal (with the exception of the Damien rain gauge dataset described in Chapter 3.2.3.3, but the data source behind this IDF curve was not available for me, so I could not evaluate this by myself).

Although there is an existing rain gauge network in Haiti, it was not possible to obtain any data, neither for me, nor for MSF. As mentioned before, the Haitian state is currently undergoing a deep crisis and therefore its administration has to fight with many problems at the same time, so, understandably, answers to and data for my academic research topic would not have priority number one. So I had to look for alternatives.

3.4. Pluvial flood modelling

All input parameters described in the following chapters (DTM including buildings, rainfall, roughness and infiltration) are exactly the same for both types of flood calculation, the “classic” 2d hydrodynamic calculation via the – in German speaking countries widely used – flood calculation software “Hydro-As-2D”[51] as well as the merely GIS-based filling-and-spilling algorithm “Safer_RAIN” (Samela et al., 2020).

All input variables are – as far as possible – deliberately kept simple to enable the best possible comparison of the two approaches. For this reason, for example, and also due to the fact that soil conditions are generally very poor in the AOI, infiltration has been set to zero. Of course, there will be some infiltration in case of a heavy to extreme rain event like in the scenarios, but as imperviousness in the urban part of the AOI is high, soil seepage capacity and tree cover both are very low, the percentage of total precipitation which infiltrates the soil instead of turning into overland flow will neglectable.

3.4.1 Model input: DTM, buildings

Like already described before, the DTM of the AOI consists of two different datasets, both of them being LIDAR data. Around 90% of the AOI are covered with the LIDAR point cloud data of the 2010 survey[52], whereas the rest is covered by the newer survey carried between 2014 and 2016[53] (see Figure 11 in chapter 2.2). Unfortunately, the point cloud data is not accessible and the gridded DTM / DSM data is available in integer data type only (see Chapter 2.2.2).

This was quite a challenge, as the integer DTM “as is” had proven to be unfeasible to serve as basis for any flood calculation, due to the “rice terrace effect”, where the water accumulated at the terraces in the 2d flood model, building ponds there instead of a steady flow downhill (see Figure 14).

But on the other hand, the availability of the point cloud data for the coastal urban region of the Carrefour AOI allowed the integration of buildings directly into the DTM as the point-cloud was poorly classified anyway and therefore predestined to kill two birds with one stone: Firstly get a suitable DTM for the flood calculation with a minimum of vegetation in it, and secondly, integrate the building structure. Via the “Zmin” algorithm, where the interpolation between the single LIDAR points rely on the lowest Z values of overlapping or nearby LIDAR points, we still have the building structure reproduced quite well, with at the same time a minimum of disturbing vegetation in the raster cell (see Chapter 2.2.1 / Figure 12 and Figure 56 on the next page)

51 https://www.hydrotec.de/software/hydroas/ (access on 2023-04-12)
52 https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.072010.32618.1 (access on 2021-12-05)
53 https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.082021.32618.1 (access on 2021-11-08)

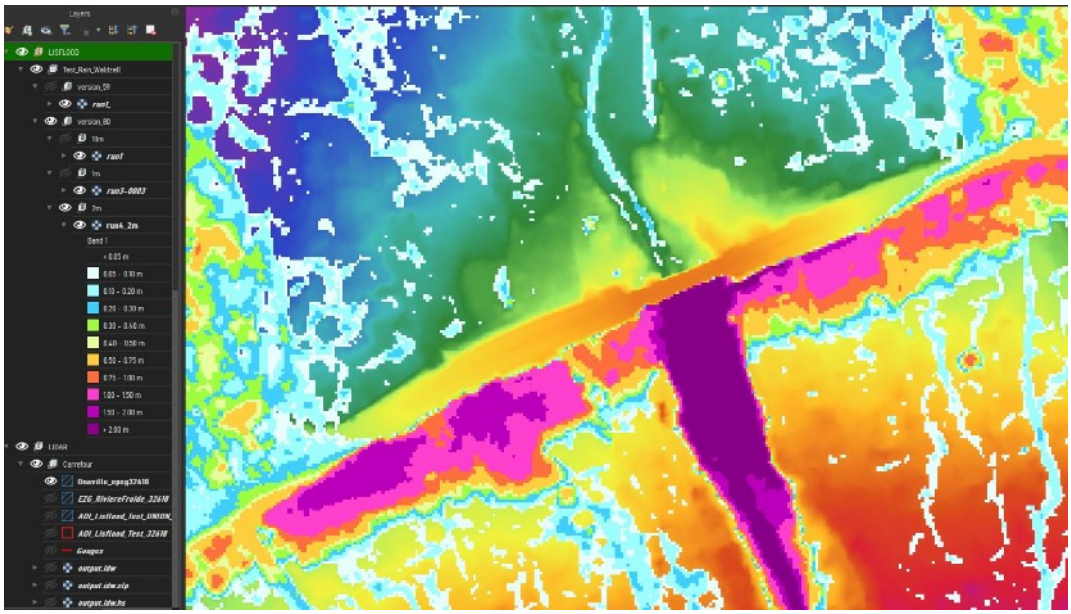


Figure 55: Manual terrain preprocessing: removing of bridges in the DTM

But not everything could be automated – the preparation of a hydraulic terrain model is lengthy and sometimes arduous as well, as in many cases, items have to be corrected manually. This begins at the removal of bridges, which is absolutely essential a sine qua non to avoid false and misleading results in the flood calculation (see Figure 55) and leads until the filling of void pixels, where there had been no valid LIDAR point in the 4 m² raster cell and therefore z values are invalid.

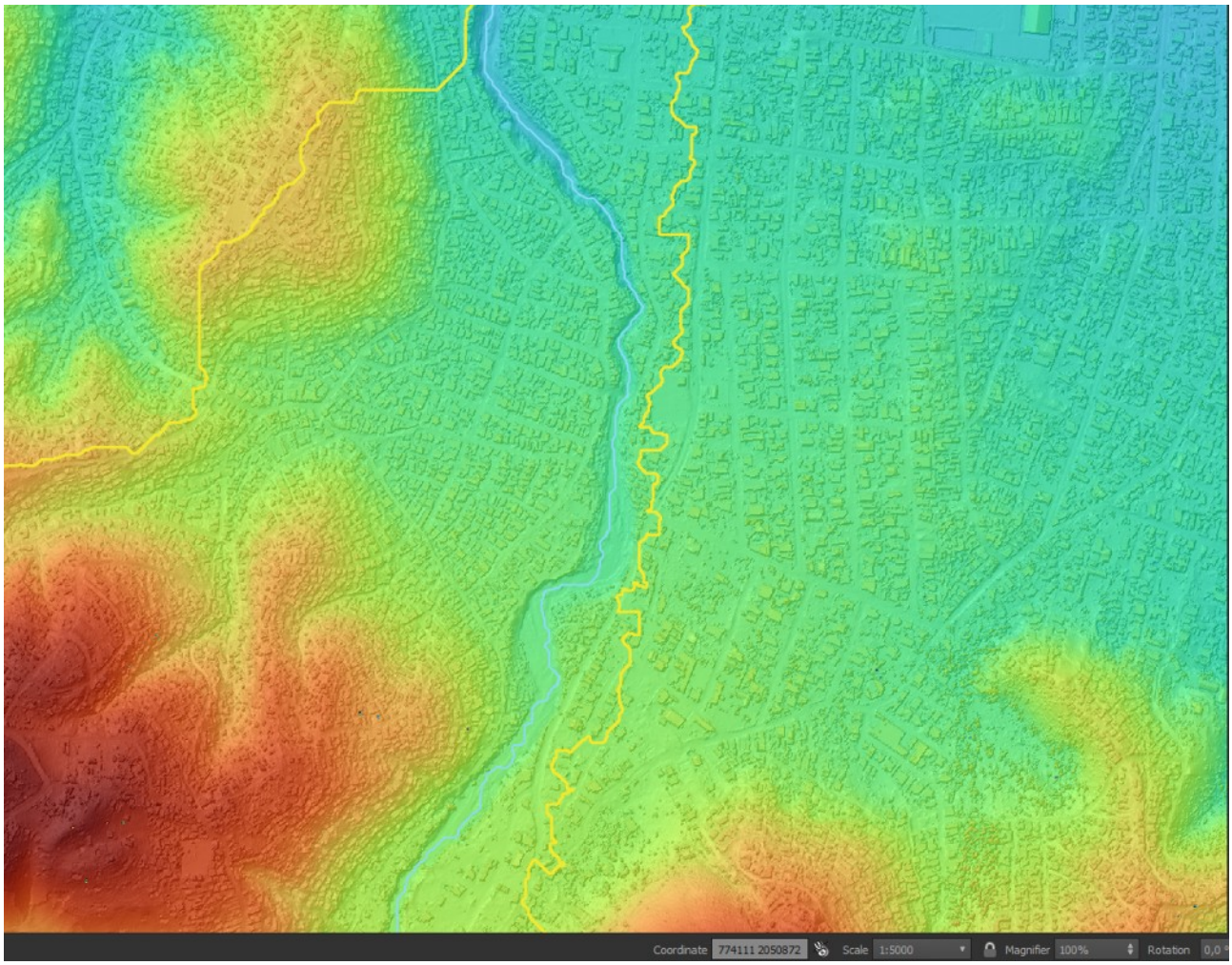


Figure 56: interpolation of the unclassified point cloud to a 2m pixel size raster DTM, with minimum z values of the point cloud per grid cell, still representing the building structure with a minimum of vegetation in it

3.4.2 Input: Rainfall

The time of concentration T_c (time required for runoff to travel from the hydraulically most distant point in the watershed to the outlet) matters for any **hydrodynamic flood calculation,** be it 1D or 2D (not for the fill-and-spill algorithm which is used by "Safer_RAIN", as flow velocity is not calculated in this approach!), where flow velocity is calculated in the simulation. In engineering practise, a wide range of different equations are available for the computation of T_c, which leads to significant variability and uncertainty in the estimation (Ravazzani et al., 2019).

In the case of our AOI Carrefour, where all catchments of the "ravines" have 3-5 km² area maximum and where we have a steep relief on plus, T_c is definitely well below one hour – but there are hardly any global or local datasets available for rainfall durations below one hour, and furthermore, the "SAFER PLACES" platform only allows calculations of rainfall durations of 1h upwards, so I chose the one hour duration for the pluvial flood scenario.

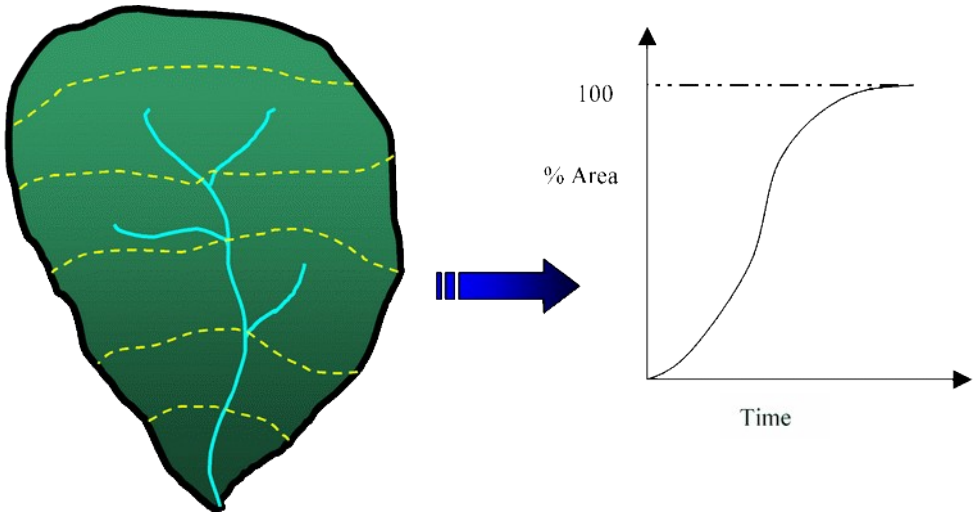


Figure 57: Illustration for the time of concentration (source: https://directives.sc.egov.usda.gov/OpenNonWebContent.aspx?content=27002.wba, access on 2023-04-08)

Tc is dependant not only from size and shape of the catchment (see Figure 57), but also from slope and terrain roughness (Strickler / Manning, see Figure 58 and following chapter 3.4.3), which all have a big influence on how fast the peak of the flood wave will reach the outlet of the catchment.

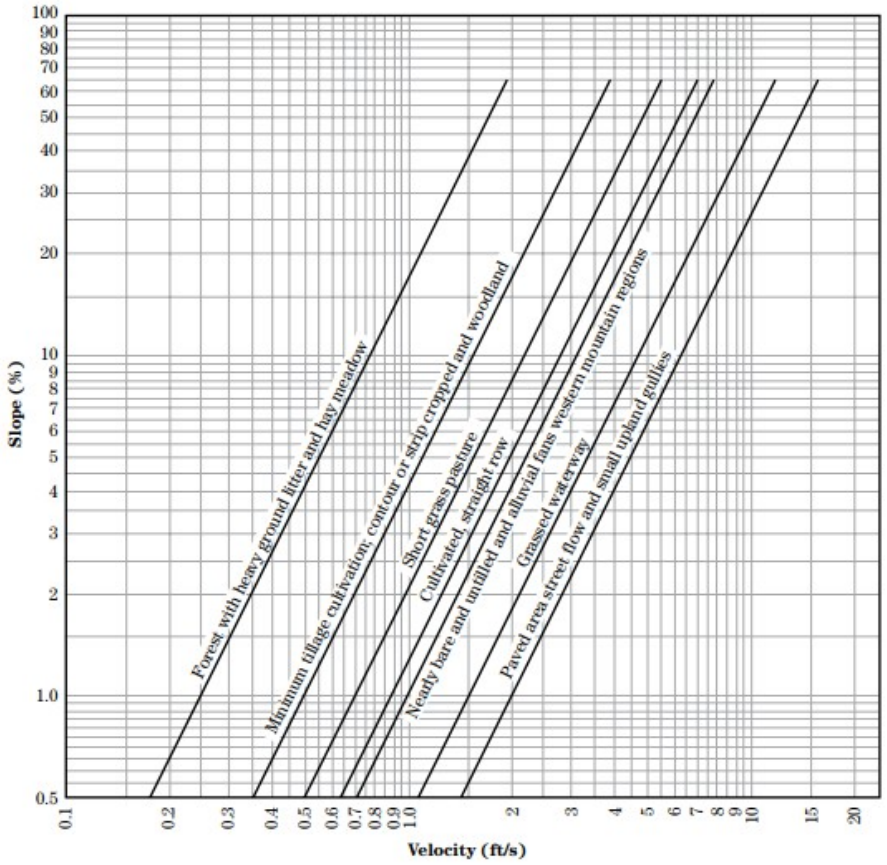


Figure 58: Flow velocity as dependant variable from Slope and Roughness (Manning). National Handbook of Engineering, https://irrigationtoolbox.com/NEH/Part630_Hydrology/NEH630-ch15draft.pdf, p. 15-6, access on 2023-04-08)

Merging all the information of the previous chapters into one single pot, I decided to carry out **three different flood scenarios** for our AOI Carrefour:

- **Scenario of a frequent rain event** (return period T2/T5, 50 mm of rainfall in 60 minutes)
- **Scenario of a rare rain event** (return period T30/T50, 90 mm of rainfall in 60 minutes)
- **Scenario of an extreme rain event** (return period T100 / T200, 130 mm of rainfall in 60 minutes)

The not-unique quotation of a single return period is in full intent, as in my estimation, it is simply not possible to assign a certain rainfall intensity in mm/h a unique return period with the prevailing data situation – the unknown parameters, first and foremost an available, coherent and long-term time series of a rain gauge nearby the AOI, are too much to state a certain rainfall intensity I as unique amount for return period T. But, as it had been shown in the chapters 2.3 and 3.2, it is possible to narrow down the probable range of rainfall intensities I for return period T, or, vice versa, the range of return periods T for a certain rainfall intensity I in mm/h.

Why vice versa? Because in the flood simulation, I cannot give a range of possible rainfall amounts as input, each model run, whether via Hydro-As-2D or via the "SaferPLACES" platform, needs a certain rainfall amount in a certain duration as input variables. So this is the compromise for a data scarce region.

3.4.3 Input: Roughness, soil and infiltration

Roughness coefficient is an essential parameter for every flood calculation, as it is – together with slope – **the** critical parameter for flow velocity. In the Anglo-Saxon world, the common parameter is the Manning coefficient, whereas in German speaking countries, the Strickler coefficient is prevailing. But anyway, Manning coefficient is just the reciprocal of the Strickler coefficient (K_s Strickler = 1/n manning), so the conversion between both is not a rocket science.

The **Gauckler–Manning coefficient n** is an **empirically derived coefficient**, which depends on **surface roughness, hydraulic radius** and **stream slope:**

$$v_m = k_{St} \cdot R^{2/3} \cdot I_E^{1/2} \quad \text{oder} \quad v_m = \frac{1}{n} \cdot R^{2/3} \cdot I_E^{1/2}$$

(Landesanstalt für Umweltschutz Baden-Württemberg, 2002, p. 12)

where applies:

v_m = average flow velocity
k_{St} = Strickler roughness coefficient
n = Gauckler-Manning roughness coefficient
R = hydraulic radius
I_E = stream slope (hydraulic gradient)
When field inspection is not possible (like in this work), the prevailing method to determine n or k_{St} is to use photographs of river channels where n has been determined using Gauckler–Manning's

formula. As I don´t have photographs either, and as I am examining overland flow on the floodplain instead of channel flow only, I used the available orthoimages to determine the Strickler coefficient.

The Strickler coefficient is common in German speaking countries and the possible value range is between 1 (extremely rough surface) and 100 (extremely smooth surface), whereas the Manning scale is between 0.01 for extremely smooth surfaces and 1.00 for an extremely rough surface, see Table 8.

Again, like in many fields in hydrodynamic engineering, there is not a single scale determining which surface / land cover has which exact Strickler or Manning coefficient, as this is empirical science where 10 authors are likely to have 10 (at least slightly different) determinations of which surface has which roughness coefficient.

Gerinne	**k_{St} [$m^{1/3}/s$]**
Natürliche Flüsse	
1) kleinere Flüsse mit HW-Abflußbreite < 30 m	
a) gleichmäßiger gerader Verlauf, kleine Gräben oder tiefe Mulden, bordvoll	30 - 40
b) gleichmäßig gewunden, einige Mulden und Untiefen	22 - 30
c) mit Stillwasserbereichen, verkrautet, tiefe Mulden	13 - 20
d) sehr verkrautete Bereiche, tiefe Mulden oder Vorländer mit dichtem Baumbestand und Unterholz	ca. 10
2) Gebirgsflüsse, ohne Vegetation im Flussbett, steile Böschungen, Bäume und Gebüsch entlang der Ufer bei HW überschwemmt	
a) Boden: Kies, Steinbrocken, einzelne Felsblöcke	20 - 35
b) Boden: Steinbrocken mit großen Felsbrocken	15 - 25
Vorländer	
1) Wiese, kein Gestrüpp	
a) kurzes Gras	30 - 40
b) hohes Gras	20 - 33
2) Gestrüpp	
a) verstreutes Gestrüpp, dichtes Unkraut	14 - 29
b) mittleres bis dichtes Gestrüpp, im Winter	9 - 22
c) mittleres bis dichtes Gestrüpp, im Sommer	6 - 14
3) Bäume	
a) dichte Weiden im Sommer	8 - 13
b) dichter Holzbestand, wenig Unterholz, HW-Stand unter den Zweigen	6 - 10
c) wie b) aber HW-Stand erreicht die Zweige	

Table 8: Empirical collection of Strickler roughness coefficients (Landesanstalt für Umweltschutz Baden-Württemberg, 2002, p. 48)

In this work, the used Strickler coefficients are very low (see Figure 59) due to the fact that almost all of the water in the model is surface runoff (instead of river or channel flow). The Strickler coefficients used for both models (Safer-RAIN and Hydro-As-2D / Sturzflut) are as follows:

- forest areas: 5
- shrubland: 9
- urban areas with low building density: 3

- urban areas with high building density: 6
- roads: 50
- water areas (sea): 100

These mostly very low Strickler coefficients may look not in row with the values stated in Table 8, but indeed, they have proven to be adequate during my 7-year practise in pluvial flood calculations at Humer consulting engineer via calibration of rainfall-runoff models where rain gauge data as well as discharge data have been available (Humer et al., 2015, seeTable 9).

Material	Mat. ID	HSG [1)]	CN	Strickler-Wert	Kf-Wert (mm/h)	Anfangsverlust (mm)	kont. Verlust (mm/sec)
Ackerland_clay loam	1	D	88	12	1	4,5	0,000278
Ackerland_loam	2	C	84	12	3,4	4,5	0,000944
Ackerland_sandy loam	3	B	76	12	10,9	4,5	0,003028
Ackerland_silt	4	C	84	12	4	4,5	0,001111
Ackerland_silt loam	5	C	84	12	6,5	4,5	0,001806
Ackerland_silty clay loam	6	D	88	12	1	4,5	0,000278
Industrie/Gewerbe_loam	7	C	95	25	3,4	0,8	0,000944
Industrie/Gewerbe_silt	8	C	95	25	4	0,8	0,001111
Industrie/Gewerbe_silt loam	9	C	95	25	6,5	0,8	0,001806
Wald_loam	10	C	73	6	3,4	8	0,000944
Wald_sandy loam	11	B	60	6	10,9	8	0,003028
Wald_silt	12	C	73	6	4	8	0,001111
Wald_silt loam	13	C	73	6	6,5	8	0,001806
Wald_silty clay loam	14	D	79	6	1	8	0,000278
Wasser_loam	15	C	100	200	3,4	0	0,000944
Wiesen_loam	17	C	79	18	3,4	3,5	0,000944
Wiesen_sandy loam	18	B	69	18	10,9	3,5	0,003028
Wiesen_silt	19	C	79	18	4	3,5	0,001111
Wiesen_silt loam	20	C	79	18	6,5	3,5	0,001806
Wiesen_silty clay loam	21	D	84	18	1	3,5	0,000278
Wiesen_clay loam	16	D	84	18	1	3,5	0,000278
Wohngebiet_loam	22	C	85	5	3,4	3	0,000944
Wohngebiet_sandy loam	23	B	78	5	10,9	3	0,003028
Wohngebiet_silt	24	C	85	5	4	3	0,001111
Wohngebiet_silt loam	25	C	85	5	6,5	3	0,001806
Wohngebiet_silty clay loam	26	D	89	5	1	3	0,000278

[1)] Hydrological Soil Group (abgeleitet aus der Bodentextur) zum Zwecke der Festlegung der CN

Table 9: Benchmark table for Strickler coefficients used in a flood model validation during my working experience at Humer consulting (Humer et al., 2015, p. 167)

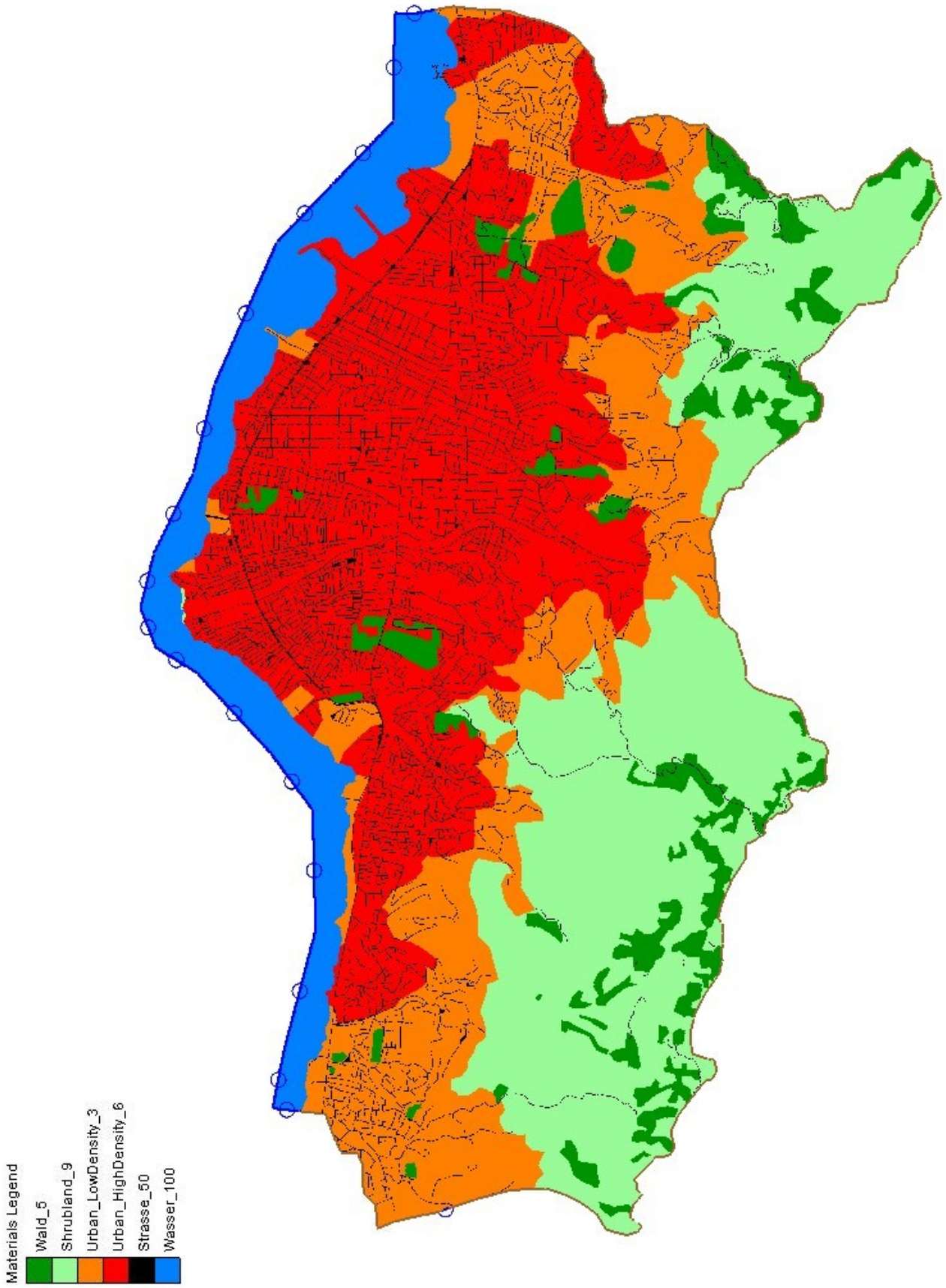


Figure 59: Strickler coefficients dependant on land cover used in both flood models (Hydro-As-2D / Sturzflut and Safer_RAIN)

3.4.4 Input: Vegetation, imperviousness and land cover

All these three parameters affect the roughness coefficient, but under real world conditions, the infiltration into soil as well. Normally, for the pluvial flood function of the hydrodynamic Hydro-As-2d model, there are two types of infiltration parameters:

- initial loss in mm (the proportion of precipitation percolating soil at the beginning of the rain event instead of running off as overland flow)
- infiltration rate in mm/h (the proportion of precipitation percolating continuously, not only at the very beginning of the rainfall duration.)

For Safer_RAIN, infiltration has to be deducted beforehand, model input is netto rainfall. In the case of our AOI, I have set the infiltration rates (both initial as well as continuous) intentionally to zero, as soil conditions are generally very poor here. Of course, there will be some infiltration in case of a heavy to extreme rain event like in the scenarios, but as imperviousness in the urban part of the AOI is high, soil seepage capacity and tree cover both are very low, the percentage of total precipitation which infiltrates the soil instead of turning into overland flow can be seen as neglectable (see Figure 60).

Furthermore, especially for extreme rain events, the role of forest / vegetation cover according to Posner and Georgakakos, 2017 remains scientifically controversial. The authors argue that there would be an increasing body of evidence that while forests may do a great job in reducing floods for small to moderate storms, this effect is increasingly reduced as precipitation events increase in magnitude (Posner and Georgakakos, 2017) Last but not least, the model assumption of not determining a fixed infiltration rate facilitates the comparison between Safer_RAIN and Hydro-As-2d.

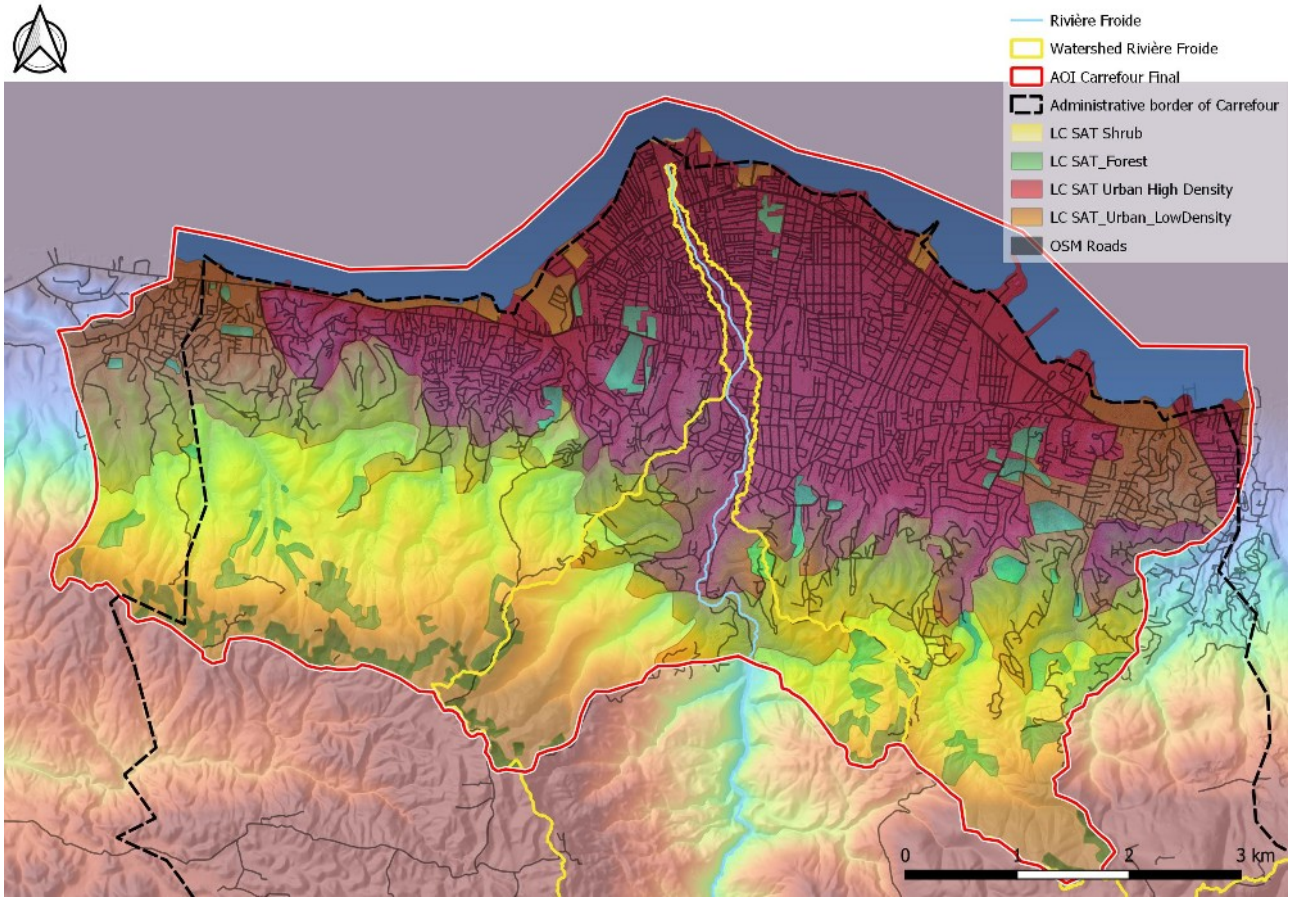


Figure 60: land cover in the AOI Carrefour, derived from OSM (street network) and orthoimages (all other land cover classes). Note the very low forest share and high degree of urban areas with high impervousness.

3.4.5 Hydrodynamic 2d flood model: “Sturzflut” extension for software “Hydro-As-2d”

There is not one methodology for running a hydrodynamic flood model (1D, 2D, coupled 1D/2D, finite volumes method vs. finite elements method vs. finite difference method,..) and to state the whole theoretical knowledge base of the underlying numerical mathematics on the solving of partial differential equations (namely the shallow-water equations) would go far beyond the scope of this work, therefore, reference is made in this context to the corresponding technical literature (i.e. Baumann et al., 2006).

But a very short theory discussion on the used approach for the modelling shall be given here. In the year 2015, during my professional work expertise at Humer consulting engineer in the province of Upper Austria, I was involved in the validation (Humer et al., 2015) of an extension called “Sturzflut” (meaning “flash flood”) for the then current version of the 2d hydrodynamic flood simulation software “Hydro-As-2d”, developed by the Unit of Hydraulic Engineering at University of Innsbruck one year before (Klar et al., 2014).

Hydro-As-2d uses the finite volumes method (FVM) for the solving of the shallow-water-equations. The FVM is a simulation method in which the calculation area is divided into discrete small

volumes during discretisation. The simulation focuses on the balance equation for each of these volumes (Baumann et al., 2006). The advantage of the finite volumes method is a maximum in stability during the computation process, also in case of steep and rapidly changing gradients.

At that time, when the extension had been developed (2014), most 2D hydrodynamic models did not allow for distributed rainfall as input nor were there any types of soil/surface interaction implemented as in hydrological models (that has changed in the meantime, new 2d models like JFLOW have been developed and others like Hydro-As-2d have integrated this ability in newer versions).

Therefore, the 2D hydrodynamic model of Hydro-As-2d (depth-averaged flow equations using the finite volume discretization), was extended to accept direct rainfall enabling to simulate the associated runoff formation. Rainfall is introduced via the modification of water levels at fixed time intervals (Klar et al., 2014, see Figure 61).

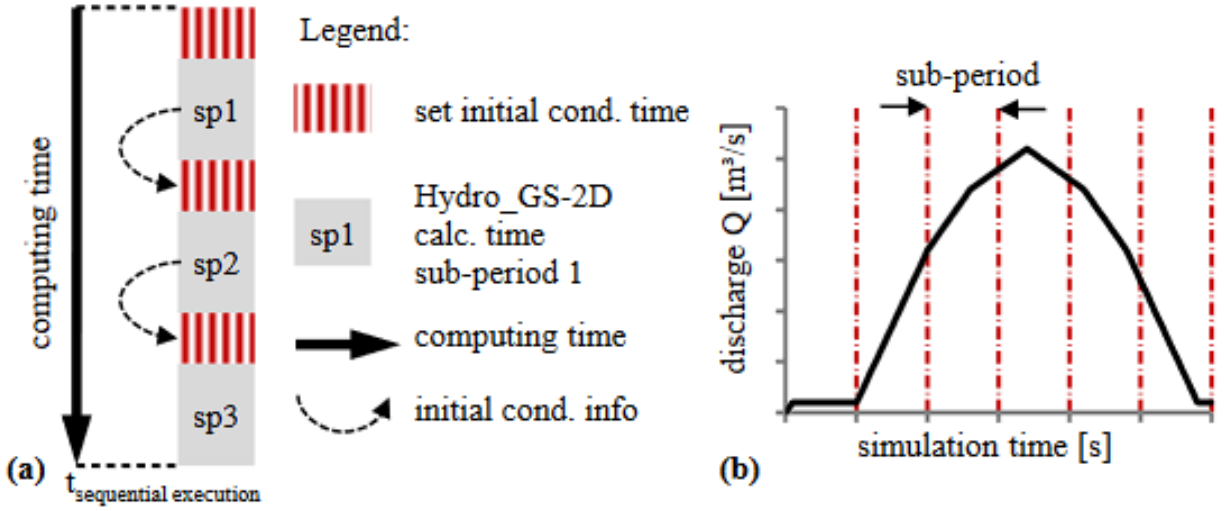


Figure 61: technical functionality of the Hydro-As-2d extension “Sturzflut”: (a) Temporal decomposition approach; (b) exemplary inflow hydrograph with indicated sub-period boundaries (Klar et al., 2014, p. 3)

In this way, it is possible to add rainfall at the beginning of every internal time step (in the range of some seconds) of the 2d model – the script stops the hydrodynamic flood calculation after each time step, adds rainfall amount X, which is defined via an input table, deducts infiltration rates (spatially variable), if they are defined in the input, and then starts the simulation again – until it stops again after the next time step, for adding rainfall and deducting infiltration and the loop continues, until the end of the calculation time.[54] **The rainfall added into the hydrodynamic calculation model can be spatially and temporarily variable**.

It was my own work in the year 2015 to validate this extension in two small catchments (with an area of 24 km² and 30 km²) in Upper Austria, were we had rain gauge data with 5 min temporal resolution and discharge curves of two gauges available, which acted as the outlet of the catchments. With the Roughness Coefficients and loss rates stated in Table 9, I managed to get the

54 The chosen rainfall duration is in most cases lower than the chosen calculation time, as normally, you want to go sure that at least the peak of the flood wave has reached the bottom of the valley or catchment – otherwise, if the calculation stopped before,

following results (see Figure 62), which fit the discharge curve of the two gauges (thick dark blue line) quite well:

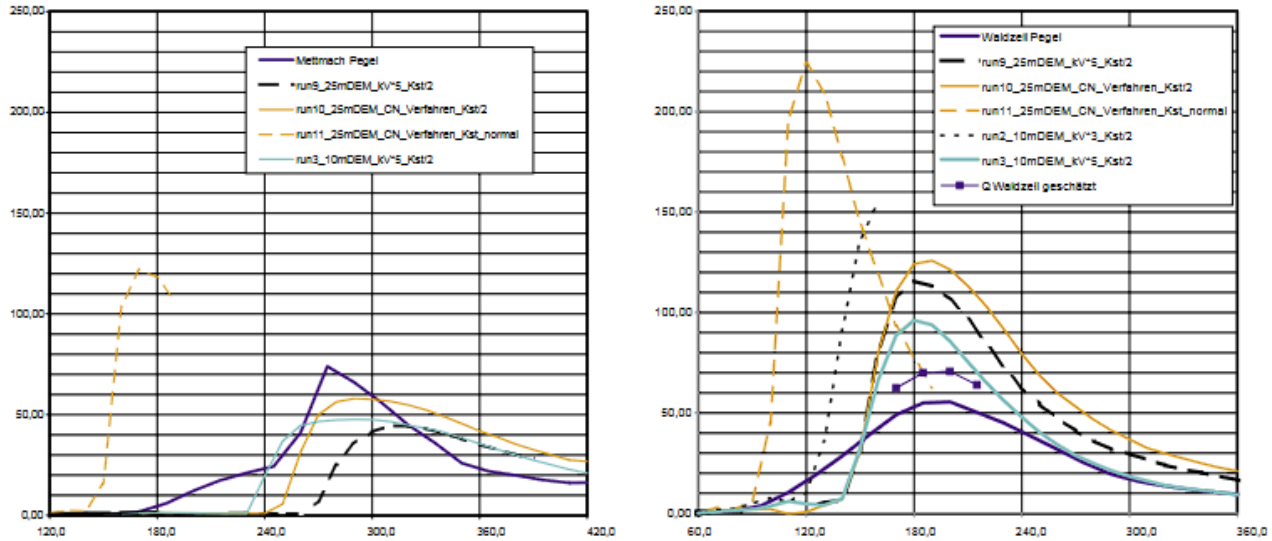


Figure 62: Discharge curves: gauges of Waldzell and Mettmach (AT, dark blue thick line))vs. discharge curves attained in the "Sturzflut" validation runs (Humer et al., 2015, p. 168)

3.4.6 GIS-based flood model: Filling-and-spilling algorithm “Safer_RAIN”

Given the steadily increasing availability of LiDAR high-resolution DEMs, several studies highlighted the potential of fast-processing DEM-based methods, such as the Hierarchical Filling-&-Spilling algorithm (HFSA). Samela et al., 2020 developed a fast-processing HFSA, named Safer_RAIN, that enables mapping of pluvial flooding in large urban areas by accounting for spatially distributed rainfall input and infiltration processes through a pixel-based Green-Ampt infiltration model.

> *The idea behind HFSAs is to identify pluvial-flooded areas on the basis of nested surface depressions extracted from high-resolution DEMs; the volume of rainfall is accumulated in depressions and, as they are filled, water spills downstream to depressions located at lower elevations. HFSAs can be classified as “non-source flooding”, meaning that all points where the elevation is below a given water level belong to the flooded area. (Samela et al., 2020, p. 4)*

Figure 63 shows the workflow for flood modelling via the “Safer-RAIN” algorithm:

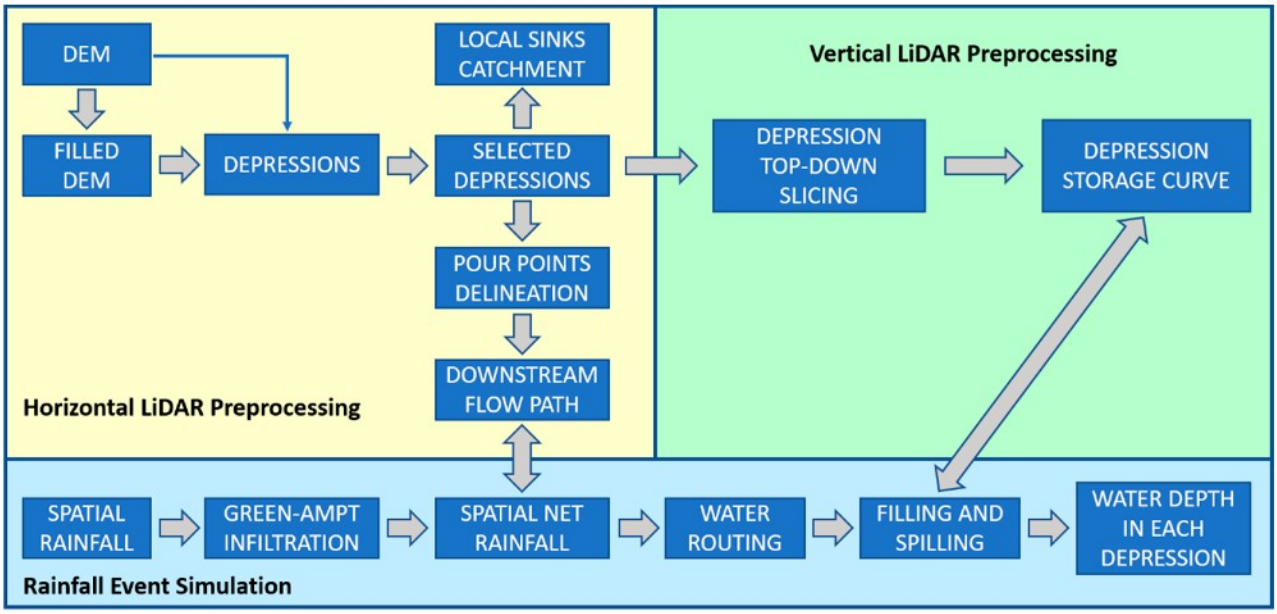


Figure 63: Safer_RAIN work flow, distinguishing between DEM preprocessing (on top) and pluvial flood simulation (below) (Samela et al., 2020, p. 4)

The “Safer_RAIN” algorithm is based on the following assumptions:

- Drainage direction is identified according to the D8 method (see Figure 64)
- Overland flow dynamics is neglected
- Net-rainfall volume accumulates into the nested depression system according to the capacity and hierarchical structure of depressions themselves
- Spatial distribution of rainfall may be variable
- Terrain may be considered impermeable, but infiltration rates may also be defined

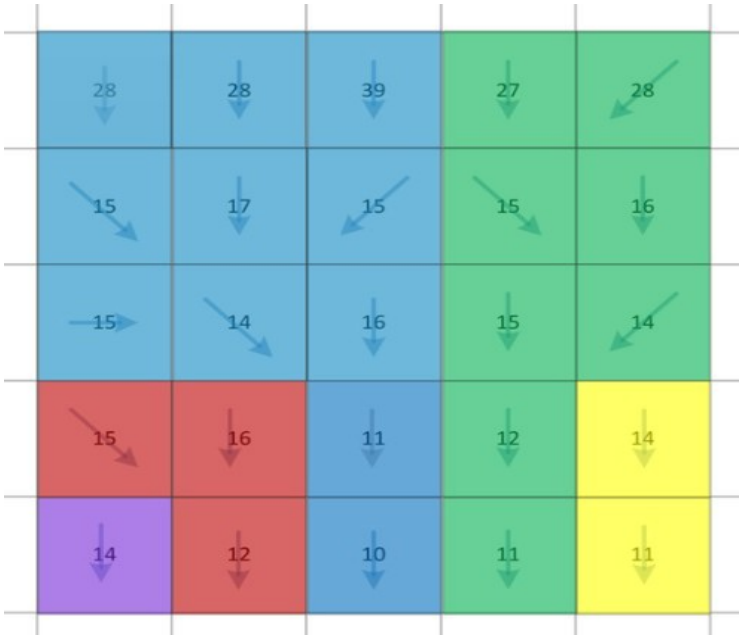


Figure 64: D8 algorithm flow directions (in use in “Safer-RAIN”, (Wang et al., 2014, p. 3)

The DEM preprocessing requires the identification of: (i) first-level depressions (here referred to as blue-spots) and depression volumes (i.e. depressions G and H in Figure 65) through the application of a DEM pit-filling algorithm; (ii) blue-spot pour-points (i.e., the lowest-lying cell along the depression upper edge), from where water would pour out if the depression is filled up with water; (iii) blue-spot contributing watersheds; (iv) the derivation of the upstream/downstream relationships and flow paths between pourpoints according to the D8 method (Samela et al., 2020).

“Safer_RAIN” implements an original flooding phase that identifies partially flooded areas for a given rainfall volume by considering all the intermediate levels (see Figure 66). This is achieved via a bottom-up level-set method used for quantifying partial filling in nested higher-level depressions:

(i) the method starts from an empty condition (see step (a) in Figure 66 the blue-spot is empty); (ii) according to their vertical hierarchy, higher-lever nested depressions are gradually filled from the bottom; (iii) gradual filling is performed step-by-step, considering in turn depressions with same hierarchical order until they are completely filled (Samela et al., 2020).

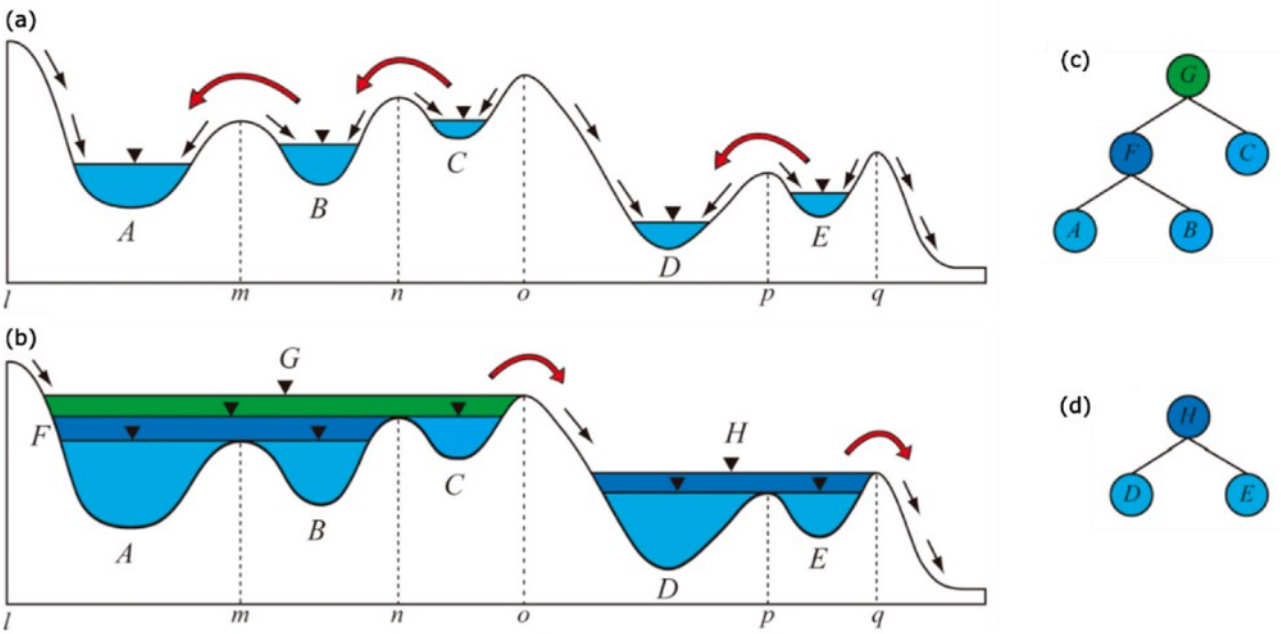


Figure 65: Schematic representation of the “Safer_RAIN” HFSA: (a) nested-depression system, and (b) blue-spots G and H under fully-filled conditions with illustration of the second-level depressions (F, C, D and E) and third-level depressions (A and B) (Samela et al., 2020, p. 5)

According to (Samela et al., 2020), “Safer_RAIN” is able to compute flooded areas very fast, as DEM preprocessing is run only once and it fully characterizes the hierarchy for filling and spilling processes. Considering that, differently from a hydrodynamic model, it does not model flood dynamics, it can produce an underestimation of maximum water levels; moreover, it does not allow for obtaining indication on timing and velocity.

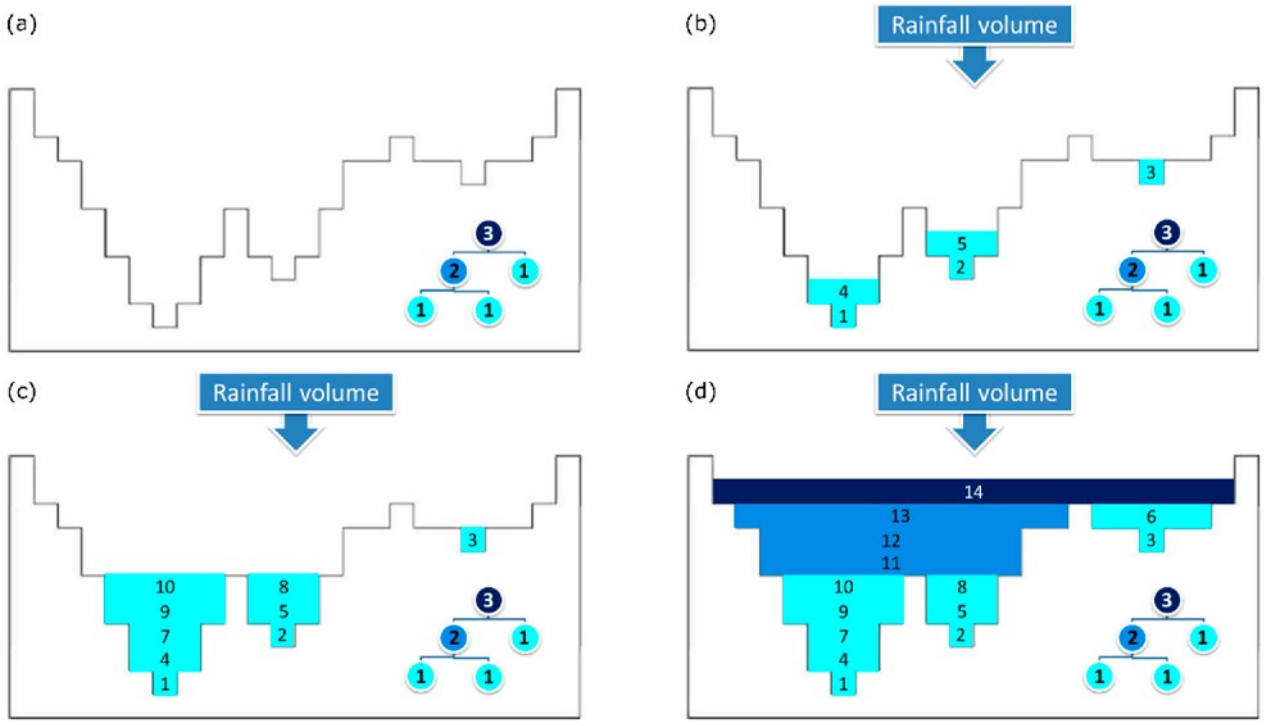


Figure 66: Example of four different intermediate steps (from a to d) in the application of the bottom-up level-set depression flooding phase: dry domain (a), bottom-up level-set methods for partial filling, while simulating the flooding resulting from a given rainfall event (c,d). (Samela et al., 2020, p. 6)

This makes a very important difference to all hydrodynamic flood calculations: There is no flow velocity in the model, and therefore, I can neither get any estimation on when the flood peak will reach grid cell or mesh node XY, nor can I produce a map with “maximum depth” of each all output intervals (i.e. 5, 10 or 15 minutes), where the peak flood extent (peak inundation depth) all all output intervals is shown in one map, pretending they are happening all at the same time (which is not possible of course).

But this theoretical state of a simultaneous maximum water depth (and therefore also flood extent) is a key visual aid for assessing flood prone areas. There is an equivalent in the HFSA approach, which is the final state of the calculation run (flood extent + water depth at the last time step / output interval), but especially in areas with a certain gradient, this can be quite misleading, when water is routed down via D8 flow algorithm from grid cell to grid cell and there are no or hardly any depressions in this area. Let´s keep this important fact in mind, when we take a look at the results in chapter 4.

3.4.7 Benchmark calculation with a raster-based 2d hydrodynamic software: LISFLOOD-FP 8.0

Before choosing Hydro-As-2d with the “Sturzflut” extension for the calculation of the 2d hydrodynamic benchmark for the Carrefour AOI, the idea was to use an also raster based approach, so that I could evaluate on raster-based flood algorithm against the other. The tool of choice was LISFLOOD-FP, in its current version 8.0. I knew the flood software from a test run against Hydro-

As-2d during my time at Humer Engineering, and in the small catchment in Upper Austria where I tested it, the results had been quite good and fitted well with the Hydro-As-2d results.

But, how life plays out, the results in this case (AOI Carrefour) have been totally different, and most notably simply unusable (see chapter 4.1.4). But let´s start at the beginning and state some key facts on LISFLOOD-FP. The software is a pioneer in the area of raster-based 2d hydrodynamic flood calculation; i.e. the nowadays successful hydrodynamic 2d flood software "JFlow", raster-based as well, has its mathematical roots in LISFLOOD-FP. Lamb et al., 2009 have found a way how to bring the abilities of LISFLOOD-FP into a very fast and efficient computation via massive parallelization on GPU cores. Like some of the possible computation variants in LISFLOOD-FP (i.e. the "Acceleration" solver, which had been used here in this work for the AOI), JFlow

> *"..uses a simplified form of the full 2D hydrodynamic equations, but captures the main controls on flood routing for shallow, topographically driven flow. It is thus easier to set up and quicker to run than full 2D hydrodynamic codes."(Bradbrook, 2006, p. 79)*

In its version 8.0, LISFLOOD-FP offers both full 2d hydrodynamic flood calculations regarding all terms of the shallow-water-equations (solvers 'FV1' and 'DG2'as well as 'Roe', see Table 10) as well as several simplified, less CPU-intensive calculation methods (solvers 'Routing', 'Flow-limited', 'Adaptive' and 'Acceleration'), which neglect some or even all ('Routing') terms of the shallow-water-equations (SWE). It shall be noted in this context, that even when two solvers both make use of all terms of the SWE, results can be different nevertheless, as it is also important which approach of solving the partial differential equations in hydraulics they use – finite volume or finite elements.

Solver	Dimensions	Shallow water terms included	Shallow water terms assumed negligible	Time step	Further technical details
Routing	1D on 2D grid	User specified velocity and bed slope direction only	All	Adaptive	Sampson et al., 2012
Flow-limited	1D on 2D grid	Friction and water slopes	Local and convective acceleration	Fixed	Bates and De Roo, 2000
Adaptive	1D on 2D grid	As above	As above	Adaptive	Hunter et al., 2005
Acceleration	1D on 2D grid, friction terms in 2D	Friction and water slopes, local acceleration	Convective acceleration	Adaptive	Bates et al., 2010; De Almeida et al., 2012
FV1	2D	All terms	None	Adaptive	
DG2	2D	All terms	None	Adaptive	
Roe	2D	All terms	None	Adaptive	Neal et al., 2012b

Table 10: Available solvers for floodplain (overland) flow calculation in LISFLOOD v8.0 (user manual LISFLOOD-FP, v 8.0, p. 11, available online at https://publications.jrc.ec.europa.eu/repository/bitstream/JRC78917/lisflood_2013_online.pdf, access on 2023-04-10)

Regarding the effectiveness and performance of the single solvers, Table 11 shows the differences in computation time in dependence of the chosen solver:

	$\Delta x = 50$ m 57 000 elements	$\Delta x = 20$ m 850 000 elements	$\Delta x = 10$ m 1.7 million elements
ACC	20 s	466 s (8 min)	1779 s (30 min)
FV1-CPU	22 s	739 s (12 min)	2188 s (36 min)
FV1-GPU	19 s	145 s (2 min)	965 s (16 min)
DG2-CPU	788 s (13 min)	4133 s (69 min)	33 009 s (9 h)
DG2-GPU	448 s (7 min)	2304 s (38 min)	13 606 s (4 h)

Table 11: Solver runtimes at grid spacings of 1x = 50 m, 1x = 20 m and 1x = 10 m. ACC, FV1-CPU and DG2-CPU solvers are run on a 16-core CPU; FV1-GPU and DG2-GPU solvers are run on a single GPU. (Shaw et al., 2021, p. 3591)

4. Results

4.1. Hydrodynamic model (Hydro-As-2D)

All 2d hydrodynamic flood calculation model runs have been carried out with the same calculation time of 120 minutes, which is exactly double the duration of the design precipitation event (60 minutes). By this means, it is ensured that the peak of the flood wave resulting from the design rainfall event has reached the bottom of our AOI, which is the shoreline to the Caribbean sea in our case.

The rainfall had been added intentionally in the simple and hypothetical form of a "**block rainfall**", which characterized as a **hypothetical rainfall event, where the intensity stays the same for the whole duration, from minute one to minute 60** (see Figure 67). **Even though this is a mere model assumption, it facilitates the comparability between the two flood calculation approaches** (2d hydrodynamic versus the merely GIS-based hierarchical filling-and-spilling algorithm), as the HFSA ("Safer_RAIN") is not able to calculate temporarily variable rainfall intensities, so for the HFSA, every rainfall input is always block rainfall.

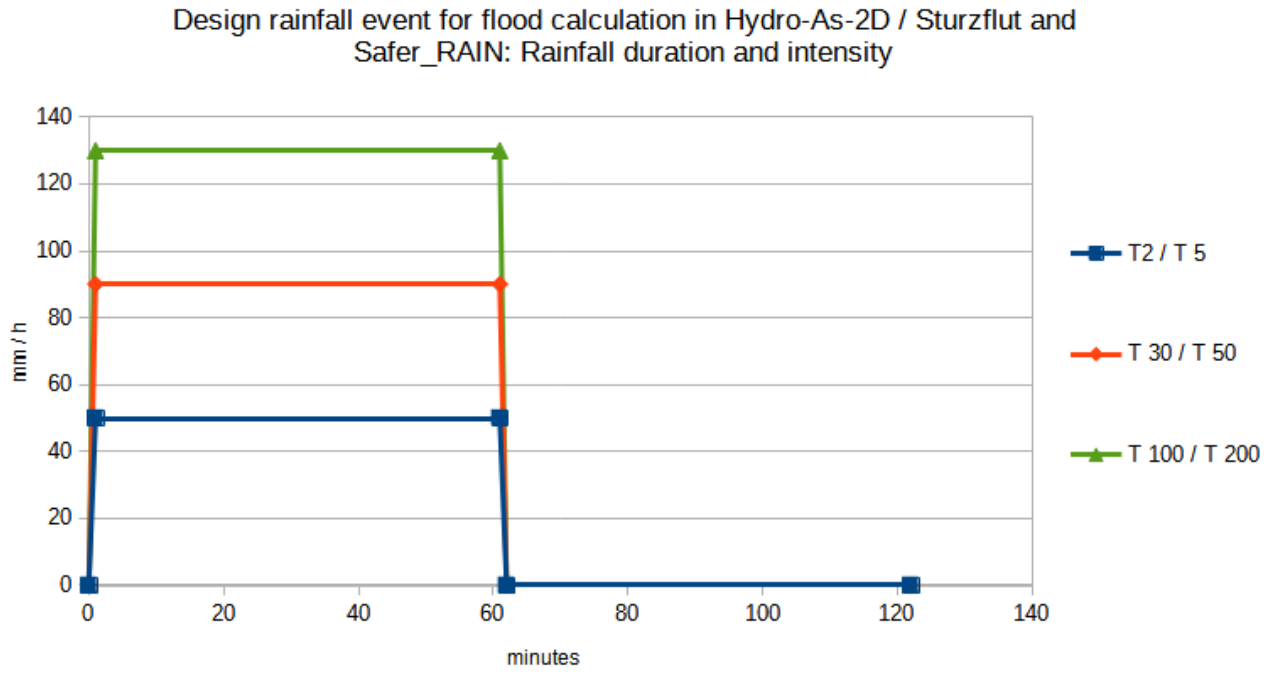


Figure 67: Intensity and duration of the three rainfall scenarios calculated in both models: frequent rain event with return period T2/T5, 50 mm in 60 minutes; rare rain event with return period T30/T50; 90 mm in 60 minutes and extreme rain event with return period T100/T200, 130 mm in 60 minutes

In the "SaferPLACES" pilot sites platform[55], where the operating team of Italian consulting engineer company GECOsistema srl[56] kindly added the AOI Carrefour to its test sites[57], the rainfall duration is 60 minutes with no additional hour for all water in the model to reach the bottom (shoreline), as the HFSA automatically calculates as long as the rainfall volume added into the model has reached its final stage – so either into various smaller and bigger pits (see Figure 66) or out of the model domain (the borderline of the area is open for passing water in the model, so no accumulation of water in front of an imaginary model borderline).

In the 2d hydrodynamic flood calculation with Sturzflut / Hydro-As-2d, from minute 60 to minute 120, no additional rainfall is added any more into the model, but the existing water still continues its way downwards. For hydrodynamic flood models, this is very important, as the flood wave should reach (respectively pass) the very bottom of the model area, for that the maximum water depth and flood extent can be delineated.

The following maps of the flood results show

- **Frequent rain event** (T2 / T5, 50mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), **whole AOI** (Figure 68)
- **Frequent rain event** (T2 / T5, 50mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), **Carrefour urban area** (Figure 69)
- **Rare rain event** (T30 / T50, 90mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), **whole AOI** (Figure 70)
- **Rare rain event** (T30 / T50, 90mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), **Carrefour urban area** (Figure 71)
- **Extreme rain event** (T100 / T200, 130mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), **whole AOI** (Figure 72)
- **Extreme rain event** (T100 / T200, 130mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), **Carrefour urban area** (Figure 73)

Summarizing, the simulation of the rainfall events with Hydro-As-2D and the “Sturzflut” extension worked very well, the results look realistic and give a good impression of which areas are in danger at which rainfall intensity (and also when the peak inundation will be reached, if one looks through the results of each output interval).

Flow velocity is **calculated in the simulations as well**, but as there is **no counterpart in Safer-RAIN** to compare it to, the velocity model output (although important for the practitioner, as flow velocity has a major impact on possible flood damage in steep areas) has **not been reviewed for this work.**

55 http://54.164.155.9/webgis/ (access on 30-03-2023)
56 https://gecosistema.com/ (access on 30-03-2023)
57 http://54.164.155.9/webgis/Haiti/ (access on 30-03-2023)

The results also show clearly that the **main inundation risk in Carrefour** is indeed **not coming out of the river Rivière Froide, but from the small *ravines*** (narrow streams which are dried-up during most of the time and only lead water after rainfall events) with a rather small catchment area between 0.5 and 5 square kilometres, but a mostly steep to very steep slope and therefore a very quick response to every precipitation event (see figures 68 until 73 on the next pages).

4.1.1 Frequent rain event (T2/T5, 50 mm in 60 minutes)

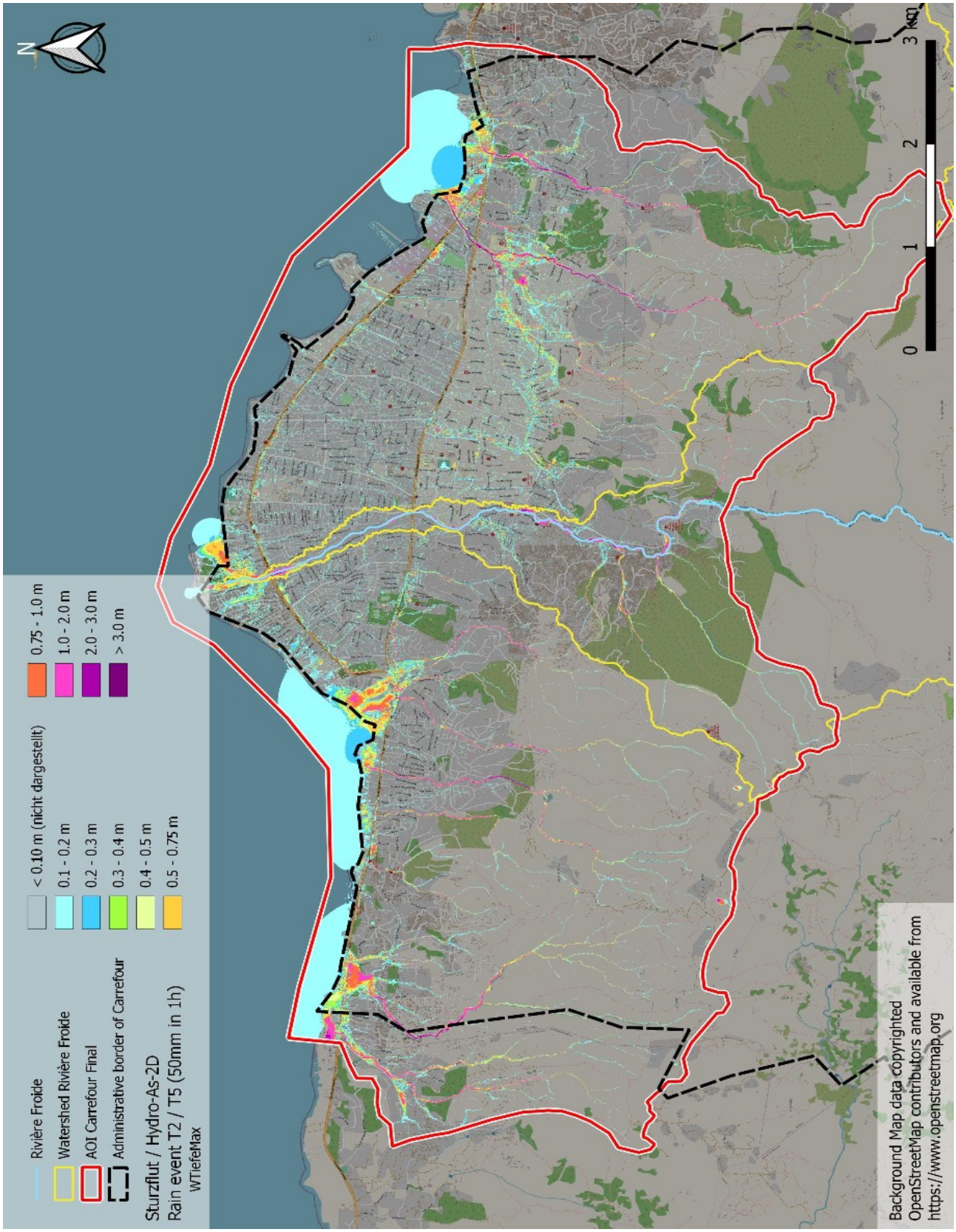


Figure 68: Frequent rain event (T2 / T5, 50mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), whole AOI

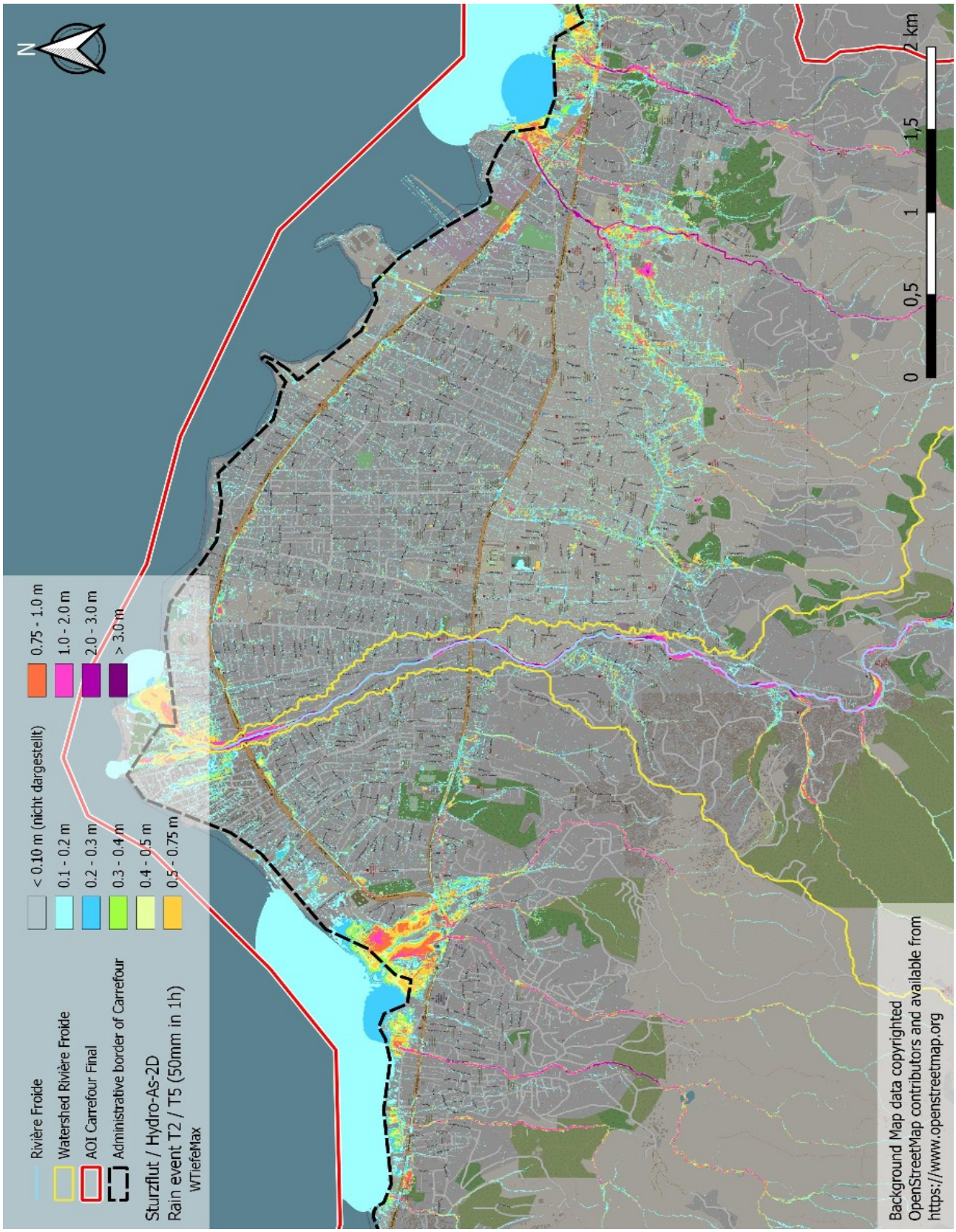


Figure 69: Frequent rain event (T2 / T5, 50mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), Carrefour urban area

4.1.2 Rare rain event (T30/T50, 90 mm in 60 minutes)

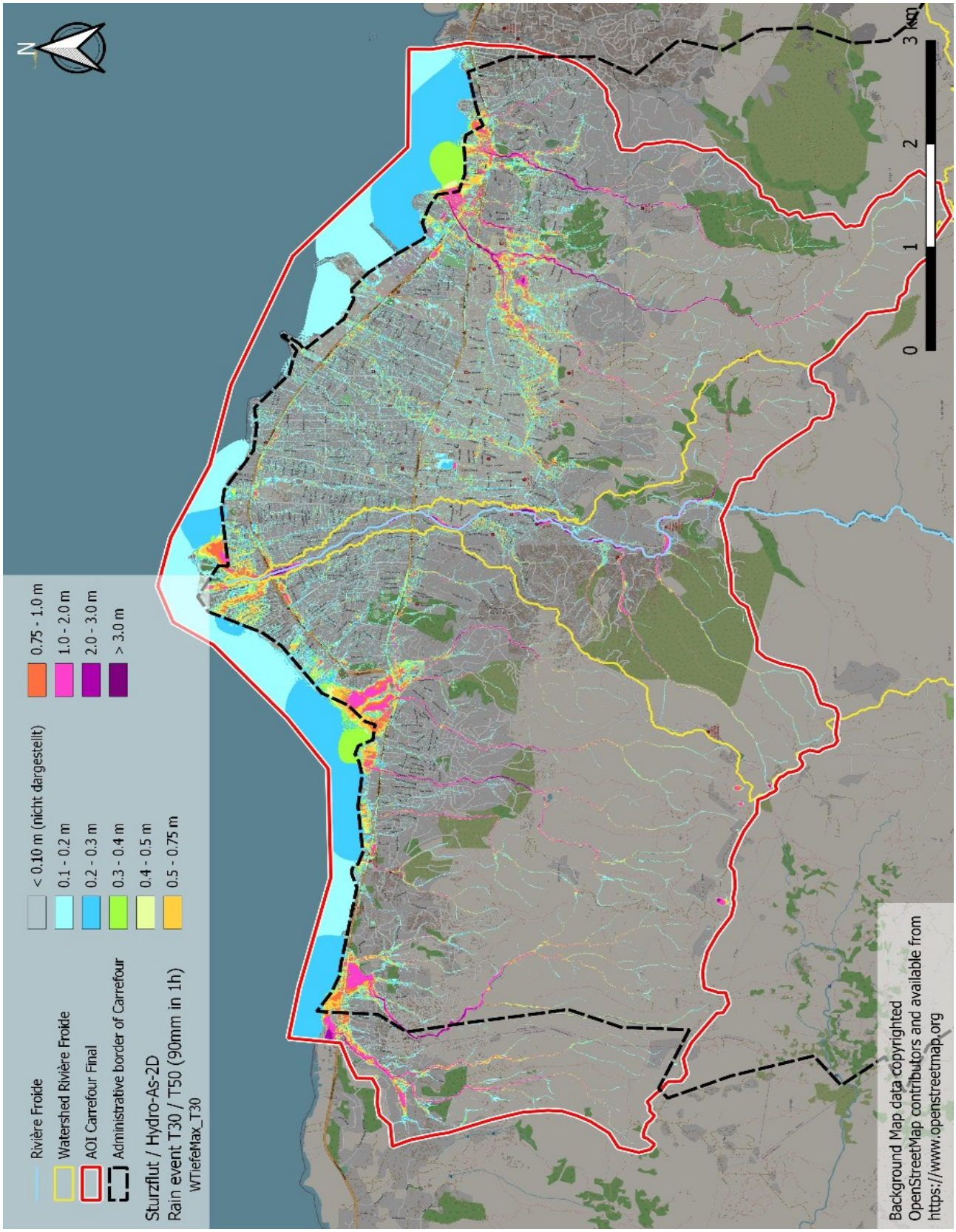


Figure 70: Rare rain event (T30 / T50, 90mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), whole AOI

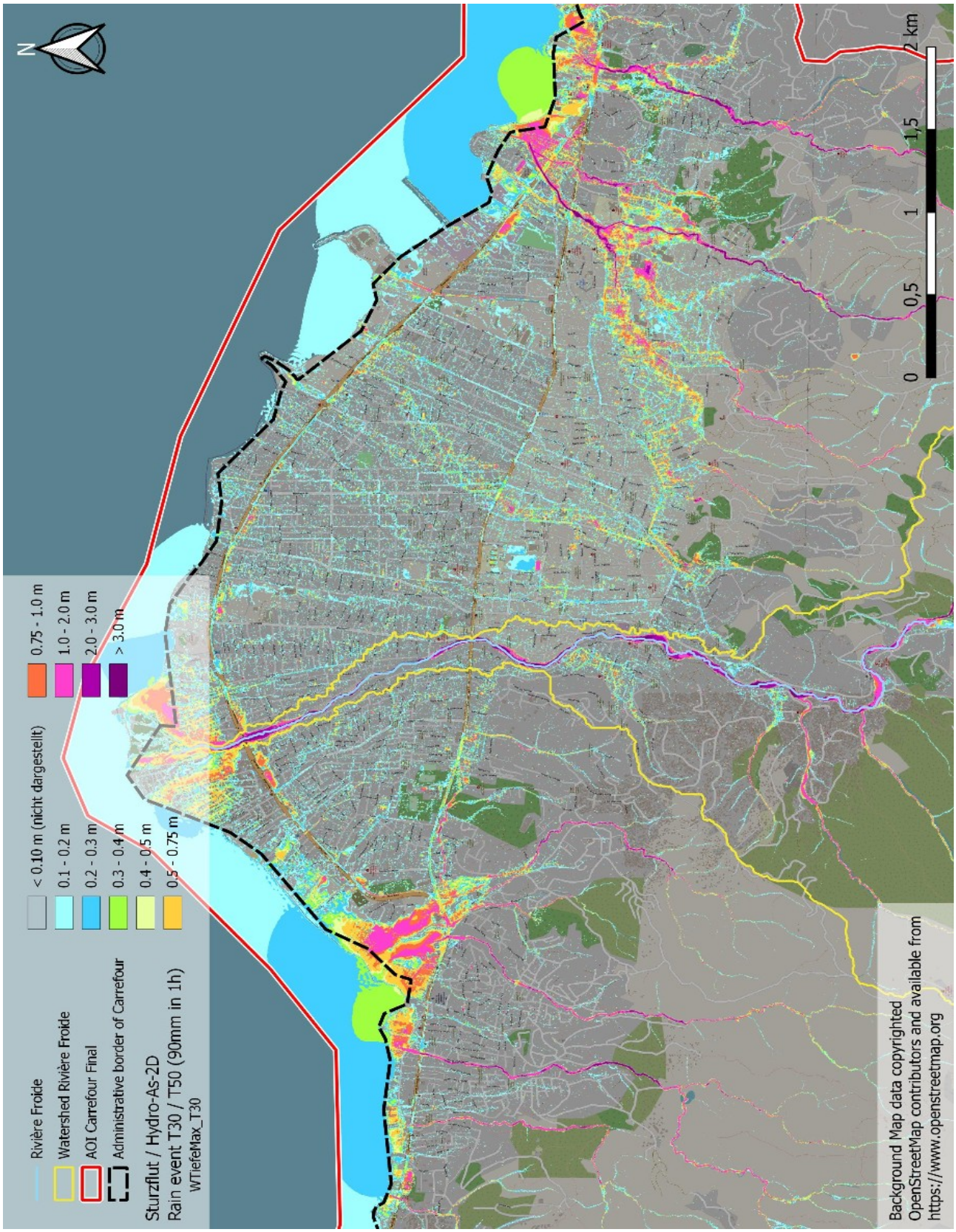


Figure 71: Rare rain event (T30 / T50, 90mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), Carrefour urban area

4.1.3 Extreme rain event (T100/T200, 130 mm in 60 minutes)

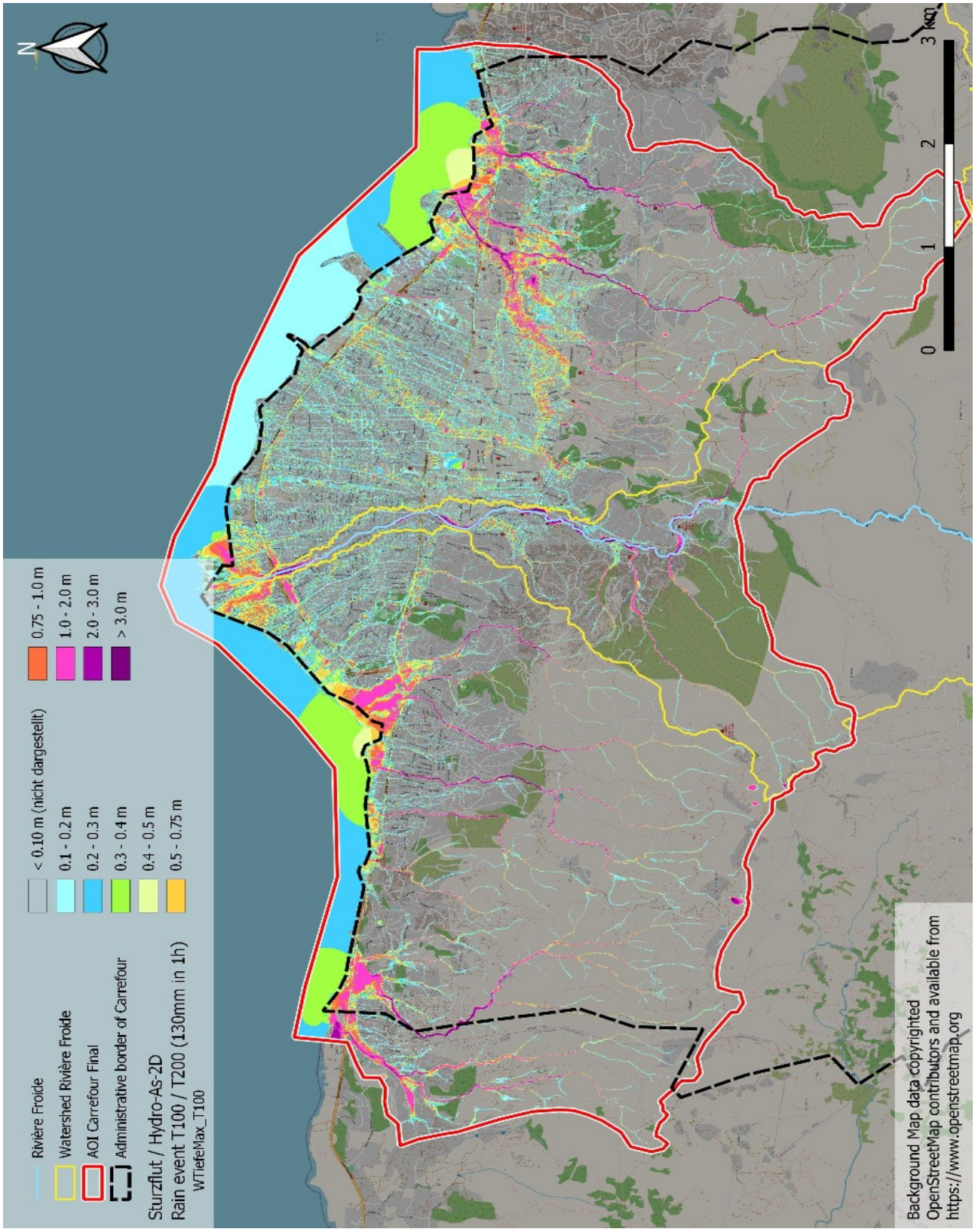


Figure 72: Extreme rain event (T100 / T200, 130mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), whole AOI

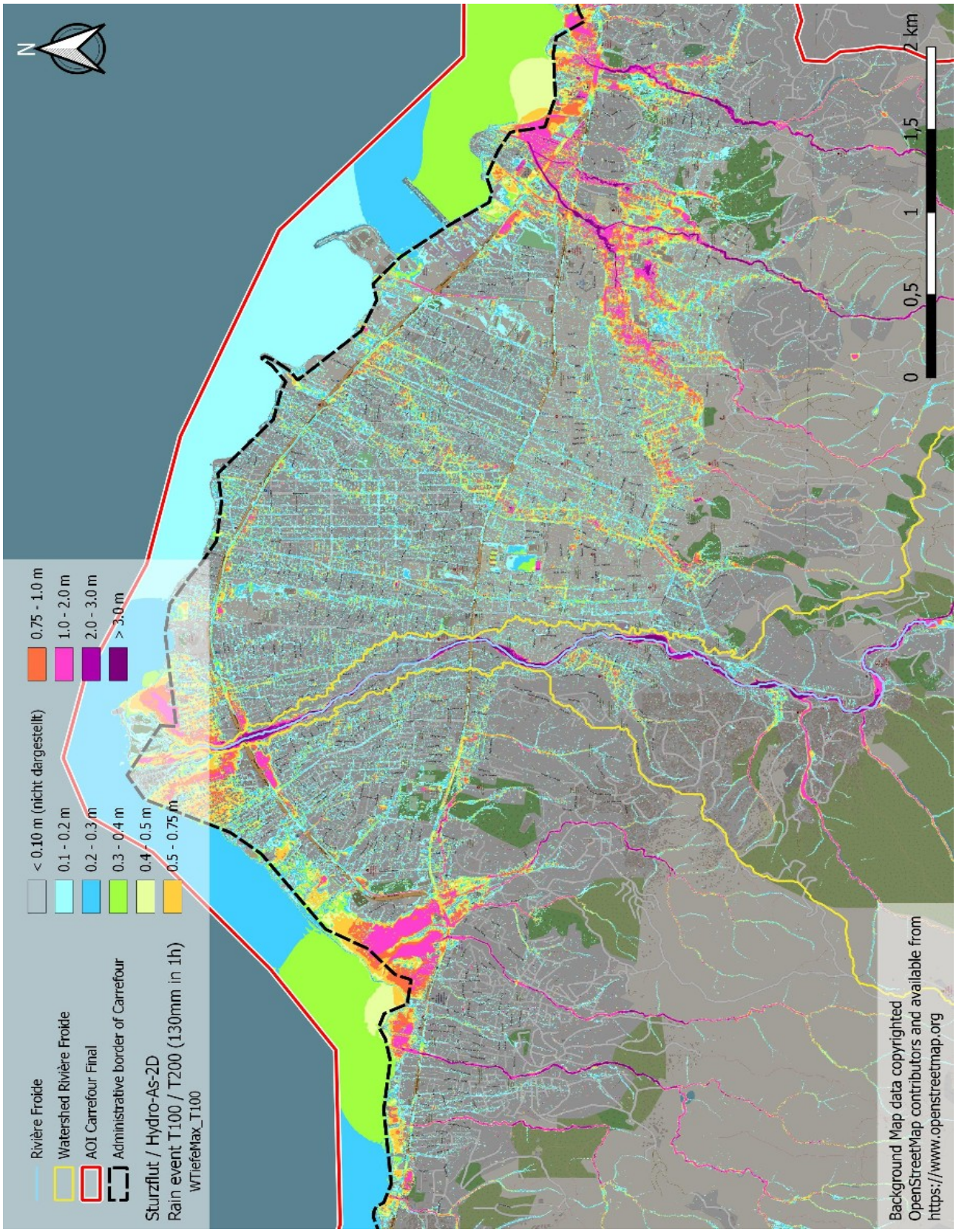


Figure 73: Extreme rain event (T100 / T200, 130mm in 1h): maximum water depth of all time steps (calculation in Hydro-As-2D / Sturzflut extension), Carrefour urban area

4.1.4 Benchmark calculation with “LISFLOOD-FP”

As already mentioned, the solver used for the test runs in LISFLOOD-FP had been the “Acceleration” solver, which I had used before in my work practice and where I knew how to proceed and prepare the dataset. “Acceleration” neglects the term of the convective acceleration in the SWE, whereas the other terms (friction, water slopes, local acceleration) are included, see Table 10 in Chapter 3.4.7.

But unfortunately, this solver has a methodological problem when used in steep reliefs in combination with a high amount of precipitation – the model became highly unstable after ~4 minutes rainfall with an intensity of 92 mm/h in my test runs, also with other rainfall intensities and durations (with 50 mm/h as minimum intensity), see Figure 74 and Figure 75.

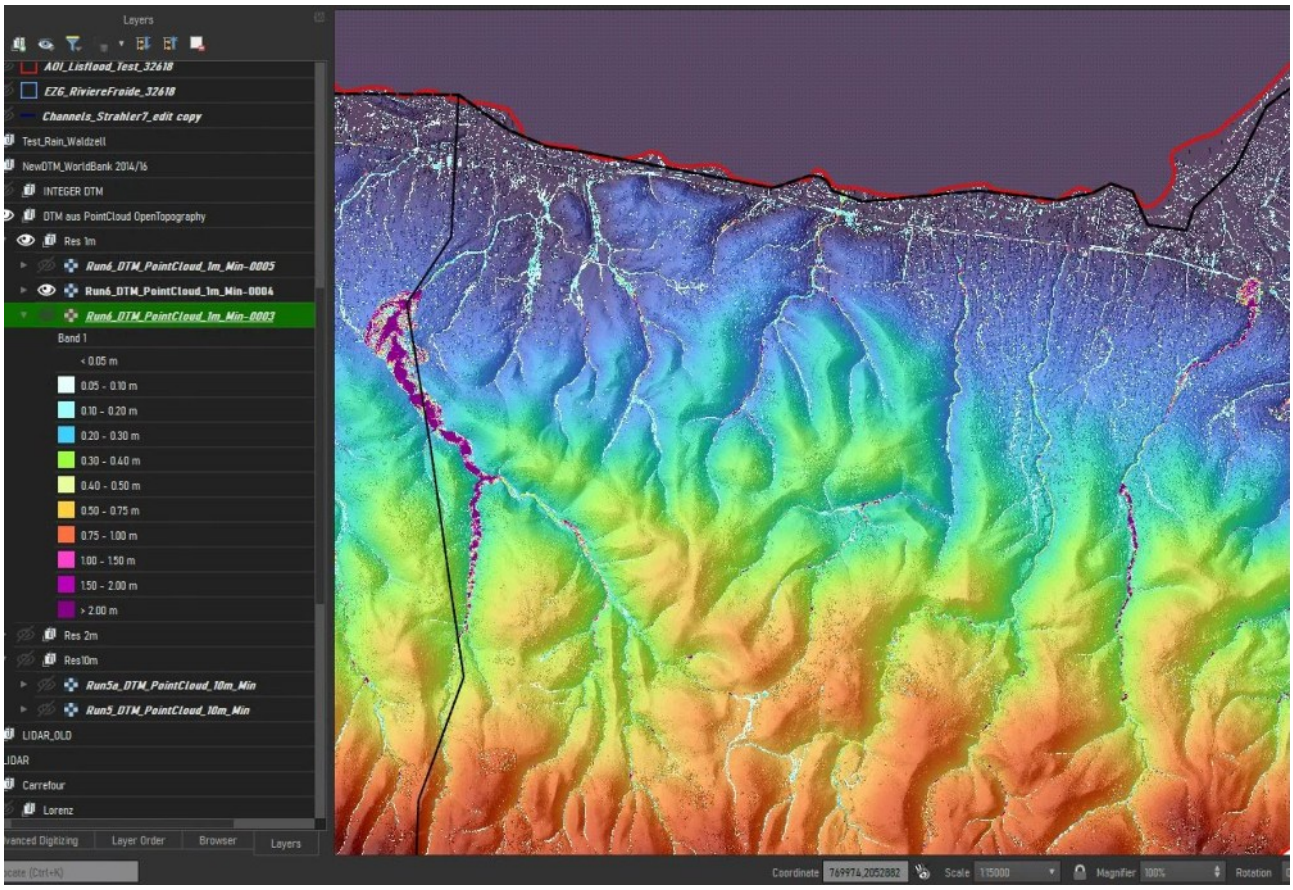


Figure 74: Testrun with LISFLOOD v8.0, I= 92mm/h, solver 'Acceleration': model beginning to become unstable at output interval 3 (= after 15 minutes), beginning to produce big "flood waves"

So, although I had already used LISFLOOD-FP in my engineering practise with the same solver algorithm successfully in other (flatter) areas with similar rainfall intensities, I came to the conclusion that, for steep reliefs like it is the case in many areas of Haiti and also in our AOI Carrefour, LISFLOOD-FP with the “Acceleration” solver cannot be the tool of choice.

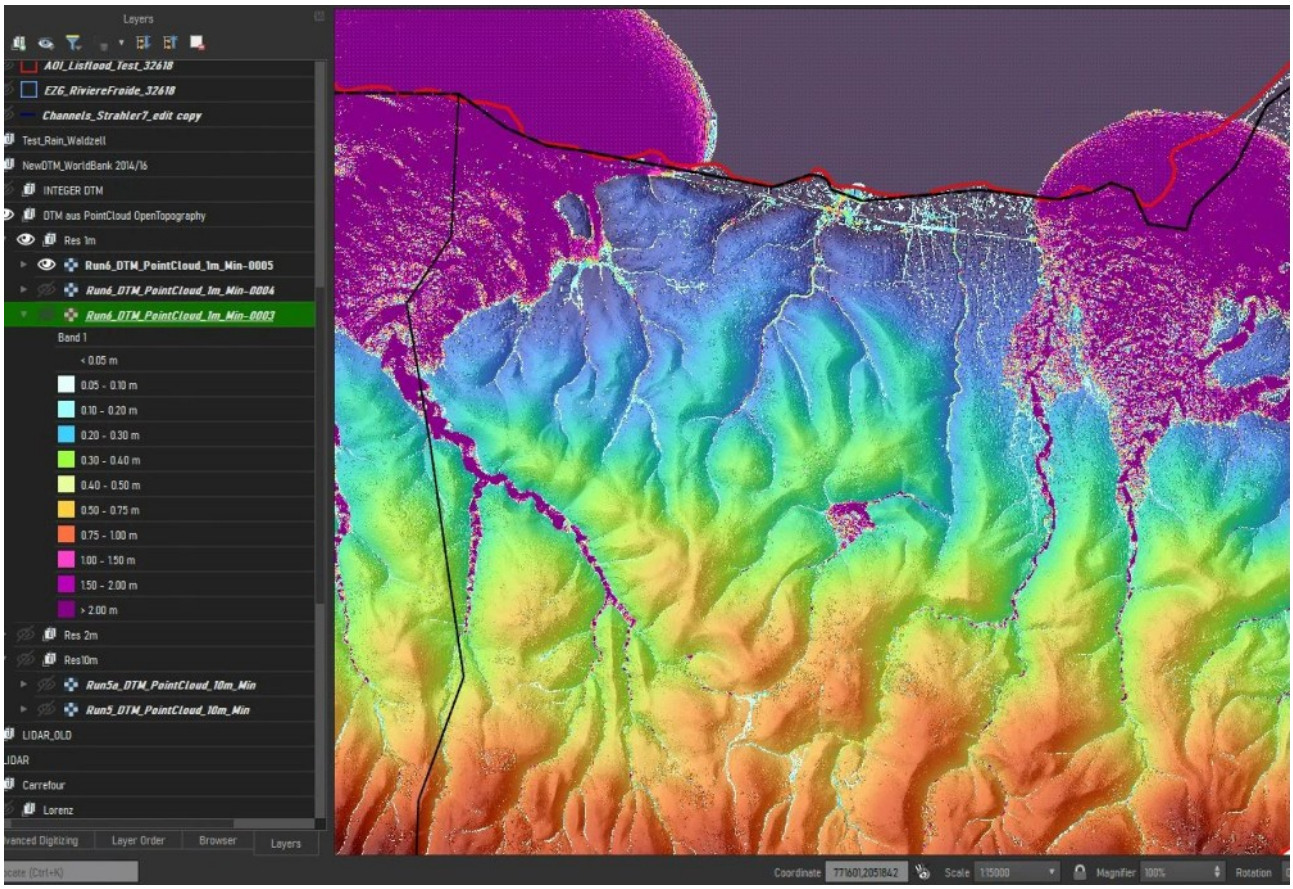


Figure 75: Testrun with LISFLOOD v8.0, I= 92mm/h, solver 'Acceleration': model highly unstable at output interval 4 (= after 20 minutes), producing enormous "flood waves"

4.2. GIS-based model (HFSA “Safer_RAIN”)

On the "SaferPLACES" Web platform for all pilot sites[58] (most of them being cities in Europe, the project is co-funded by the “Climate-KIC” initiative from the European Commission), the AOI Carrefour had been uploaded and can be tested there by oneself (see Figure 76), with a basic knowledge on GIS and hydrology sufficient to run an instant calculation on your own. The calculation itself is very quick, the results are ready in some minutes, or even seconds (depending on the size of the area and the rainfall amount).

So we have a **very clear advantage in calculation velocity for “Safer_RAIN”** compared to the 2d hydrodynamic calculation in Hydro-As-2D, which lasted for about 48 hours for each scenario run. Even clearer becomes this advantage when using the global platform (Figure 77):

If we suppose, we have no ready-to-use raster or mesh dataset with the DTM, the roughness coefficient, the infiltration rate and the rainfall intensity, then this global platform is a super fast way to get a first impression on what the possibly endangered areas in case of flooding are.

58 http://54.164.155.9/webgis/ (access on 30-03-2023)

Furthermore, the **platform is not limited to the calculation of pluvial flood hazards**; also the **fluvial** and a possible **coastal** flood hazard can be calculated in some minutes or seconds, and they can be **calculated each of them for its own or combined together**, with a multiple hazard map as result.

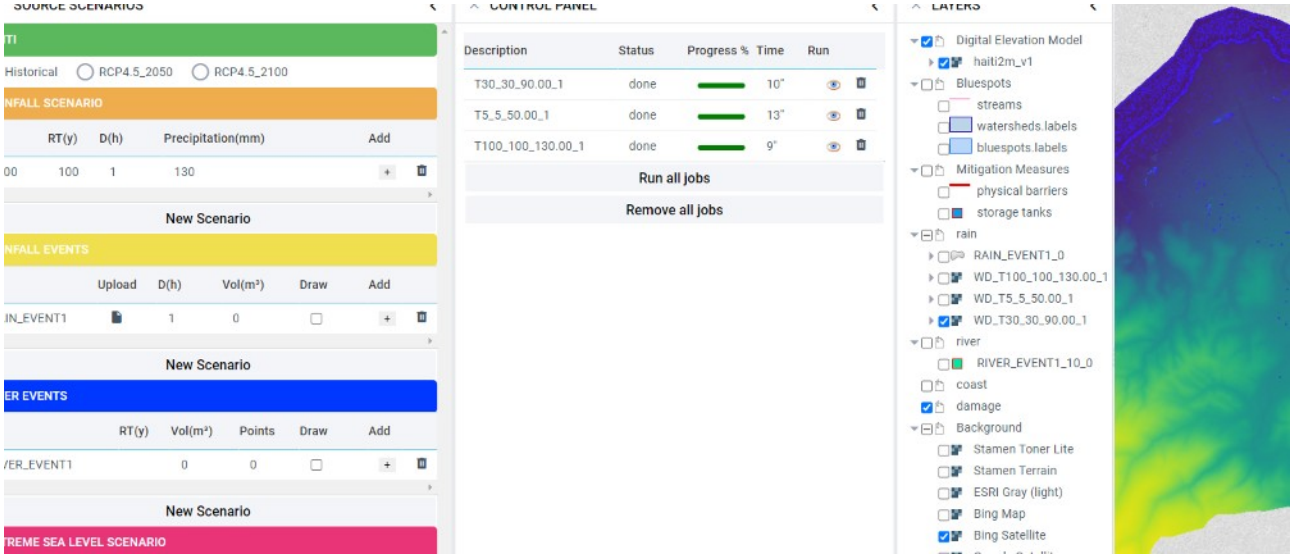


Figure 76: Screenshot of the "SaferPLACES" pilot area "Haiti" (=AOI Carrefour)

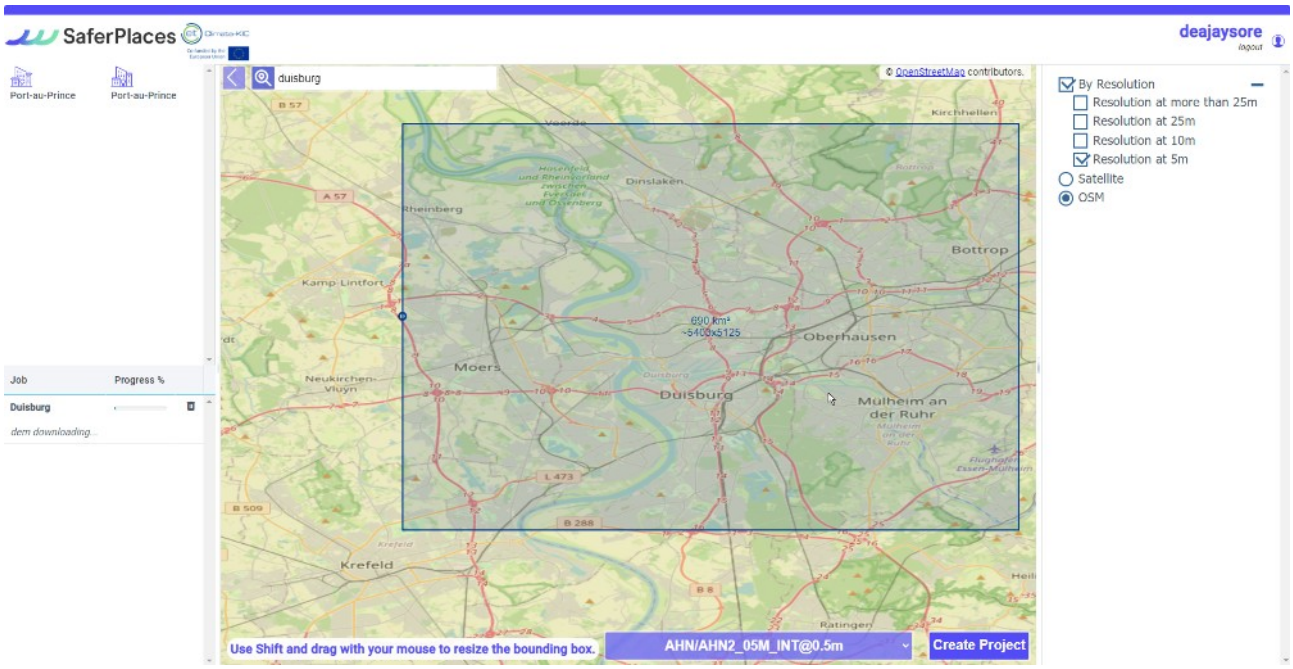


Figure 77: Screenshot of the "SaferPLACES" global platform, where an AOI of own choice can be defined and calculated throughout the world

But, on the other hand, the **precise function of the HFSA** (hierarchical filling-and-spilling approach) is **very important for the interpretation of the results**. As the algorithm works (like the name suggests) with the filling of sinks and the spilling over their edge, as soon as it is filled with water, the non-existence of sinks in a terrain with a continuous gradient seems to cause serious problems for the algorithm, as there basis from where the filling begins, is missing. **When no water accumulates in local depressions, but rather all the water in the model is just being routed downwards by the D-8 flow algorithm, the flood extent and water depth in the result is accordingly low or simply not present.**

This is the case in the AOI Carrefour – we can remember from Chapter 2.1 that the urban centre of Carrefour is situated on an alluvial fan, where the slope is continuously falling from south to north. So apart from some minor local depressions, this are conceivably unfavourable conditions for the HFSA to achieve good and reliable results. I was not aware of this fact when the AOI had been suggested by MSF, but I saw it later on the basis of the calculations.

The results from the calculation via the HFSA “Safer_RAIN” can either be viewed in the web platform directly or exported as a GeoTIFF raster file. For the maps of the comparison, I chose the latter, as it enables me to display all results from both tools (Hydro-As-2d and Safer_RAIN) in the same way, with the same colour attributes.

The following maps of the flood results show

- **Frequent rain event** (T2 / T5, 50mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), **whole AOI** (Figure 78)
- **Frequent rain event** (T2 / T5, 50mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), **Carrefour urban area** (Figure 79)
- **Rare rain event** (T30 / T50, 90mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), **whole AOI** (Figure 80)
- **Rare rain event** (T30 / T50, 90mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), **Carrefour urban area** (Figure 81)
- **Extreme rain event** (T100 / T200, 130mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), **whole AOI** (Figure 82)
- **Extreme rain event** (T100 / T200, 130mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), **Carrefour urban area** (Figure 83)

4.2.1 Frequent rain event (T2/T5, 50 mm in 60 minutes)

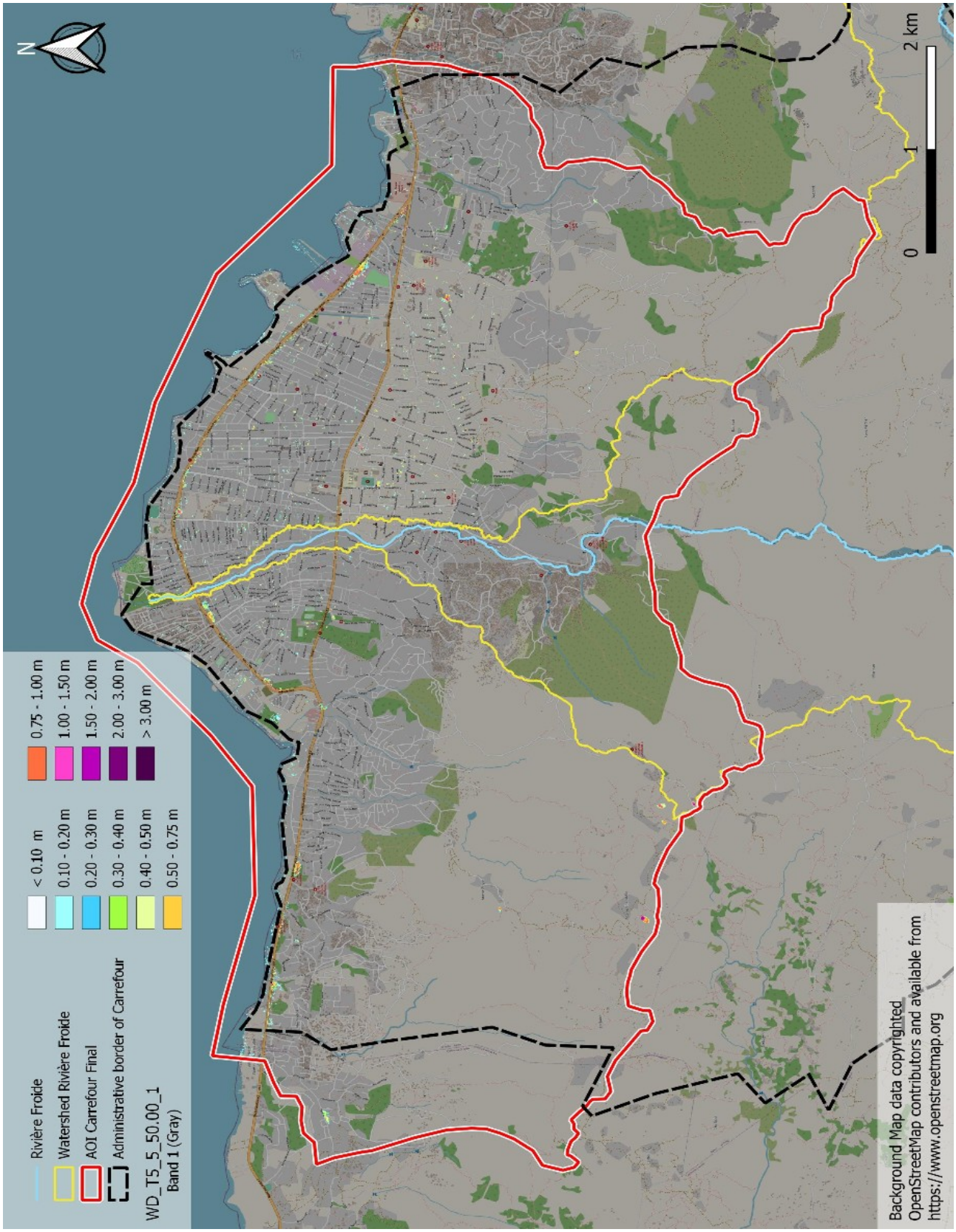


Figure 78: Frequent rain event (T2 / T5, 50mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), whole AOI

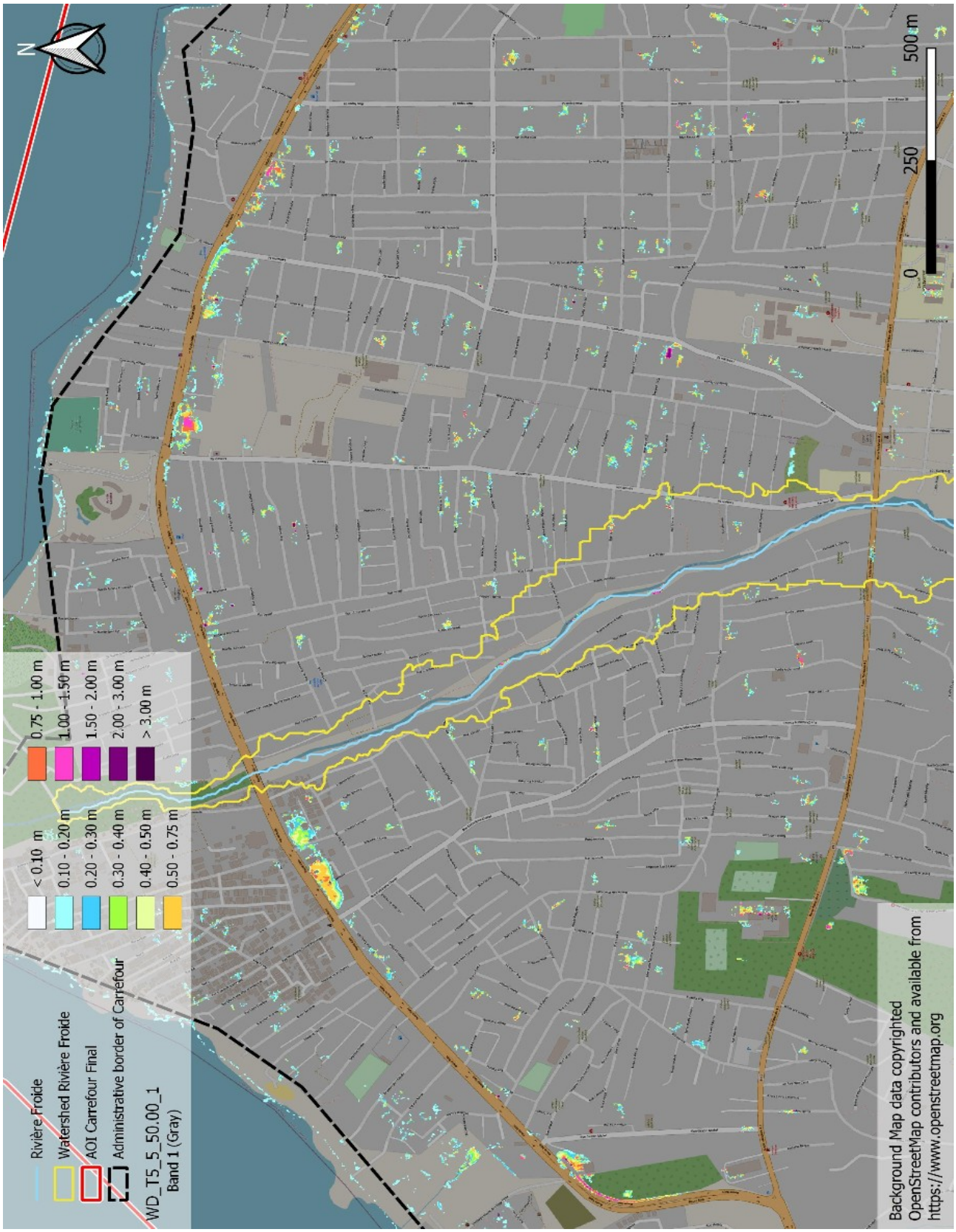


Figure 79: Frequent rain event (T2 / T5, 50mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), Carrefour urban area

4.2.2 Rare rain event (T30/T50, 90 mm in 60 minutes)

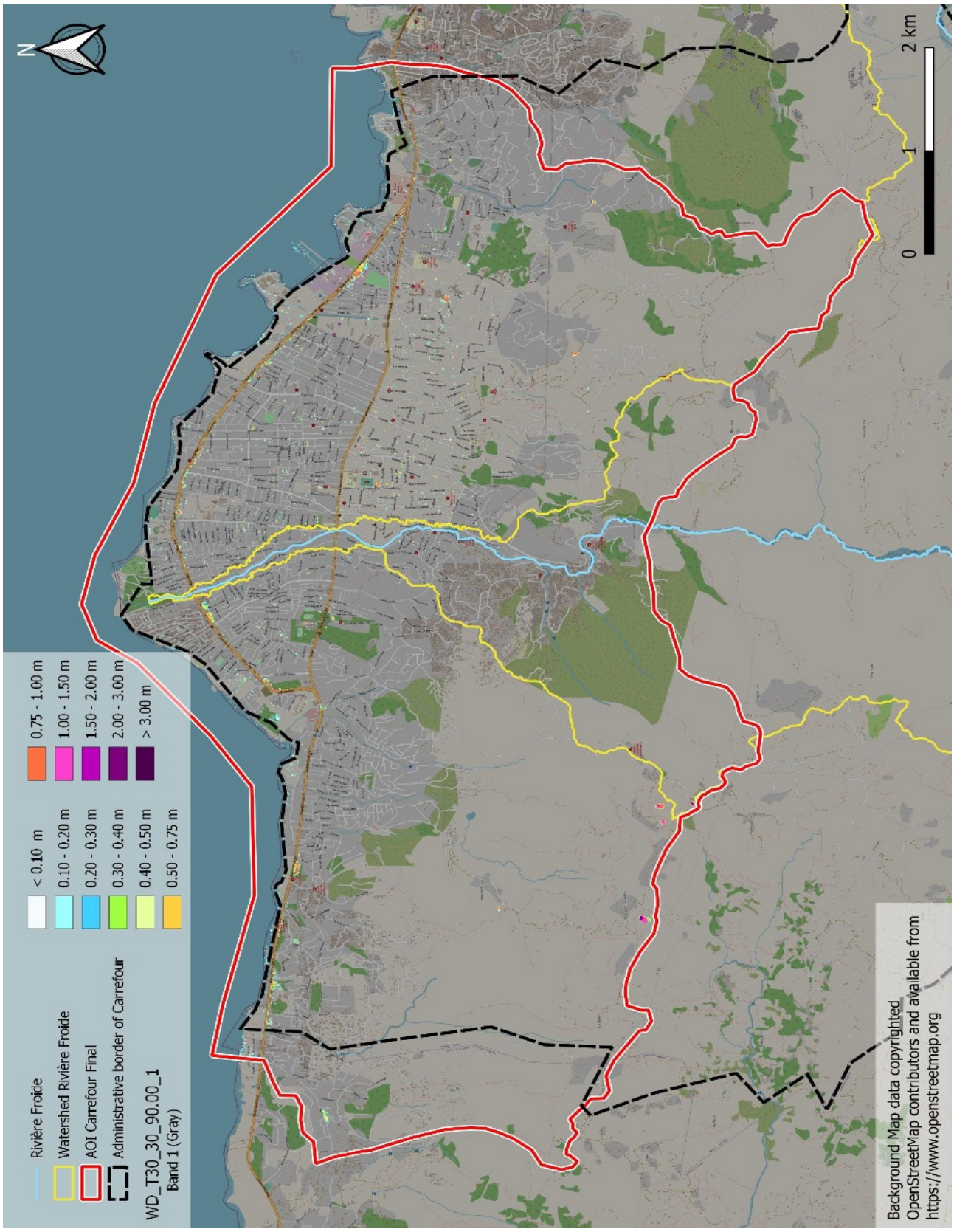


Figure 80: Rare rain event (T30 / T50, 90mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), whole AOI

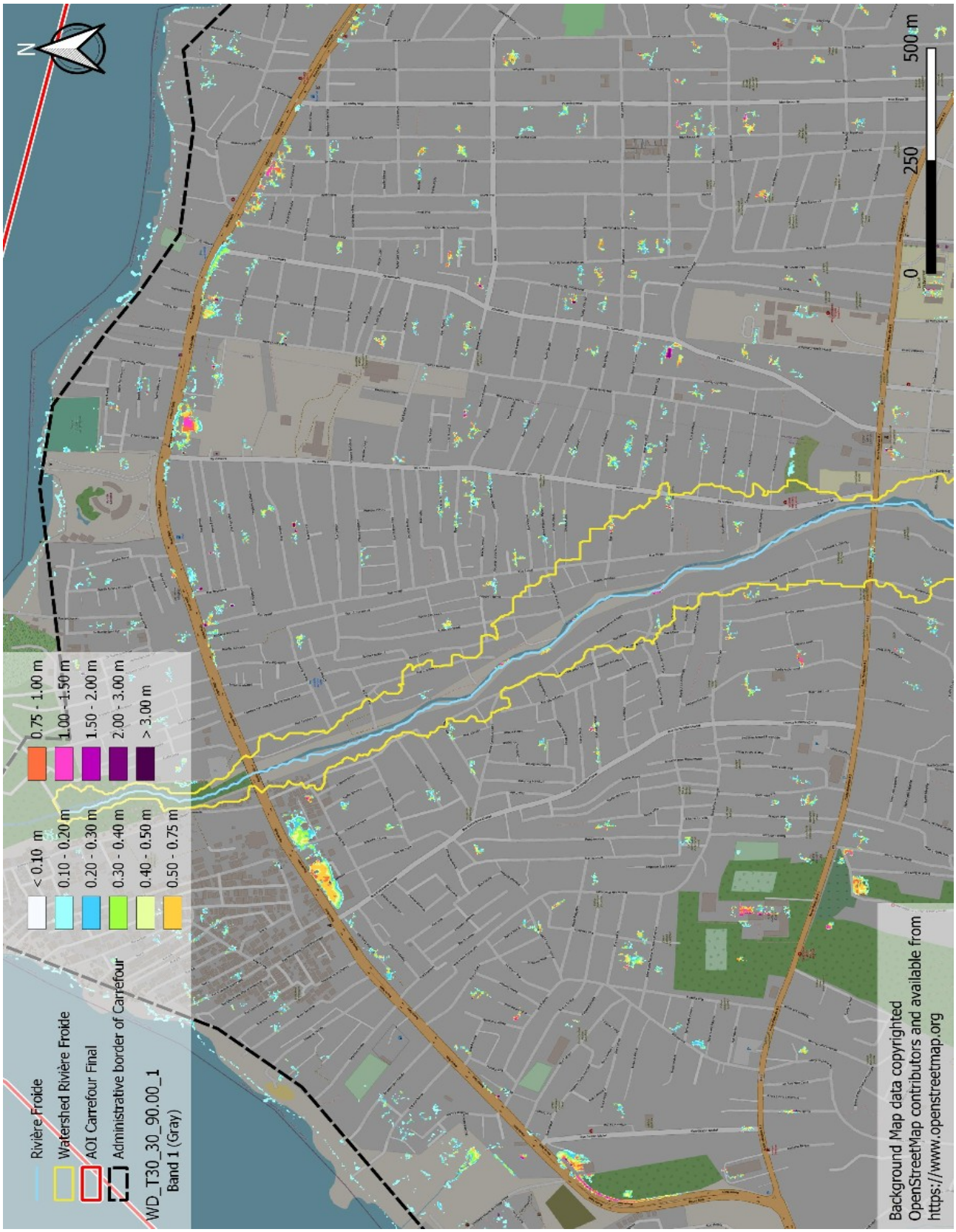


Figure 81: Rare rain event (T30 / T50, 90mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), Carrefour urban area

4.2.3 Extreme rain event (T100/T200, 130 mm in 60 minutes)

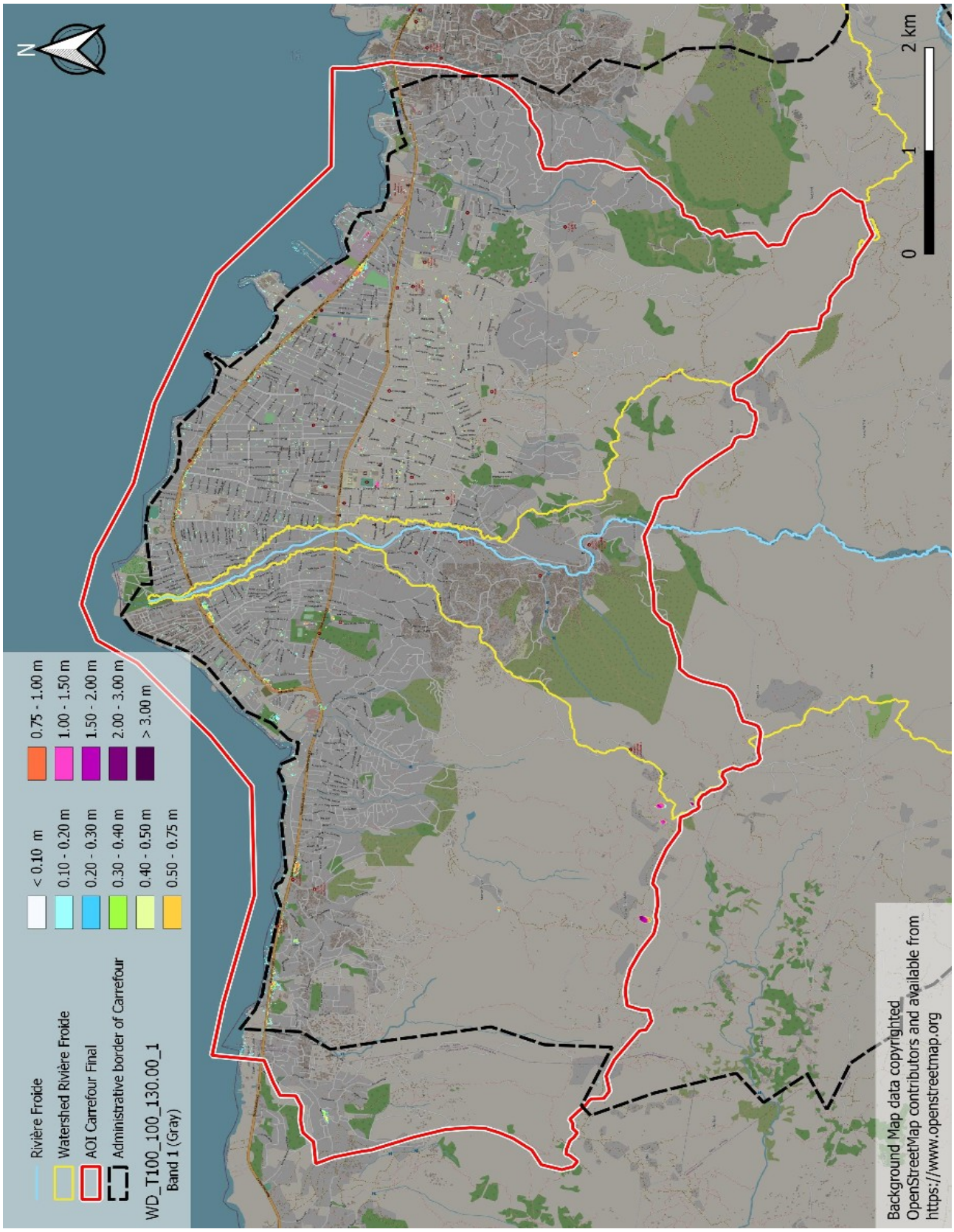


Figure 82: Extreme rain event (T100 / T200, 130mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), whole AOI

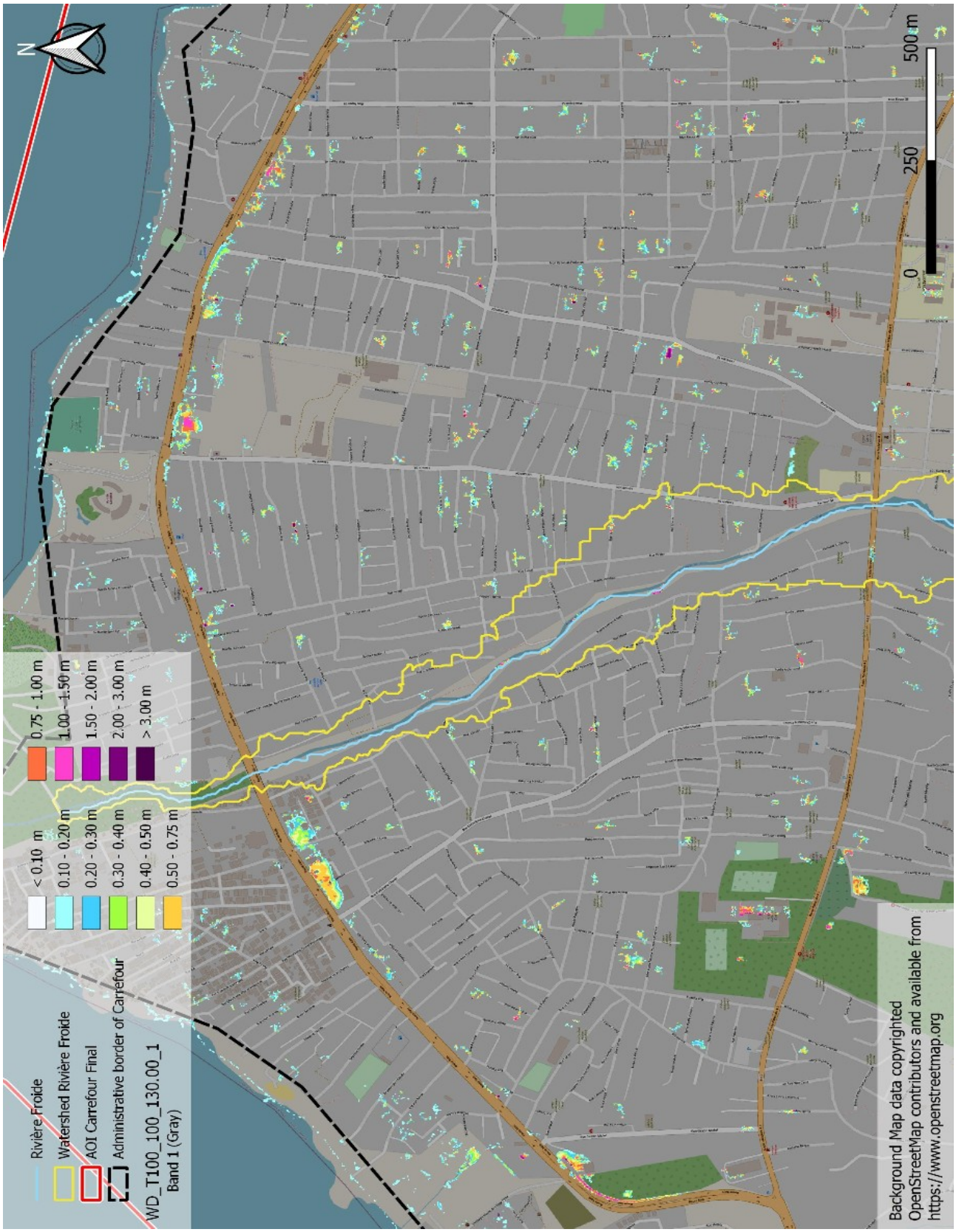


Figure 83: Extreme rain event (T100 / T200, 130mm in 1h): Final water depth (Calculation in "SaferPLACES"web platform), Carrefour urban area

4.2.4 Benchmark calculation for extreme rain event (T100/T200, 130 mm in 60 minutes): Rimini (IT)

As mentioned, the Carrefour AOI has conceivably unfavourable conditions for the HFSA to achieve good and reliable results. So, to remain fair to the tool and the algorithm as such, there shall shortly be described a counterexample with a totally different terrain, the city of Rimini (Italy). This area has been validated against the benchmark calculation of Hydro-As-2d with the "Sturzflut" extension (Figure 84), and I am able to confirm the good results in this case, as I carried out the benchmark calculation to "Safer_RAIN" with Hydro-As-2D in the year 2020 by myself.

The results of "Safer_RAIN" are of course not identical to the 2d hydrodynamic calculation, but they deliver a generally good accordance for the flood extent, as has also been presented in the paper from (Samela et al., 2020):

"We present the first applications of the algorithm to two case studies in Northern Italy. Safer_RAIN output is compared against ground evidence and detailed output from a 2D hydrological and hydraulic numerical model (overall index of agreement between Safer_RAIN and 2D benchmark model: sensitivity and specificity ***up to 71% and 99%****, respectively), highlighting potential and limitations of the proposed algorithm for identifying pluvial flood-hazard hotspots across large urban environments." (Samela et al., 2020)*

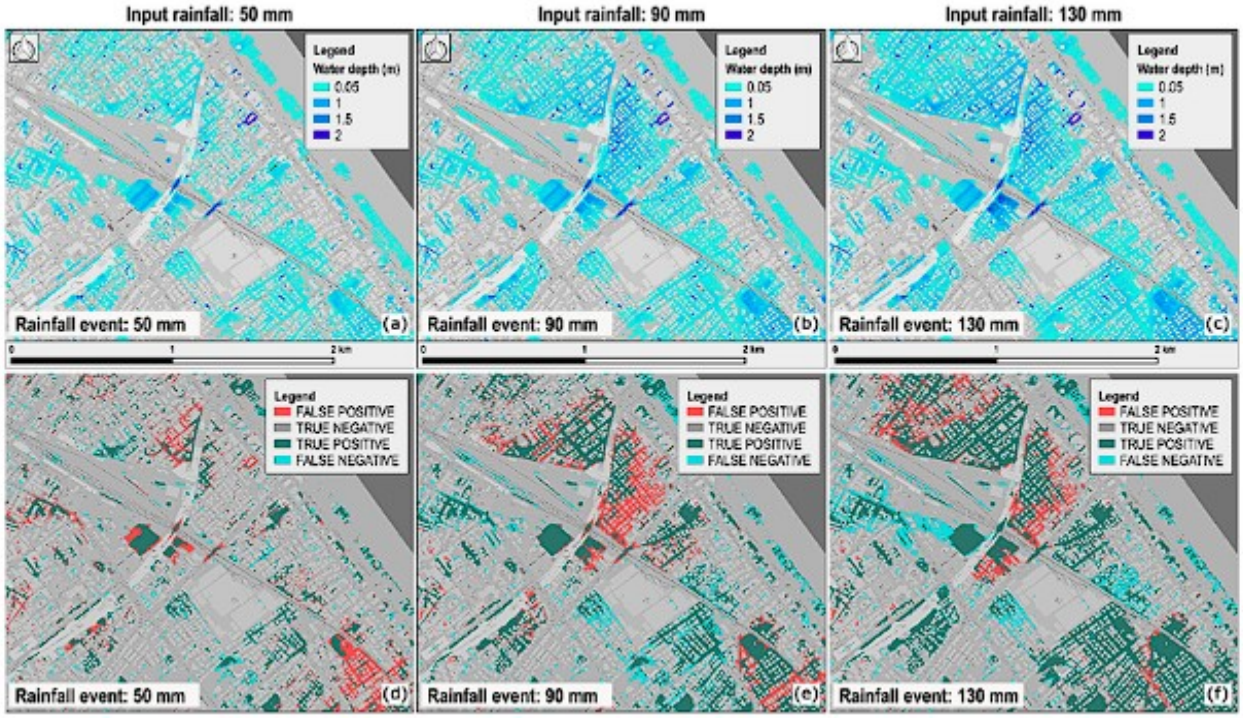


Figure 84: Rimini case study: flooding phase output. (a–c) Water depth maps obtained under the assumptions of spatially uniform rainfall and impervious terrain and relative to three different synthetic rainfall depths (color-scale represents water depth values). (d–f) Comparison between Safer_RAIN and Hydro_AS-2D outputs: images report a classification of pixels into false positive (FP), true positive (TP), false negative (FN), and true negative (TN) (Samela et al., 2020, p. 12)

Indeed, calculating the Rimini area via the “Safer_RAIN” HFSA on the "SaferPLACES" platform, the results are totally different to the results of the Carrefour area (see Figure 85 and Figure 86), which illustrates the huge impact of the DTM on the final result in this case.

As the **filling-and-spilling algorithm largely depends on the form of the terrain, where the calculation is carried out, this input parameter is also far more important than in any hydrodynamic flood calculation**. So for the HFSA, everything rises and falls with the DTM. If the DTM is suitable (rather flat, at least some local depressions, where water can accumulate in the model), the results are generally good, if the DTM is not suitable like in Carrefour, with either a high slope gradient or generally no or only very few local depressions, the result will be poor.

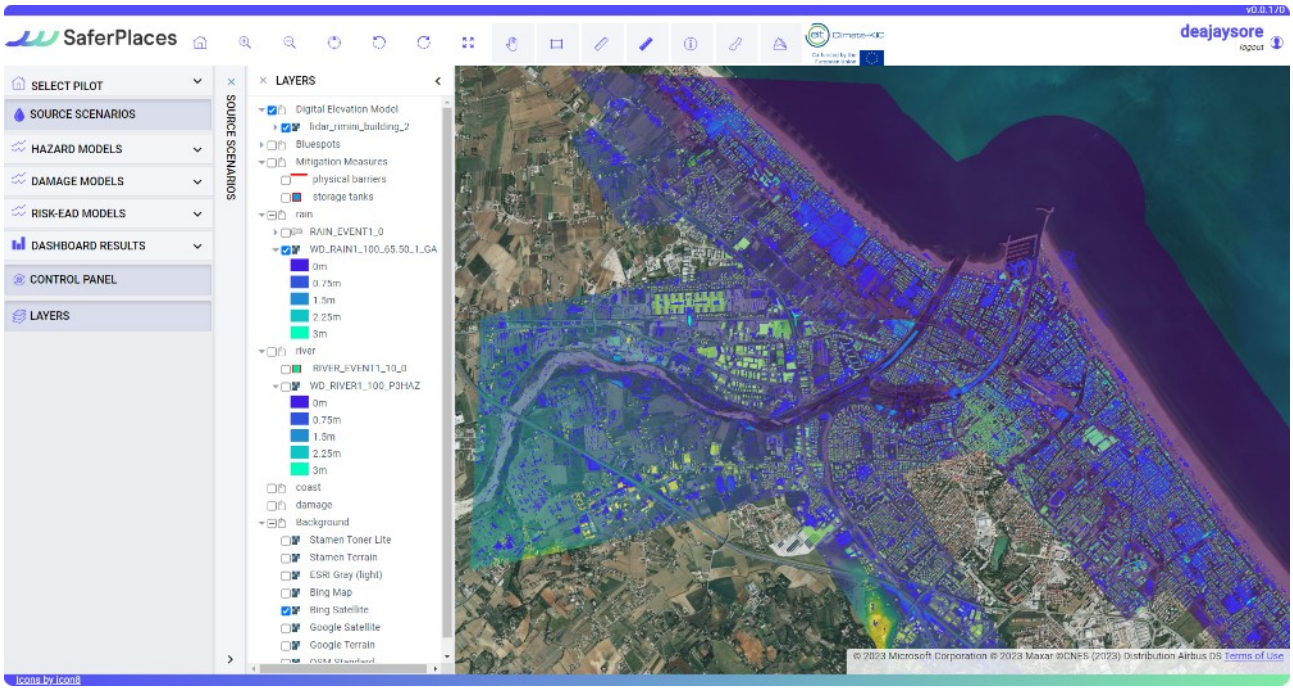


Figure 85: Extreme rain event (130mm in 1h): Final water depth, pluvial flood scenario only (Calculation in "SaferPLACES"web platform), benchmark calculation for city area of Rimini, Italy

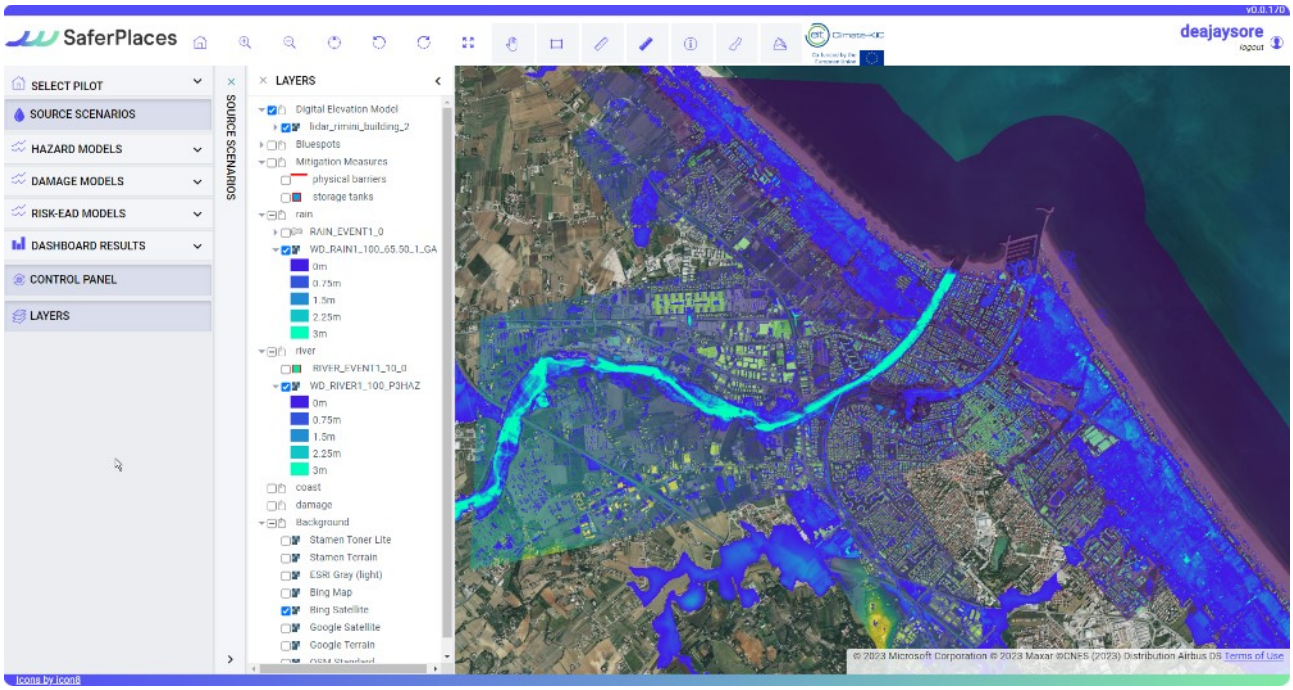


Figure 86: Extreme rain event (130mm in 1h): Final water depth, combined fluvial and pluvial flood scenario (Calculation in "SaferPLACES"web platform), benchmark calculation for city area of Rimini, Italy

4.3. Comparison: GIS-based vs. hydrodynamic

As already mentioned in the introduction, the filling-and-spilling algorithm, on which “Safer_RAIN” is based, has to compete with a methodological issue in our area of interest. This is due to the fact that the whole algorithm is based upon sinks, the filling of sinks and the spilling out of sinks. In many (rather flat) urban areas around the world, this is not a problem, as there are always areas where water can accumulate in the model.

But in the case of Carrefour, the majority of the city is situated on the alluvial fan of the Rivière froide river, and due to the steep terrain in general, we hardly have any really flat areas or bigger local depressions in the area, like there are i.e. in the counterexample of Rimini (almost flat terrain, railroad embankment acting as natural barrier, where the water can perfectly accumulate). This fact of course has a big impact on the results in “Safer_RAIN”.

The major difference between the two tools is in my evaluation the non-consideration of flow velocity and dynamics as such in the “Safer_RAIN” hierarchical filling-and-spilling approach. If overland flow is only routed via D-8 algorithm and the calculation of the inundation areas and water depth as such is done via a “backwards-approach”, where the grid cells are inundated due to water accumulation in local depressions, the algorithm automatically has an issue with steeper terrains including a continuos gradient and no or only few physical barriers like road and railway embankments and small natural hills.

Table 12 shows a comparison of the inundation extent (definition of inundated area: water depth of 1 cm or higher) and the average water depth for all three scenarios:

	Return period	Rainfall amount in mm	Rainfall duration in h	Hydro-As-2d / Sturzflut	Safer_RAIN HFSA
Flood extent in km²	T 2 / T 5	50	1	5.156	0.704
	T 30 / T 50	90	1	8.811	0.718
	T 100 / T 200	130	1	11.115	0.724
Average water depth in m	T 2 / T 5	50	1	0.071	0.055
	T 30 / T 50	90	1	0.111	0.061
	T 100 / T 200	130	1	0.143	0.064
Maximum water depth in m	T 2 / T 5	50	1	6.410	5.969
	T 30 / T 50	90	1	7.501	6.819
	T 100 / T 200	130	1	8.942	6.819

	Return period	Rainfall amount in mm	difference: Safer_RAIN in % of Hydro-As-2d	Hydro-As-2D: Diff 90 / 130 mm to 50mm %	HFSA: Diff 90 / 130mm to 50mm %
Flood extent in km²	T 2 / T 5	50	**-86.34**		
	T 30 / T 50	90	**-91.85**	**70.87**	**1.99**
	T 100 / T 200	130	**-93.49**	**115.55**	**2.71**
Average water depth in m	T 2 / T 5	50	**-22.54**		
	T 30 / T 50	90	**-45.05**	**56.34**	**10.91**
	T 100 / T 200	130	**-55.24**	**101.41**	**16.36**
Maximum water depth in m	T 2 / T 5	50	**-6.88**		
	T 30 / T 50	90	**-9.09**	**17.02**	**14.24**
	T 100 / T 200	130	**-23.74**	**39.50**	**14.24**

Table 12: Comparison on flood extent, mean and maximal water depth (reference Hydro-As-2d: DepthMax = aggregated maximum of all time steps, reference “Safer_RAIN” = depth at last time step, as DepthMax cannot be calculated in the HFSA) between Hydro-As-2d / Sturzflut and Safer_RAIN

One can easily see that the differences are not only obvious in visual form via the maps, but of course also here in the statistics. Therefore, the mesh output of Hydro-As-2d has been converted into raster format with the “Safer-RAIN” output grid as reference grid, so that the pixel centre is exactly the same.

The most obvious facts for the AOI Carrefour are:

- The **flood extent** is between 86% and 93% (130mm rainfall event: 11.11 km² in Hydro-As-2D but only 0.72 km² in “Safer_RAIN”) smaller via the “Safer_RAIN” calculation than in the Hydro-As-2D calculation
  - **The difference is becoming bigger with increasing rainfall intensity**
  - In Hydro-As-2D, flood extent increases by 71% between the 50mm frequent rain event and the 90mm rare rain event, and by 116% between the 50mm frequent rain event and the 130mm extreme rain event
  - In “Safer-RAIN”, flood extent only increases by 2.0% between the 50mm frequent rain event and the 90mm rare rain event, and by 2.7% between the 50mm frequent rain event and the 130mm extreme rain event
- The **average** (arithmetic mean, water depth 0.01m and higher) **water depth** is between 22% and 55% (130mm rainfall event: 0.143m in Hydro-As-2D but only 0.064m in Safer_RAIN) smaller via the “Safer_RAIN” calculation than in the Hydro-As-2D calculation
  - **The difference is significantly becoming bigger with increasing rainfall intensity**
  - In Hydro-As-2D, the average water depth increases by 56% between the 50mm frequent rain event and the 90mm rare rain event, and by 101% between the 50mm frequent rain event and the 130mm extreme rain event

- In “Safer-RAIN”, the average water depth only increases by 11% between the 50mm frequent rain event and the 90mm rare rain event, and by 16% between the 50mm frequent rain event and the 130mm extreme rain event, which is a very low increase in comparison to Hydro-As-2D

- The **maximum water depth** is between 17% and 40% (130mm rainfall event: 8.94m in Hydro-As-2D but only 6.82m in Safer_RAIN) smaller via the “Safer_RAIN” calculation than in the Hydro-As-2D calculation
  - **The difference is becoming bigger with increasing rainfall intensity**
  - In Hydro-As-2D, the maximum water depth increases by 17% between the 50mm frequent rain event and the 90mm rare rain event, and by 40% between the 50mm frequent rain event and the 130mm extreme rain event
  - In “Safer-RAIN”, the average water depth increases by 14% between the 50mm frequent rain event and the 90mm rare rain event, and also by 14% between the 50mm frequent rain event and the 130mm extreme rain event

These differences in raw numbers of course can also be seen visually at the first glance by looking at the result maps. But, at this point, it is very important to draw once again the attention to a fundamental difference between the two flood calculation approaches: The concept of **time**.

As described in chapter 3.4.5, Hydro-As-2D is a hydrodynamic model solving the shallow-water-equations via a finite-volumes method and thus working with flood dynamics. So we basically know how much water is passing through a cross section during which time span – a discharge curve for every virtual cross section X can be calculated out of the 2d model. As we know the water depth for every output interval (in the case of this work: every 5 minutes / 300 seconds), we can easily aggregate the water depths of all output intervals into one single output, where only the maximum of all output intervals is being displayed (the so-called “DepthMax”).

The methodology of the hierarchical filling-and-spilling algorithm used in “Safer-RAIN” does not allow this, as the model is GIS-based and not hydrodynamic. This fundamental difference lies in the very nature of the approach, and in rather flat areas with low flow velocities and water accumulating in local depressions, the approach is a competitive alternative to a conventional hydrodynamic flood calculation, as the benchmark calculations for the Italian city of Rimini showed (Samela et al., 2020).

But in the case of a terrain like in Carrefour, with very steep hills rising directly after the coastline and a continuously rising terrain between the coast and the foot of the hill slopes, the algorithm clearly has a methodological problem. The surface flow is very fast, and just as fast as the water has come, it is gone, like a quickly passing flood wave. I would like to illustrate this in the next maps, which show the water depth of the extreme rainfall scenario (T100 / T200, 130 mm in 1h) in the Hydro-As-2d / Sturzflut model at different time steps (output intervals), namely at 15 min (Figure 87), 30 min (Figure 88), 60 min (Figure 89) and 120 min (Figure 90).

It illustrates how fast the pluvial flood wave is moving through the city of Carrefour – having the peak inundation in most parts of the city at the 60 minutes time step, directly after the block rainfall event stopping after 60 minutes, and then draining quickly into the sea.

So, if we took the water depth after 120 minutes as basis for the comparison of the two approaches, things would look different: the flood extent as well as the water depth is already much lower then than at the peak inundation after 60 minutes. And this basis for the comparison has also been set for the Rimini benchmark (Samela et al., 2020): Here, the water depth of the **last time step** of the "Sturzflut" / Hydro-As-2D model has been compared with the water depth in "Safer-RAIN", not the aggregated maximum water depth of all time steps.

In the case of Rimini, the difference between the maximum water depth and the water depth of the last time step is much lower, as the calculation ended directly after the block rainfall event in the model and the terrain is totally different (almost flat, with local depressions due to road at railway embankments). But here, in Carrefour, the time component makes a huge difference.

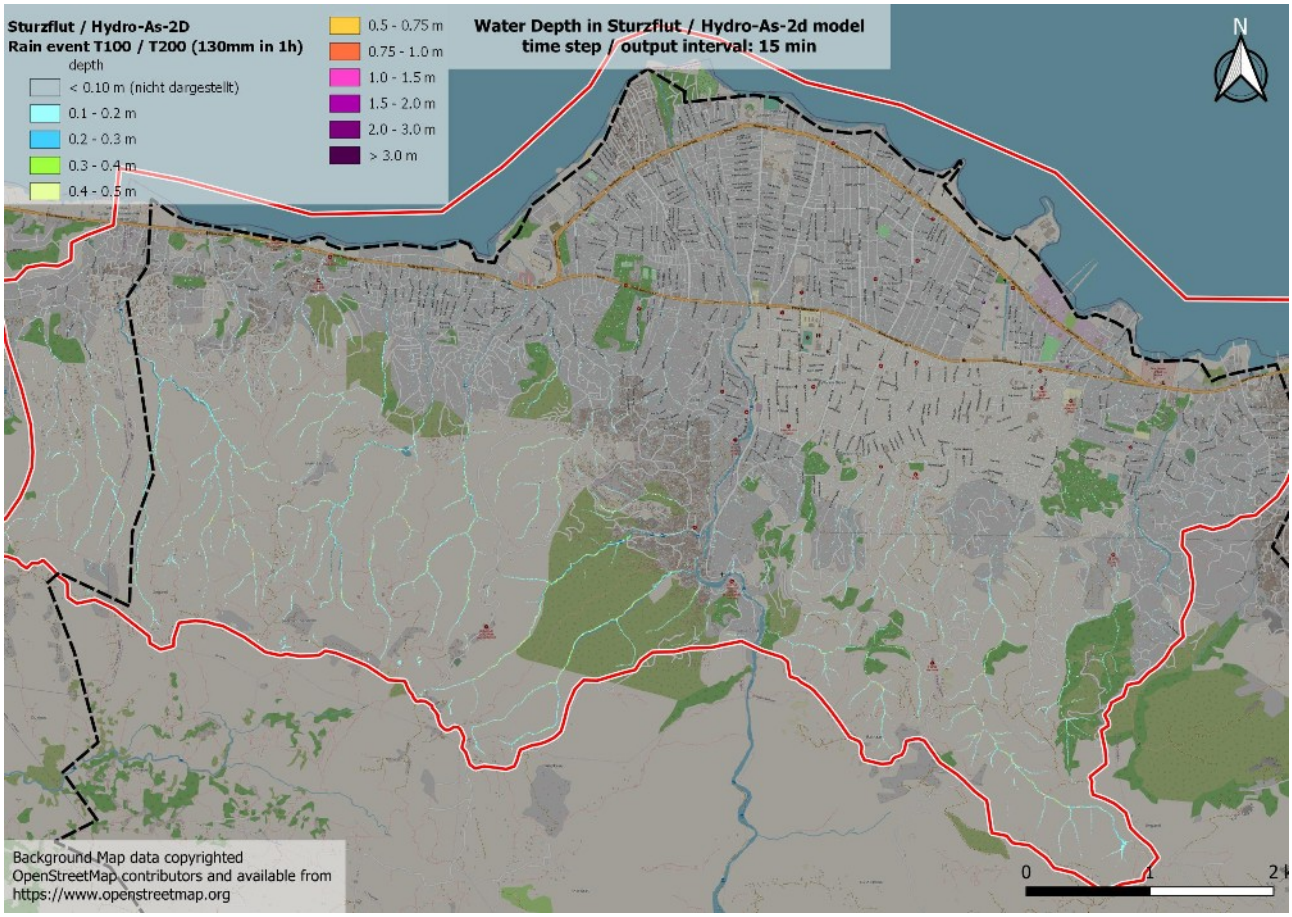


Figure 87: Water Depth in Sturzflut / Hydro-As-2d model (extreme rainfall scenario, 130mm in 1h, at TS 15 min)

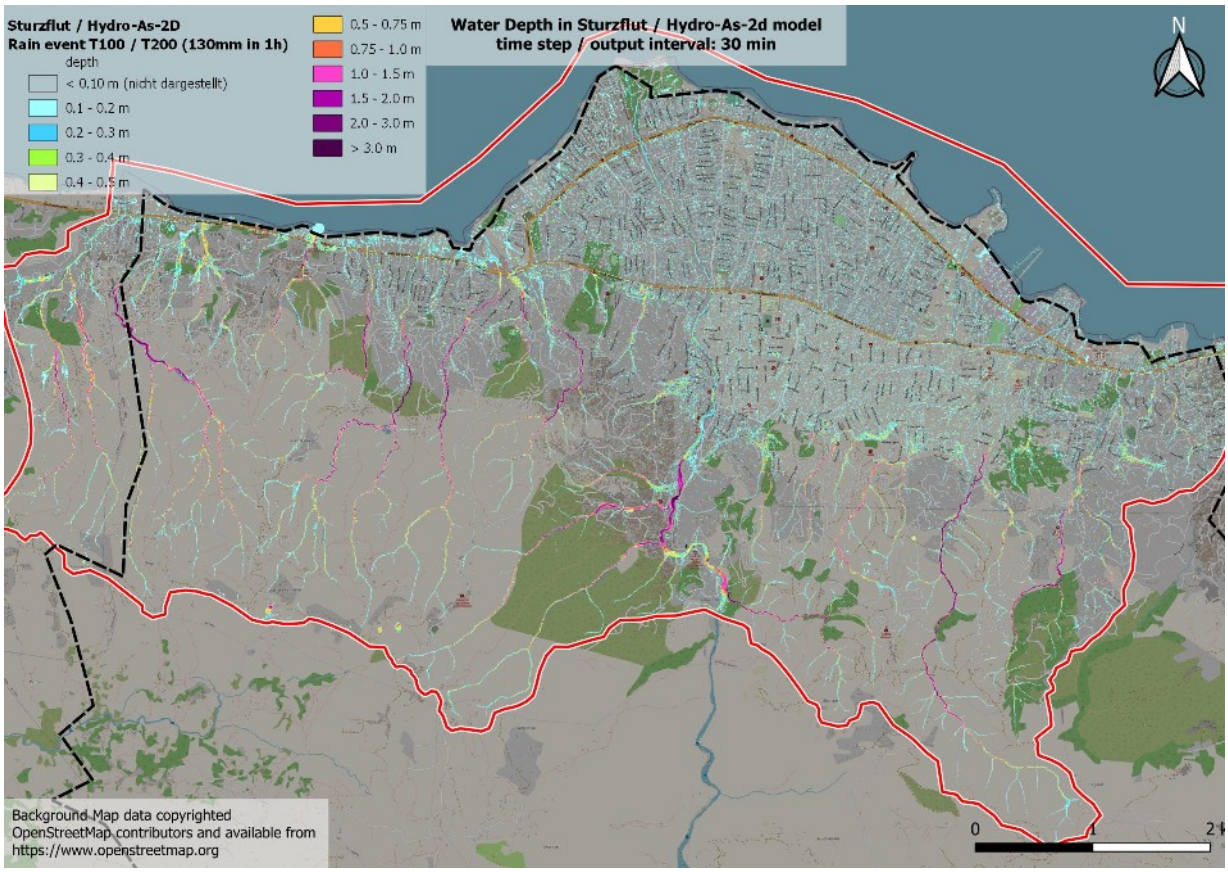


Figure 88: Water Depth in Sturzflut / Hydro-As-2d model (extreme rainfall scenario, 130mm in 1h, at TS 30 min)

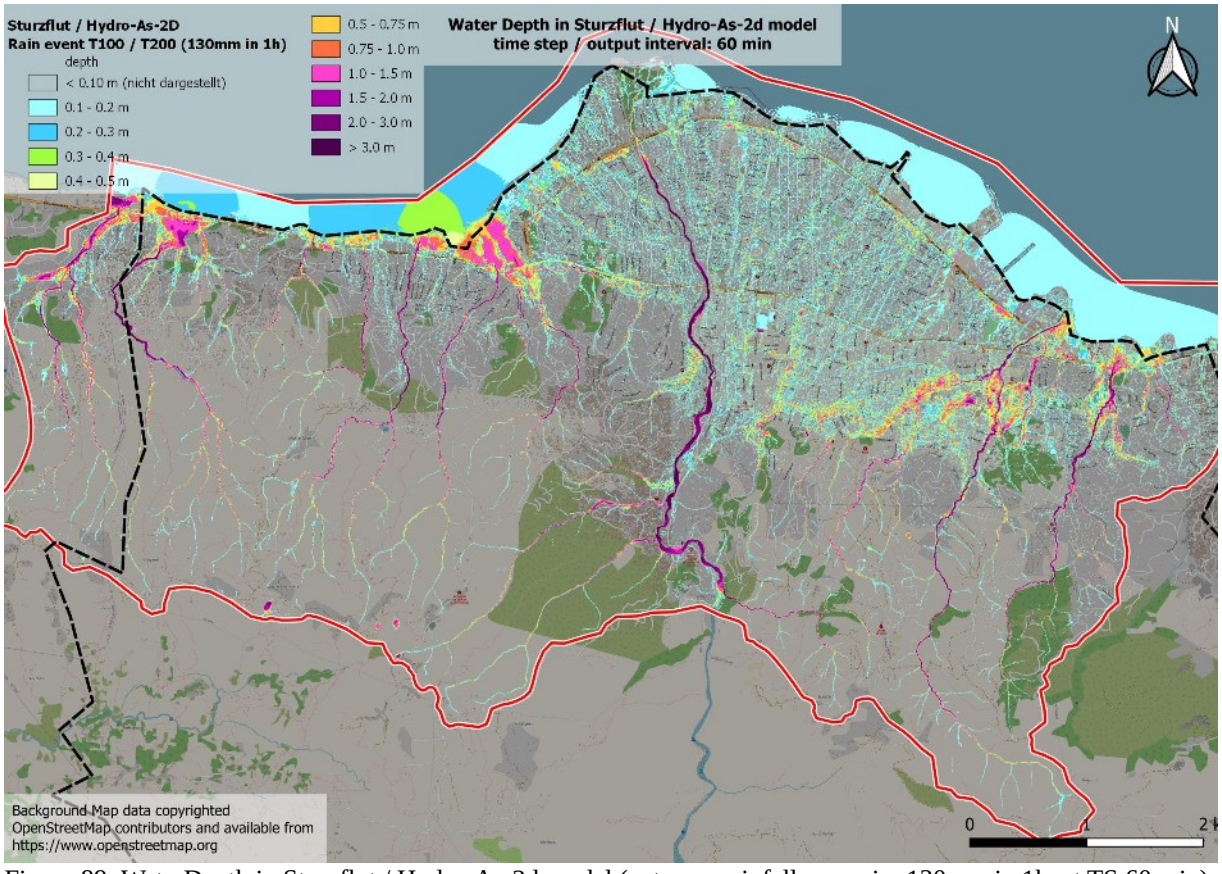


Figure 89: WaterDepth in Sturzflut / Hydro-As-2d model (extreme rainfall scenario, 130mm in 1h, at TS 60 min)

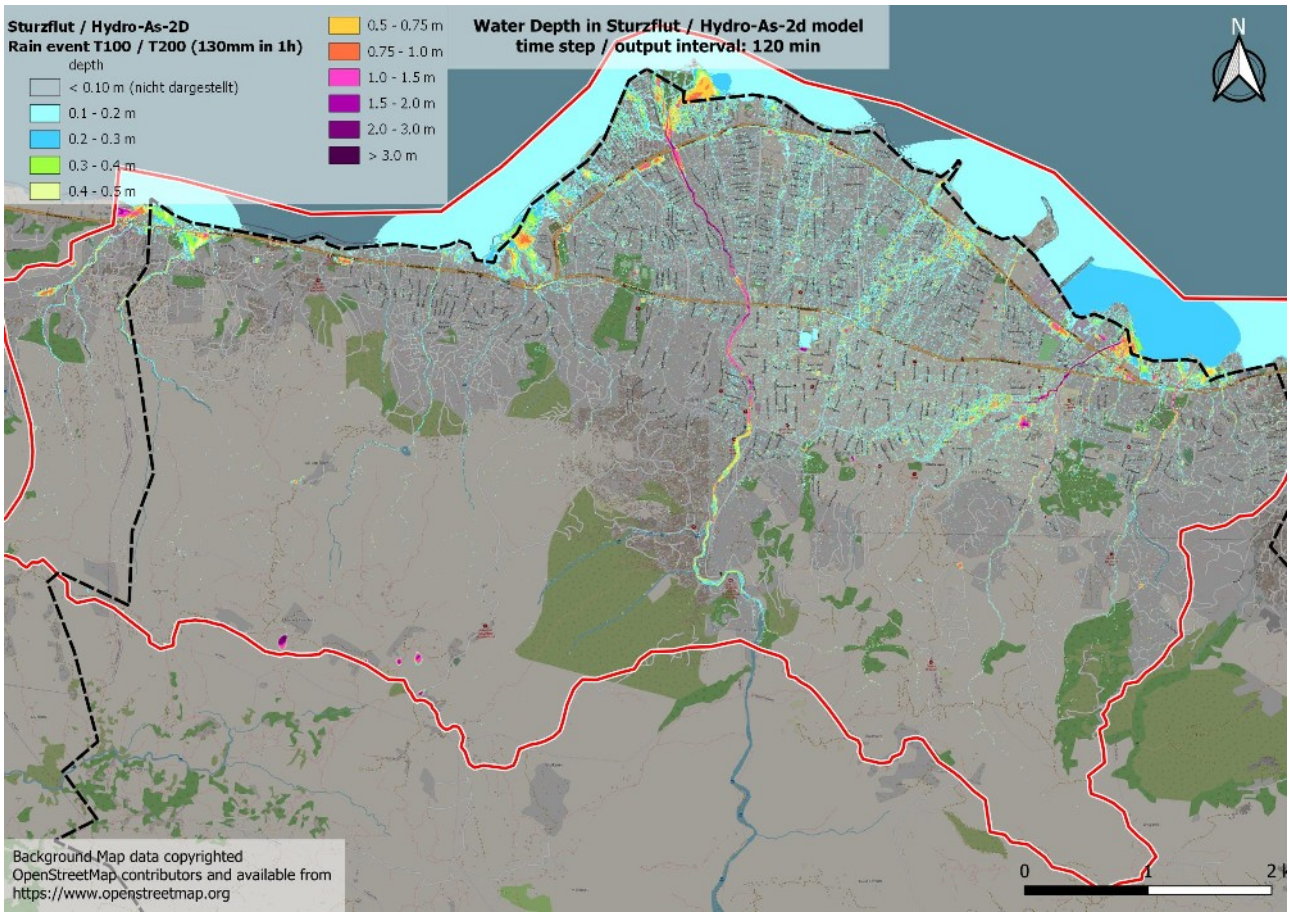


Figure 90: WaterDepth in Sturzflut / Hydro-As-2d model (extreme rainfall scenario, 130mm in 1h, at ts 120 min)

4.4. Flood risk assessment for buildings and critical infrastructure

The second objective of this work was a flood risk assessment for buildings and critical infrastructure in Carrefour at different flood scenarios. Due to the methodological problems of the HFSA in the area of interest, I only took the results of the hydrodynamic calculation as basis for the evaluation.

First step was to define which amenities should fall under this category and then, in a second step, extract its position. My chosen approach was to extract this information from OpenStreetMap via the QGIS plug-in “QuickOSM”, which enables the user to export exactly the desired data from OSM as a separate layer into QGIS in no time – as long as the information is available in OSM, we are depending on user generated content in the case of OSM.

This approach worked amazingly well for the extraction of critical infrastructure (defined as schools, kindergartens / nursery schools and hospitals / clinics, but rather bad for the building substance as a whole (see chapter 2.2.1) – for the whole city of Carrefour, there are only 15351 buildings documented in OSM, which might be about 10%, but not more. A possible solution could

be an object-based image analysis (OBIA) of high-resolution satellite images, but that would have gone beyond the scope of this work, which focuses on the data fundamentals and the model comparison. For this reason, I deliberately limited the evaluation to the available OSM data.

For the AOI is a whole, out of OSM could be extracted the following data as basis for the floor risk assessment:

- 15,351 buildings
- 34 clinics / hospitals
- 258 educational facilities, ranging from kindergarten over primary and secondary school until college / university.

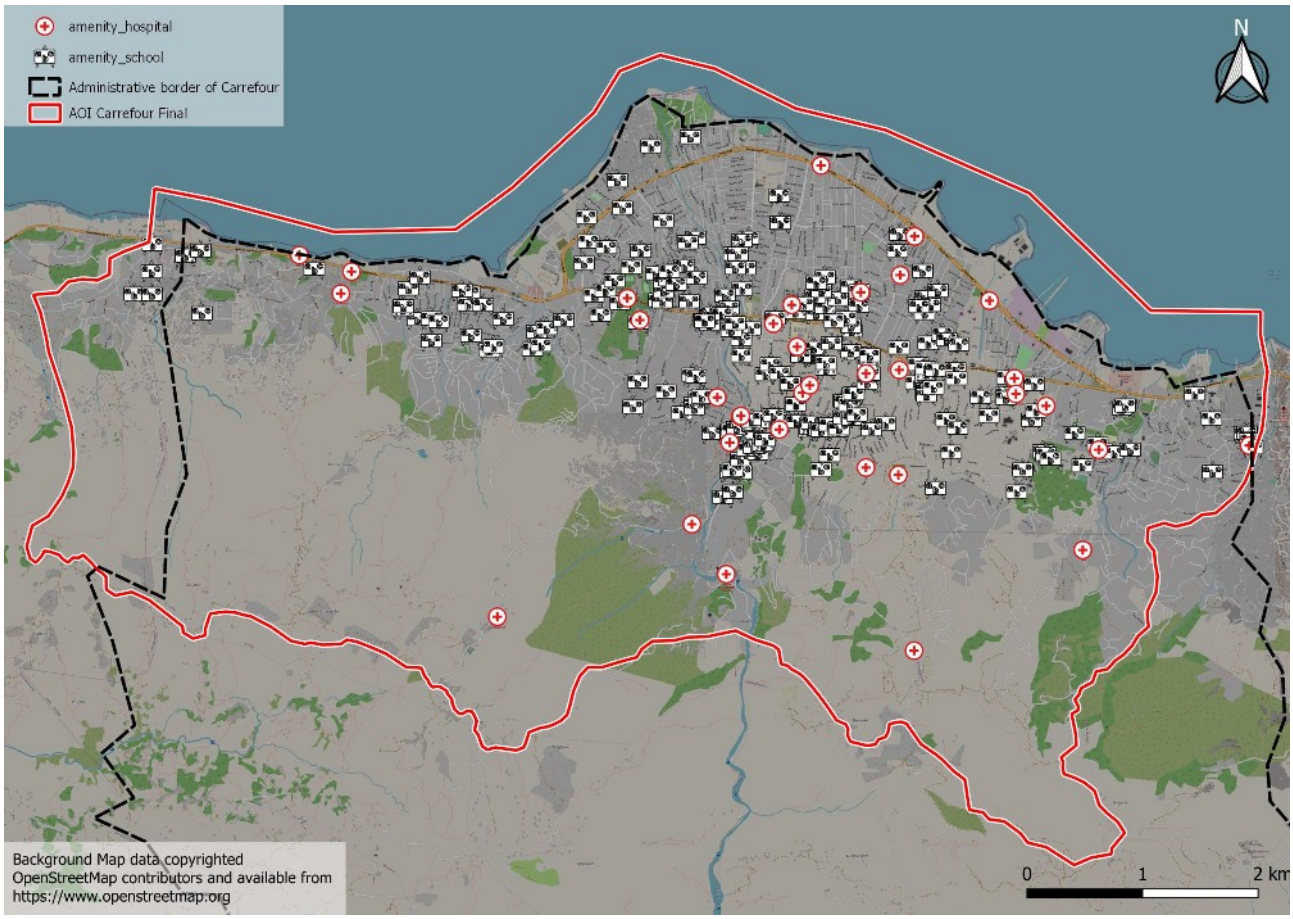


Figure 91: Critical infrastructure (schools, hospitals) extracted out of OpenStreetMap

The extracted data is shown in Figure 91. The next step was to convert the aggregated maximum water depth (Hydro-As-2D output) from mesh data format into raster format and to reclassify it into 3 different depth classes:

- 0.1 – 0.5 meter
- 0.5 – 1.0 meter
- > 1.0 meter

The intersection between the different depth classes (derived from the hydrodynamic calculation with Hydo-As-2D / Sturzflut add-on, reclassified aggregated maximum depth of all time steps), the 3 rainfall scenarios (50mm, 90mm and 130 mm) and the 2 layers with the amenities of the critical infrastructure ("Schools": schools, kindergartens / nursery schools, colleges, universities; "Hospitals": hospitals and clinics) brought the following results (see Table 13):

Rain event	Amenities	WD 0.1m-0.5m	WD 0.5m-1.0m	WD >= 1.0m
T2/5 50mm 1h	Schools	31	2	1
T2/5 50mm 1h	Hospitals	5	0	1
T2/5 50mm 1h	Buildings	1687	678	355
T30/50 90mm 1h	Schools	66	6	1
T30/50 90mm 1h	Hospitals	8	1	1
T30/50 90mm 1h	Buildings	2460	1003	632
T100/200 130mm 1h	Schools	80	16	2
T100/200 130mm 1h	Hospitals	9	2	1
T100/200 130mm 1h	Buildings	2506	1197	843

Table 13: Evaluation of flood-endangered buildings, schools and hospitals (source data: OpenStreetMap / QuickOSM)

Table 13 shows a **sharp increase in affected critical infrastructure** especially **between the frequent rainfall scenario** (T2/5, 50mm in 1h) **and the rare rainfall scenario** – the number of affected schools, hospitals and buildings in the depth class 0.1-0.5m roughly doubles, whereas the increase between the latter scenario and the extreme rainfall scenario is lower, especially in the depth class 0.1-0.5m.

The absolute number of affected buildings is not representative, as the buildings contained in OSM are distributed randomly, just as it had been digitized by OSM volunteers (see Figure 92). A better and most notably more consistent distribution of all buildings could be obtained via an object based image analysis. **But still, also the number of affected buildings in this statistic speaks a clear language** – the number of inundated buildings has a sharp increase in all depth classes especially between the frequent rainfall scenario (T2/5, 50 mm in 1h) and the rare rainfall scenario (T30/50, 90 mm in 1h), with a factor between 1.5 and 2.

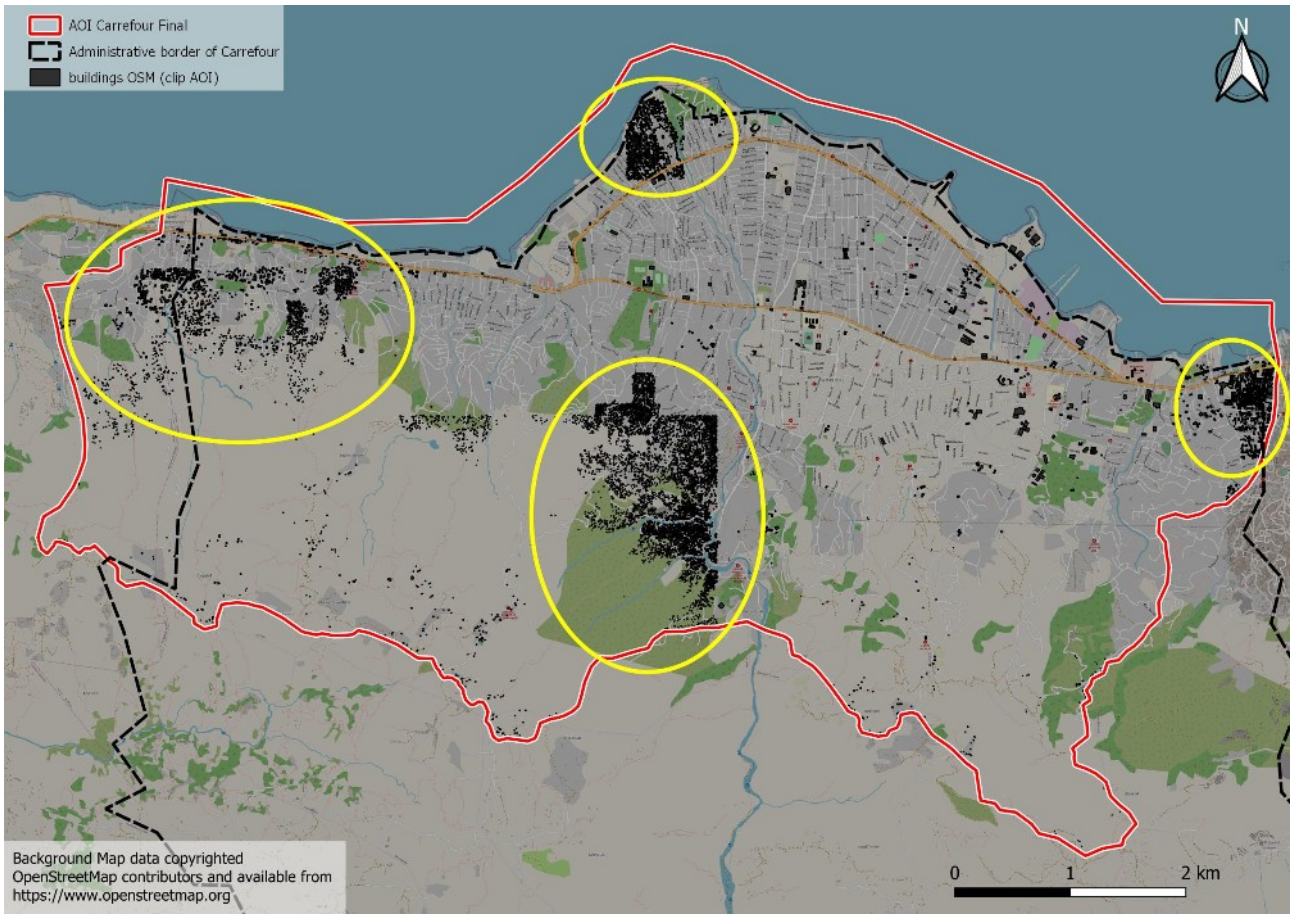


Figure 92: Buildings extracted from OSM - note the irregular (random) distribution

The maps of affected critical infrastructure show the following results:

- maximum water depth in Sturzflut / Hydro-As-2d model: **frequent rainfall scenario** (T2/5, 50 mm in 1h), **intersection with critical infrastructure** (schools, hospitals – see Figure 93)
- maximum water depth in Sturzflut / Hydro-As-2d model: **rare rainfall scenario** (T30/50, 90 mm in 1h), **intersection with critical infrastructure** (schools, hospitals – see Figure 94)
- maximum water depth in Sturzflut / Hydro-As-2d model: **extreme rainfall scenario** (T100/200, 130 mm in 1h), **intersection with critical infrastructure** (schools, hospitals – see Figure 95)

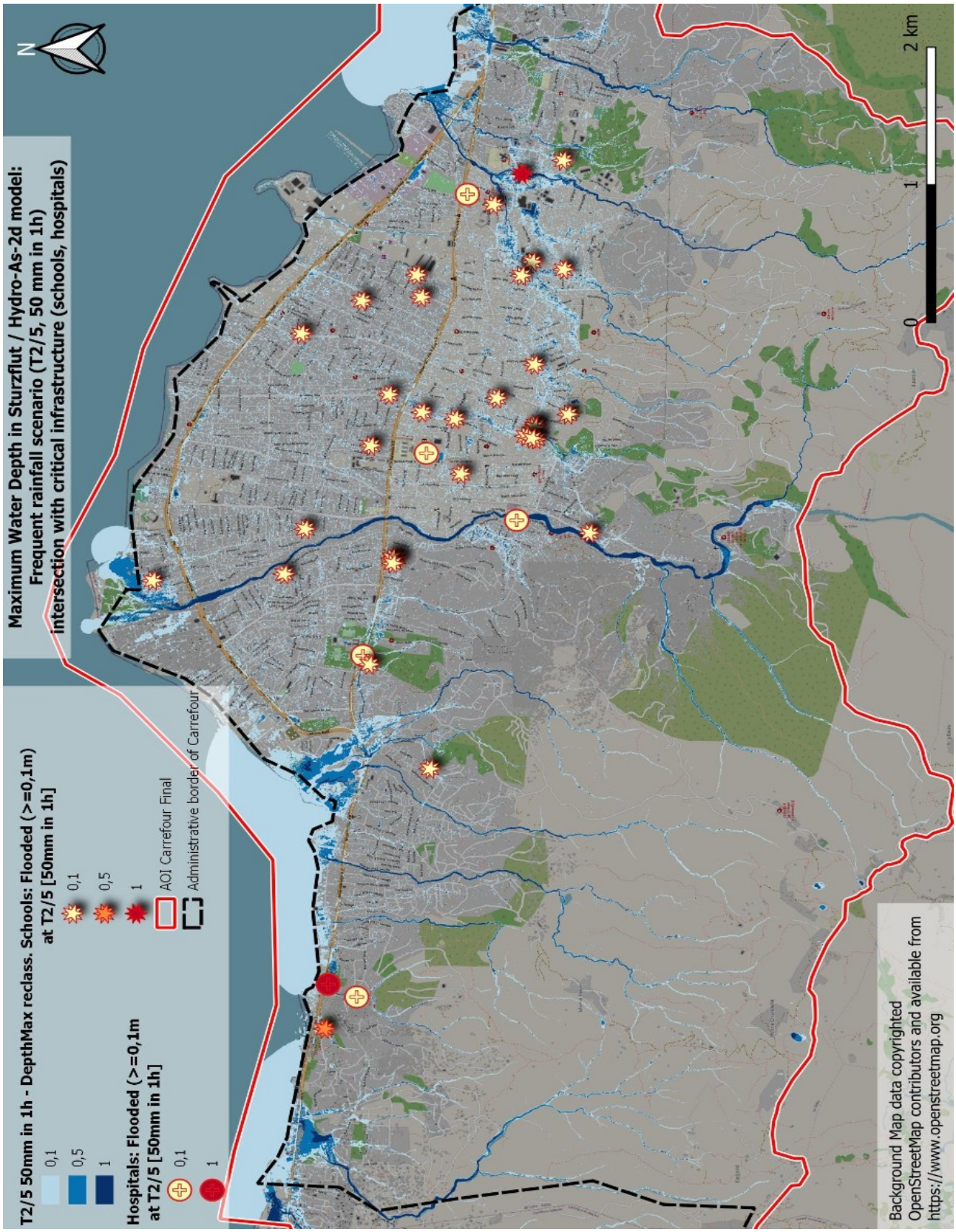


Figure 93: maximum water depth in Sturzflut / Hydro-As-2d model: frequent rainfall scenario (T2/5, 50 mm in 1h), intersection with critical infrastructure (schools, hospitals)

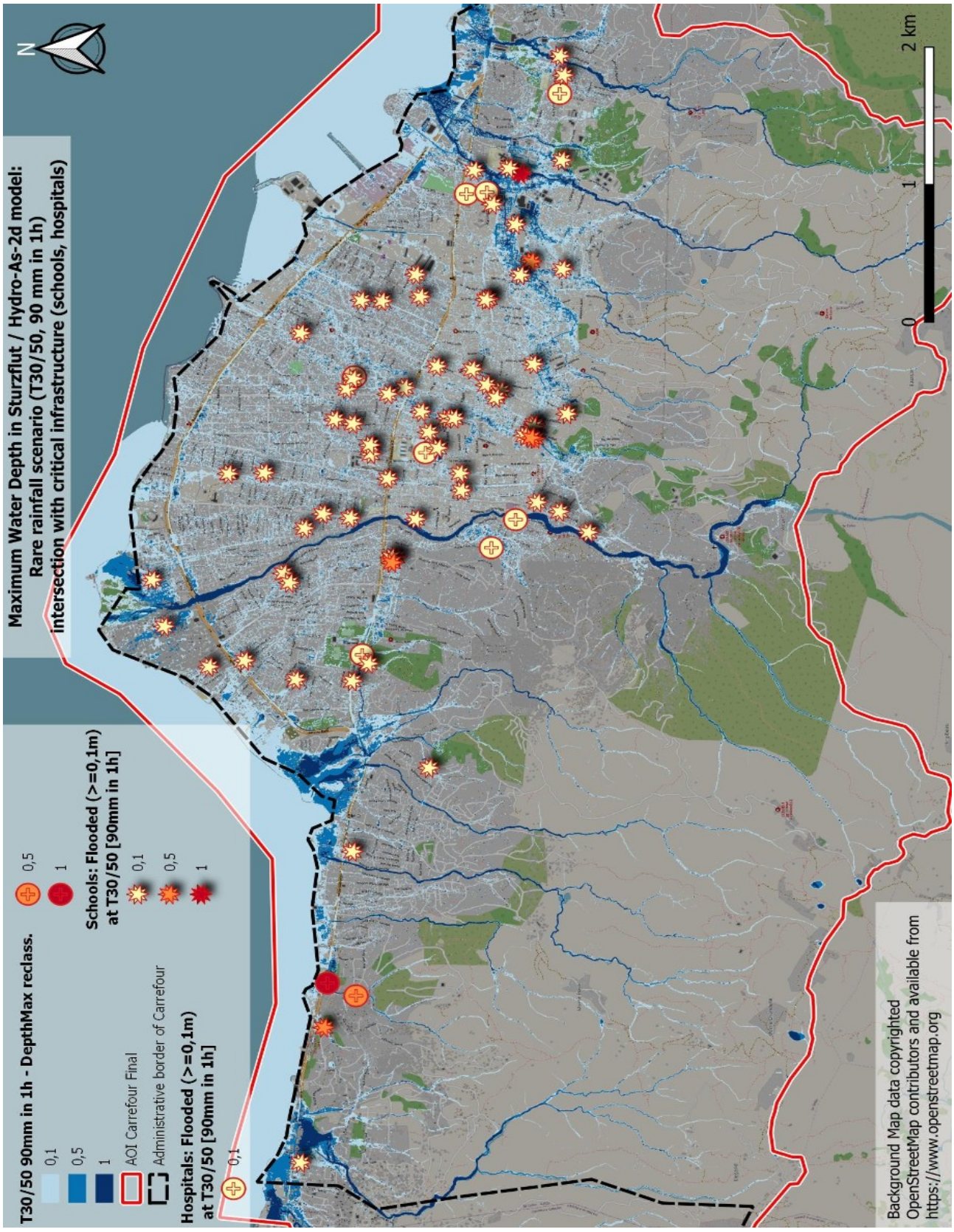


Figure 94: maximum water depth in Sturzflut / Hydro-As-2d model: rare rainfall scenario (T30/50, 90 mm in 1h), intersection with critical infrastructure (schools, hospitals)

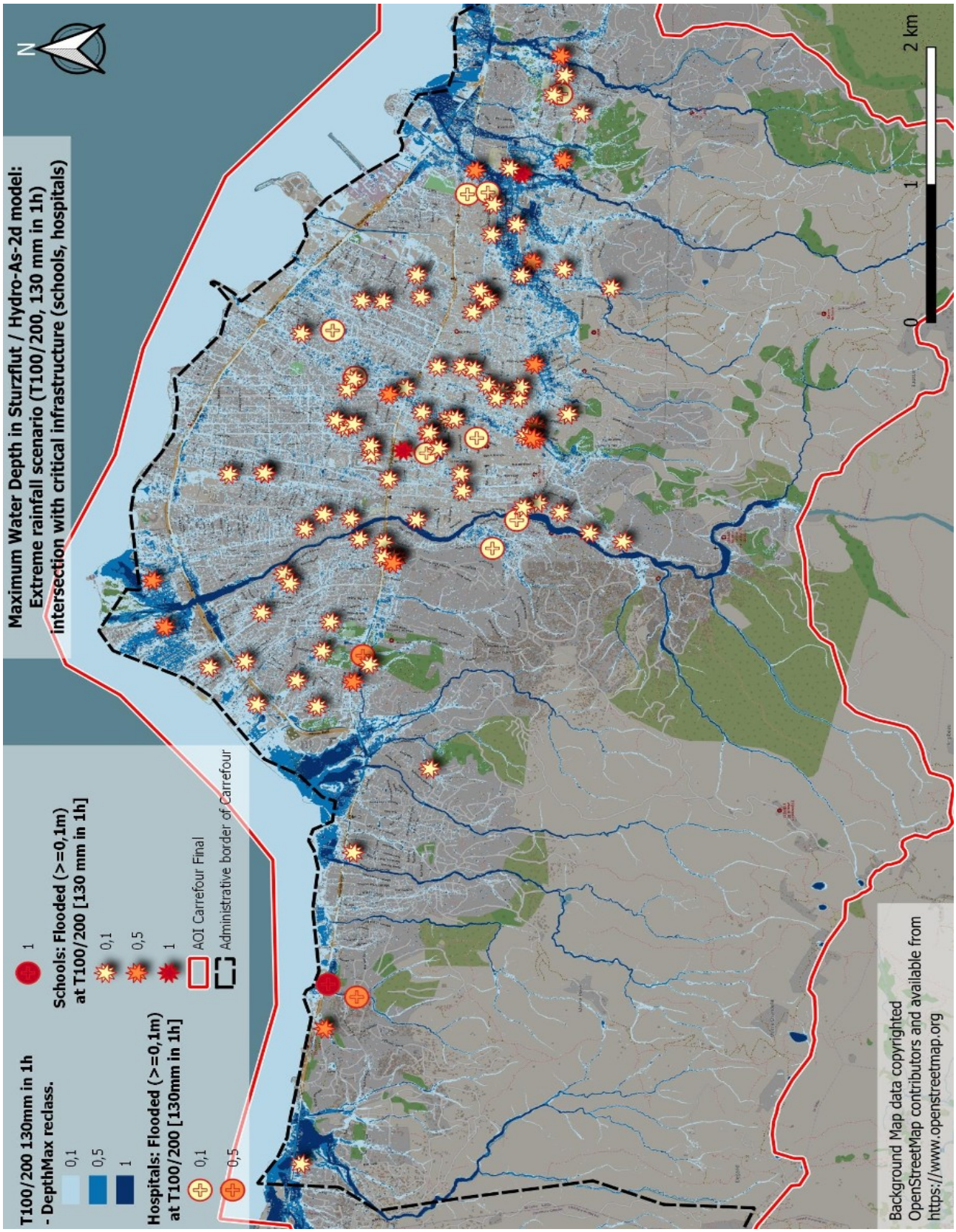


Figure 95: maximum water depth in Sturzflut / Hydro-As-2d model: rare rainfall scenario (T100/200, 130 mm in 1h), intersection with critical infrastructure (schools, hospitals)

5. Discussion of results, conclusion and outlook

5.1. Strengths and Weaknesses of each model

As already pointed out in the previous chapters, the two toolsets for calculating pluvial flood hazard and -risk are basically very different, in methodology a well as in the focus and the abilities. **In my opinion, the hierarchical filling-and-spilling algorithm "Safer_RAIN" has the following strengths (S) and weaknesses (W):**

- S: It is very efficient in calculation speed as well as in the necessary preparation work (global platform is available for the whole world, although with limitations in resolution and accuracy).
- S: "Safer_RAIN" is good in computing the flood hazard in flat or rather flat terrain, and is well suited for flat urban areas.
- S: Not the HFSA as such, but the platform "SaferPLACES" is able to combine different risk scenarios, namely fluvial, pluvial and coastal flood hazard.
- S: It is very cost efficient in comparison to either buying a very expensive software license for commercial hydrodynamic flood calculation software, be this Hydro-As-2d or Jflow.
- S: Also persons, who are no experts in the fields of hydrology, hydraulics and geographic information systems (GIS) are able to carry out the flood hazard calculations. Nevertheless, a certain degree of technical expertise in this fields is necessary to give the calculation model a reasonable input and therefore enable the tool to potentially deliver a reasonable output (otherwise, the "garbage-in-garbage-out" principle will strike).
- W: The HFSA "Safer_RAIN" has a serious methodological weakness in the flood calculation for areas with a certain gradient, when there are no or hardly any local depressions in the DEM which is used for the calculation. In extreme cases, this may lead to very few and low inundation areas in the model, which can lead to a dangerous misinterpretation of the result, if the person who carried out the analysis has no technical expertise in the field of flood hazard calculation and blindly relies on the results.
- W: The algorithm and therefore also the tool are not able to deliver flow velocity, which is an important variable in assessing the flood hazard. For the same reason, it is neither possible to give an estimation on when a certain area will be flooded in rainfall scenario X. We only that it will be inundated or not.

The 2d finite-volumes hydrodynamic tool "Hydro-As-2D" with the "Sturzflut" extension in my opinion, has the following strengths (S) and weaknesses (W):

- S: 2d hydrodynamic flood calculation software which consider all terms of the shallow-water-equations is the **current technical state-of-the-art tool** for assessing pluvial flood

hazard. It is accurate and able to deliver reasonable results also in steep areas and / or areas with no depressions.

- S: 2d hydrodynamic flood calculation software can provide information not only on flood extent and the water depth at the final stage of the calculation, but also for the maximum depth of all time steps aggregated in one dataset, as well as the probable time, when an area X in is danger of being inundated, and of course flow velocity.
- S: Therefore, also further analysis can be made, when the preparatory work is done and the model is ready for use. So the high initial investment can pay off via the possibility to calculate follow-on-products.
- W: The necessary effort for the preparatory work is very high, the scarcer the data sources are, the higher the effort for finding useable datasets. A lot of manual correction work may incur in the DTM preprocessing (like the removal of bridges, or correcting narrow streambeds, where a DTM with poor resolution my produce errors).
- W: The software licenses for the hydrodynamic tools are very cost-intensive, as it is specialist software with high requirements and a low number of users in comparison to standard software for the mass market.
- W: The calculation itself is very computationally intensive, needs a lot of CPU and GPU power and RAM. Currently, I am not aware of any cloud solutions for the calculation.

5.2. Cost-benefit ratio & fitness for purpose

From the point of view of the AOI Carrefour, the HFSA “Safer_RAIN” is not a very good option for assessing the flood risk / flood hazard of the area in case of heavy precipitation events, as the algorithm is not able to deliver good and reliable results in terrains like this, with a steady gradient and only a few sinks / local depressions.

Also LISFLOOD-FP had serious problems with flood calculation for this terrain, at least with the “Acceleration” solver, where one term of the shallow-water-equations is being left out.

So in fact, for the Carrefour AOI and places with similar topography, there is actually only one alternative out of the three evaluated toolsets: The Hydro-As-2D software with the “Sturzflut” extension, which was able to deliver good and reliable results.

5.3. Conclusion and Outlook

Let us remember our **two central research questions** from chapter 1.2.2:

First, *how can precise and spatially high-resolution flood calculations be carried out in areas with scarce input data?*

As it had been shown, performing precise and spatially high-resolution flood calculations in areas with scarce input data can be very challenging, but this work has shown several steps that can be taken to tackle this issue:

- **Use of remote sensing datasets**: The remote sensing rainfall datasets experience a rapid change and development and the methodology and the sensors are becoming more and more an alternative to the exclusive use of in-situ-station data only. But one has to be careful when comparing the different products with each other and take into account concepts like the area reduction factor (ARF).
- **For the vectorization of buildings**, object-based image analysis (OBIA) would most probably lead to better results than the buildings extracted out of OpenStreetMap, as this user-generated content often has huge data gaps in data scarce regions. However, the extraction of critical infrastructure delivered credible results in the Carrefour AOI.
- **Validation of the results is often a problem**, when there are no local data sources against which they can be validated. Again, remote sensing may be an option for this, if one is able to find a suitable satellite image of the (preferably) peak flood extent (which may be very challenging, as rainfall cause the existence of clouds, and clouds are bad when you would like to have a clear sight on the ground. If you are lucky enough to find one, then the model output can be validated against the measured extent coming out of the DOPs. The same would be possible with geo-tagged photographs. Unfortunately, neither suitable sat images nor geo-tagged photographs have been available for this work and the AOI Carrefour. For a perfectly validated flood calculation model, you would make use of rain gauge station data, a discharge curve out of local river gauges for the validation of the model and high-resolution satellite images and / or geotagged photographs from the situation on the ground.
- This is also a **question of focus**: Which results am I expecting from my analysis? Do I need data output with a very high resolution and accuracy or do I just want to get a first impression of flood-prone areas in case of a heavy rain event, and a DTM with 20 or 30 meters grid cells and the assumption of block rainfall and impermeable soil will be enough? Do I need the output of flow velocity? Is it important to know the time of inundation for my area? All these are important factors for the decision making, which product is most likely to get me to my research goal.

*And **second**, what are the differences in flood hazard calculation between an easy-to-setup merely GIS based approach ("Safer_RAIN"), which can be calculated via a global web platform and a conventional 2d hydrodynamic flood calculation, carried out with the industry standard software "HYDRO-AS 2d"?*

- This question has been answered in high detail in the previous chapters – the differences between the two approaches lie in the methodology and mathematics on which they are basing. While Hydro-As-2D is a full 2D hydrodynamic modelling software, considering all terms of the shallow-water-equations, the hierarchical filling-and-spilling algorithm "Safer_RAIN" works completely different on a merely GIS-based calculation of flood

extent and water depth out of the filling and spilling of sinks and local depressions from the rainfall added into the model.

- This of course may have serious impacts on the resulting flood extent and water depth (other parameters cannot be compared, as the HFSA only delivers there two), especially when the terrain of the survey area is not flat and has a constant gradient, where the water added into the model quickly leaves the domain instead of accumulating in sinks and local depressions. The resulting difference in these two parameters (flood extent and water depth) can be tremendous, as shown in Table 12.
- Simplified, all merely GIS-based flood calculation approaches have one Achilles' heel in common: the extremely high dependence from the underlying DTM. The success of the algorithm is highly dependant from the quality and accuracy, and, in case of “Safer_RAIN”, also from the morphology of the terrain, whereas in 2d hydrodynamic models, other factors like roughness coefficient, infiltration (though both of these two parameters are integrated in “Safer-RAIN”), flow velocity and concentration time have a big influence on the result as well; the dependence on the underlying DTM as lower than in merely GIS-based approaches.

In my estimation, **for data-scarce regions like our AOI Carrefour and Haiti as a whole**, it could be a good way to

- **bundle the effort for data extraction:** tackle the initial effort once, get good and reliable datasets for the whole island of Hispaniola
- **validate against existing data-sources with known quality standards**,
- **recycle collected data:** re-use the collected and evaluated data sources as much as you can
- **make use of different rainfall scenarios:** set the focus on the calculation of several rainfall scenarios with different durations and intensities (i.e. rainfall intensities ranging from 50 to 130 mm/h), **compare them with past events** (via ground data if available and remote sensing datasets like MSWEP or ERA5/ERA5-Land) to get a classification of which rainfall scenario corresponds to which historical flood event (if there have been any) and then **calculate them all at once.** By this manner, responsible bodies and local authorities have a good foundation for delineating
  - a) **“No-build-areas”**: areas where there should better be a ban on the construction of new buildings or where special regulations become enacted, i.e. the construction of low walls to prevent water from entering the building
  - b) **“High-danger-areas”**: delineate the most flood-prone areas for rainfall events of the past and possible rainfall events in the future
  - c) **“Safe areas”**: delineate the least flood prone areas, where can people be evacuated to from inundated neighbourhoods, which areas are suitable for the construction of critical infrastructure like hospitals and schools,

- d) **“safe and unsafe passages”**: which roads are highly likely to be inundated at which rainfall scenario? Which roads on the other hand will most probably still be able to pass?
- c) **“the time factor”**: When will the most flood prone areas be in danger? Hydrodynamic models have the advantage that one (indirect) model output is time of inundation, which enables the local authorities and the rescue forces to assess when a certain building, area or road passage is in acute danger of getting flooded.

6. Appendix

6.1. Extraction of IDF Curves from GPEX dataset

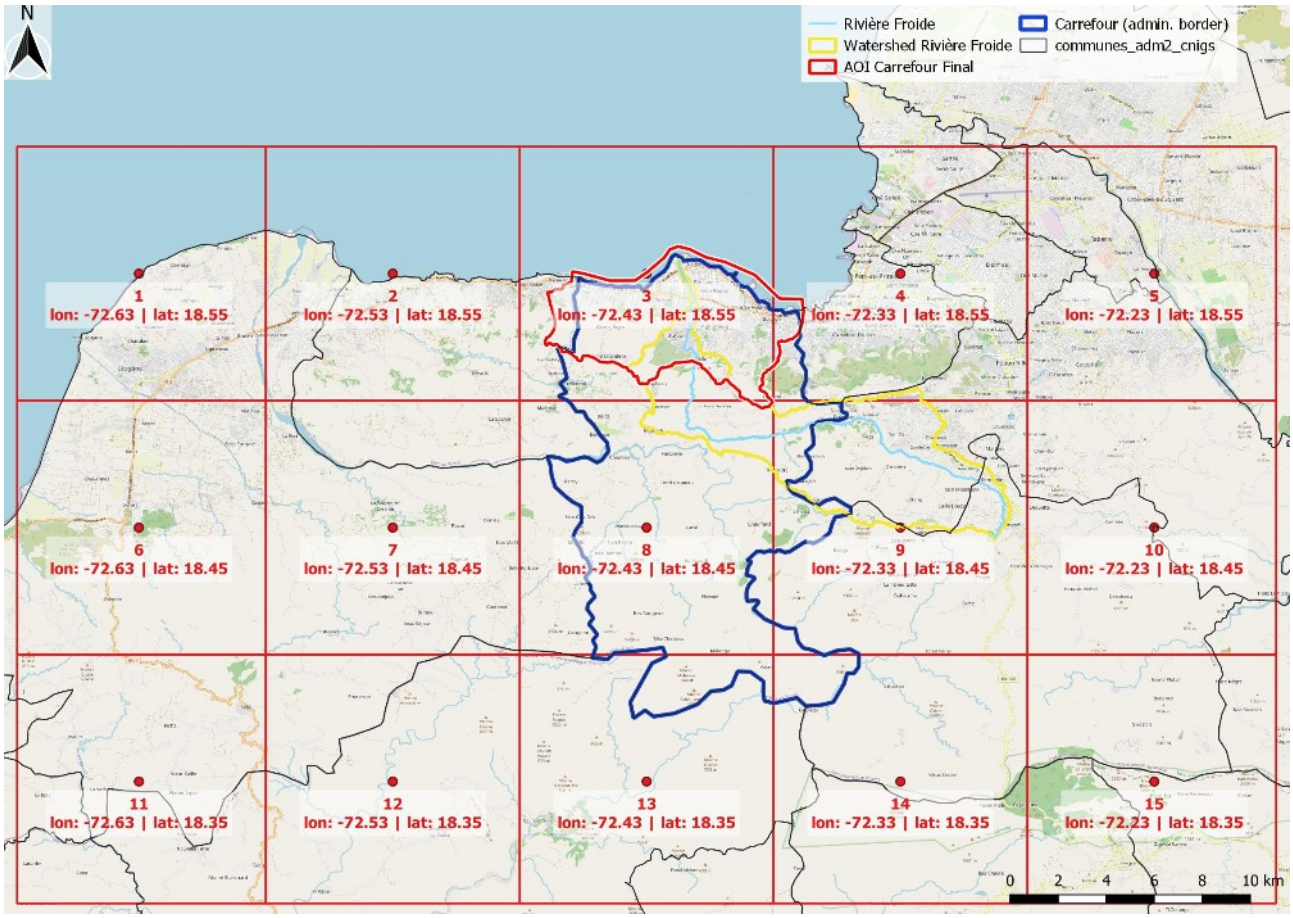
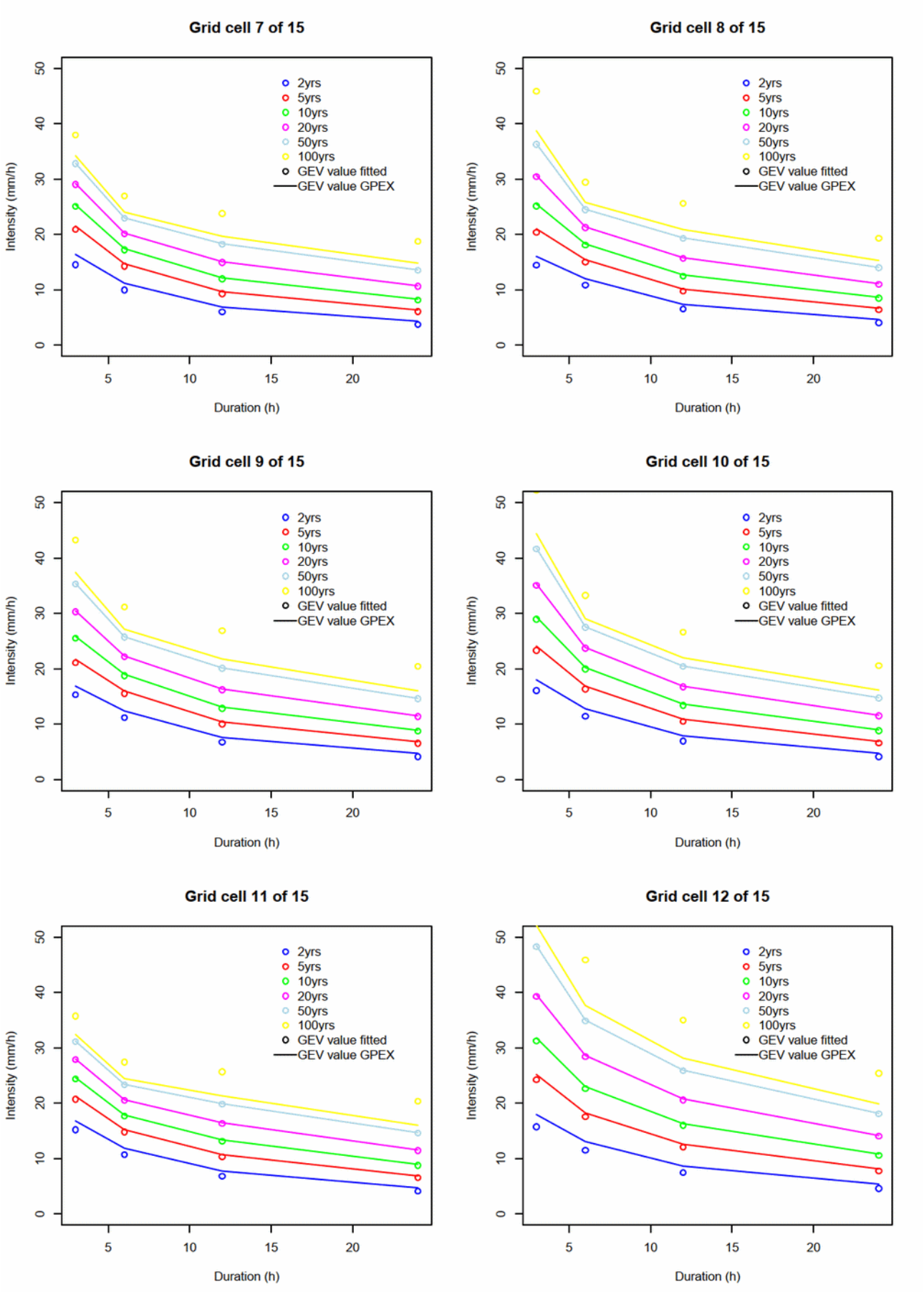


Figure 96: Position of the GPEX raster cells

6.1.1 GEV (GPEX) estimates vs. GEV fitted from observed values in GPEX dataset


Grid cell 7 of 15
Grid cell 8 of 15
Grid cell 9 of 15
Grid cell 10 of 15
Grid cell 11 of 15
Grid cell 12 of 15
2yrs
5yrs
10yrs
20yrs
50yrs
100yrs
GEV value fitted
GEV value GPEX
Intensity (mm/h)
Duration (h)
0
10
20
30
40
50
5
10
15
20

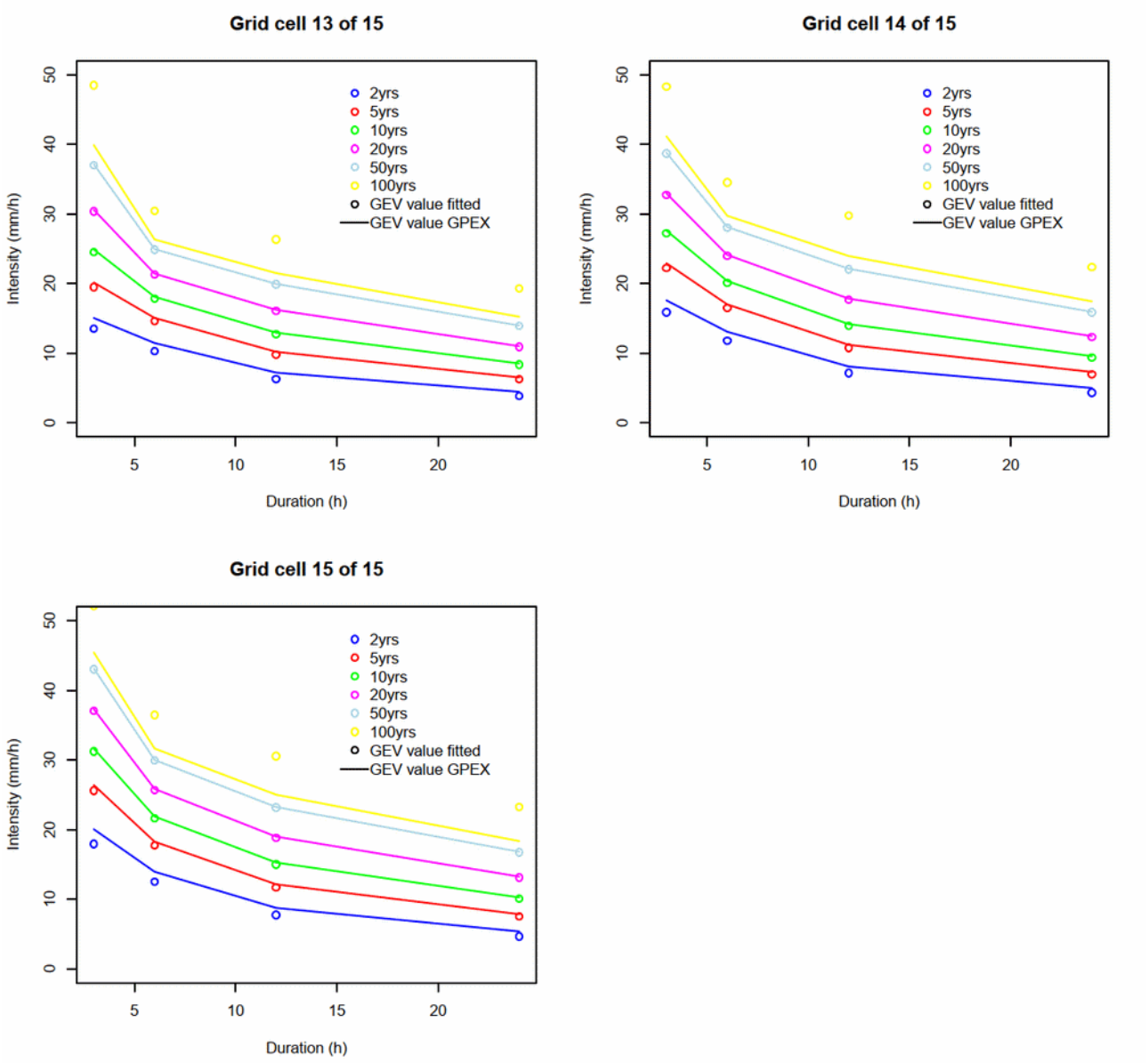
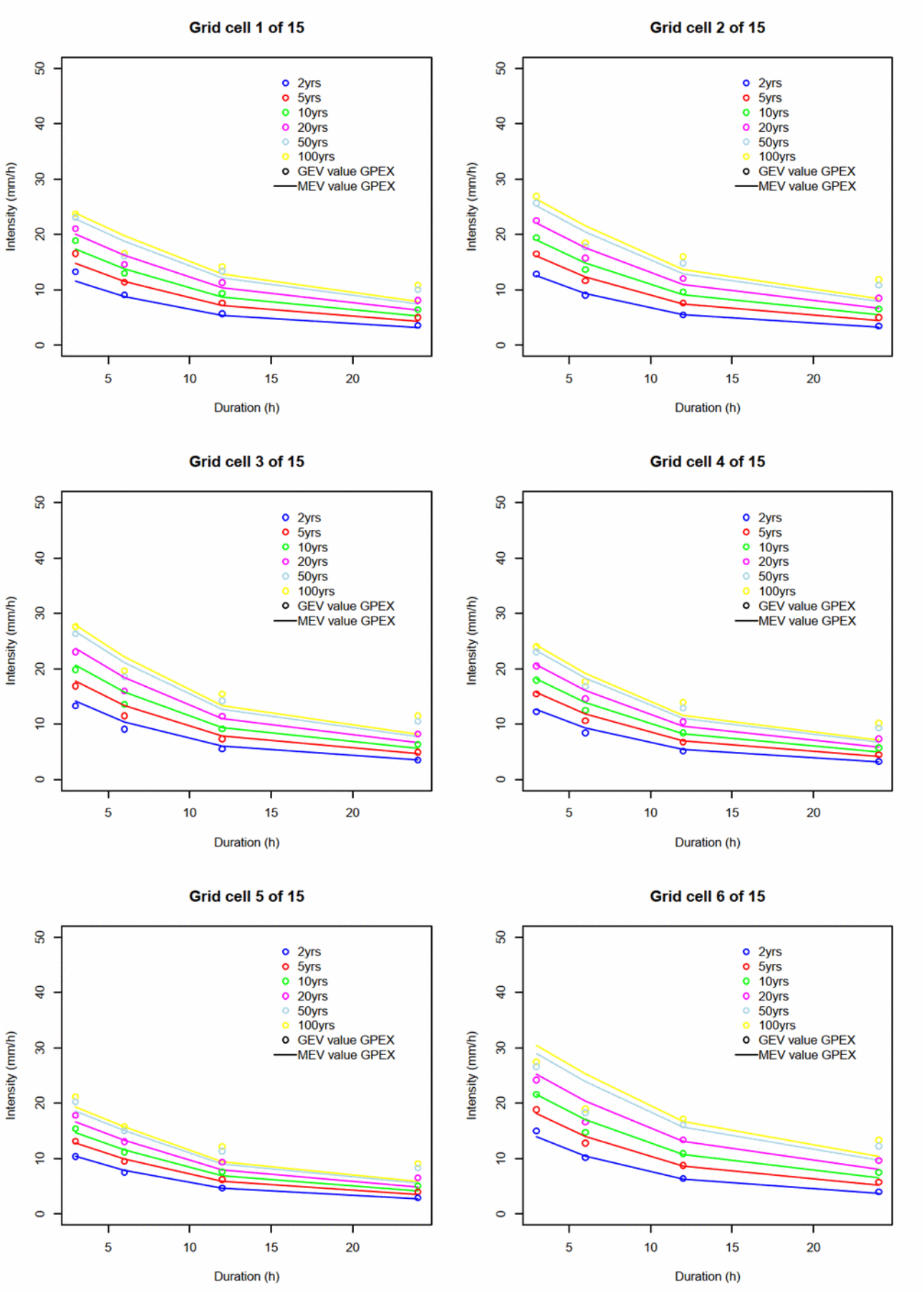
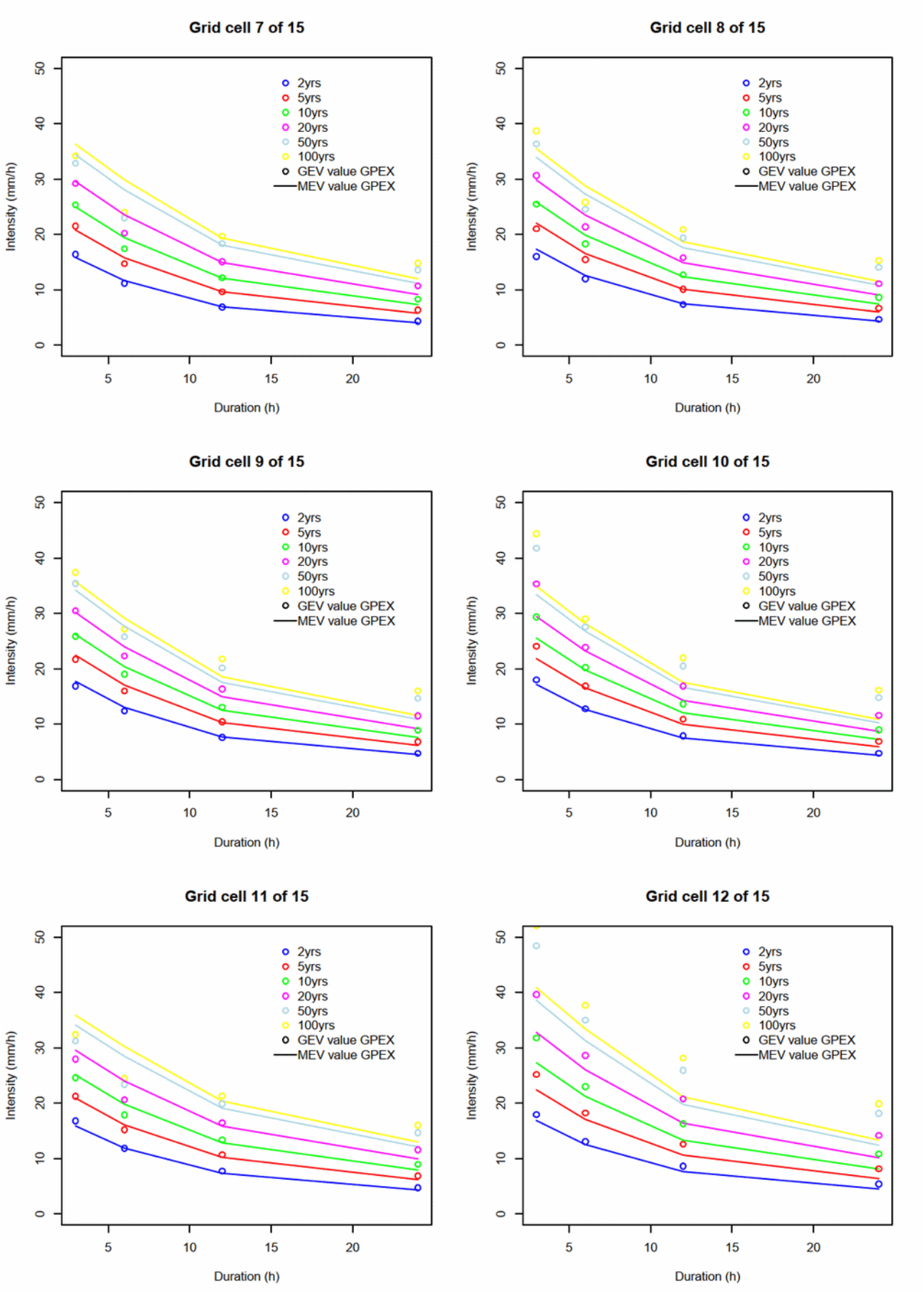


Figure 97: GEV (GPEX) estimates vs. GEV fitted from observed values in GPEX dataset

6.1.2 GEV (GPEX) estimates vs. MEV (GPEX) estimates:


Grid cell 1 of 15
Grid cell 2 of 15
Grid cell 3 of 15
Grid cell 4 of 15
Grid cell 5 of 15
Grid cell 6 of 15
2yrs
5yrs
10yrs
20yrs
50yrs
100yrs
GEV value GPEX
MEV value GPEX
Intensity (mm/h)
Duration (h)
0
10
20
30
40
50
5
10
15
20


Grid cell 7 of 15
Grid cell 8 of 15
Grid cell 9 of 15
Grid cell 10 of 15
Grid cell 11 of 15
Grid cell 12 of 15
Intensity (mm/h)
Duration (h)
2yrs
5yrs
10yrs
20yrs
50yrs
100yrs
GEV value GPEX
MEV value GPEX

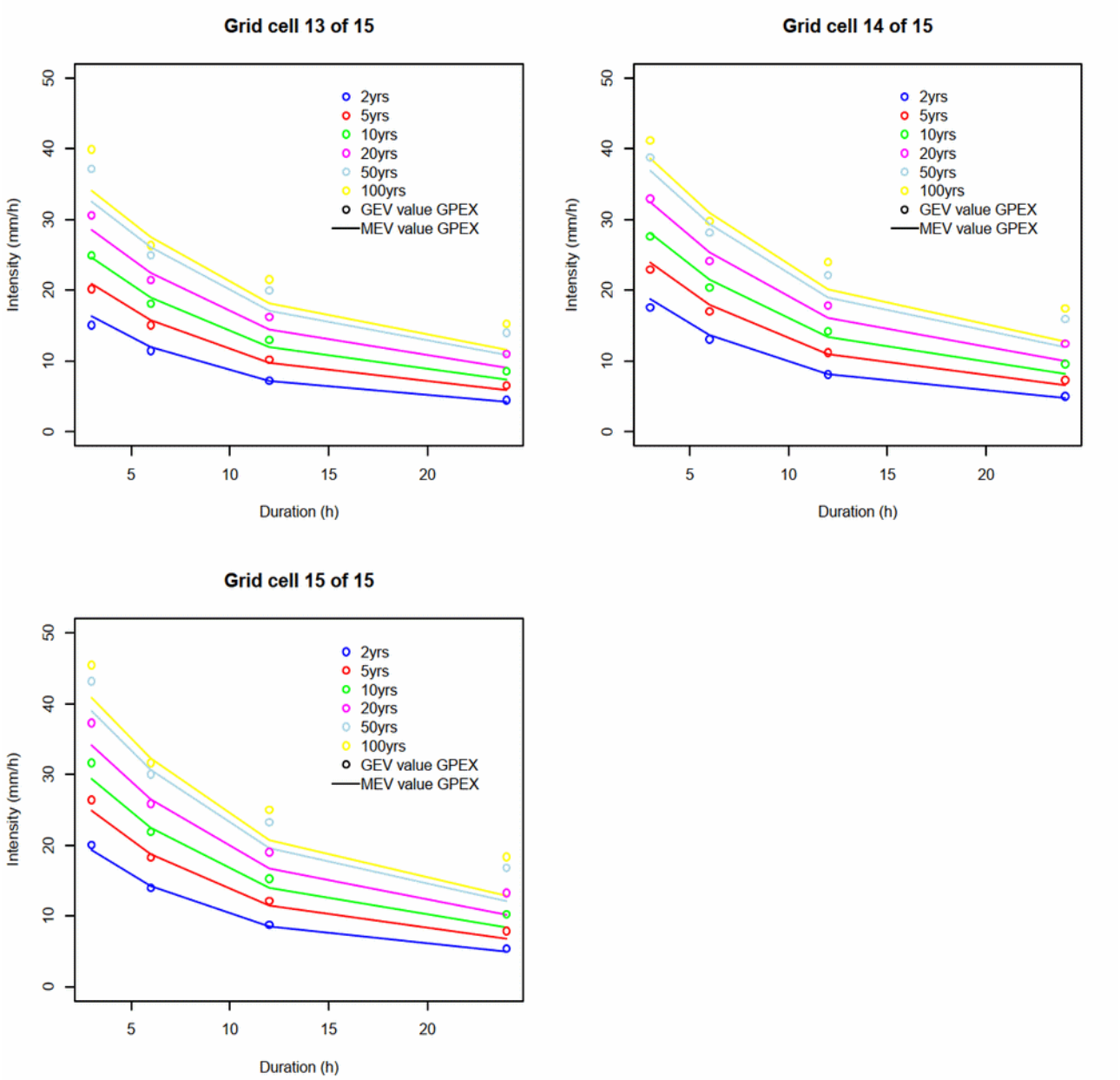
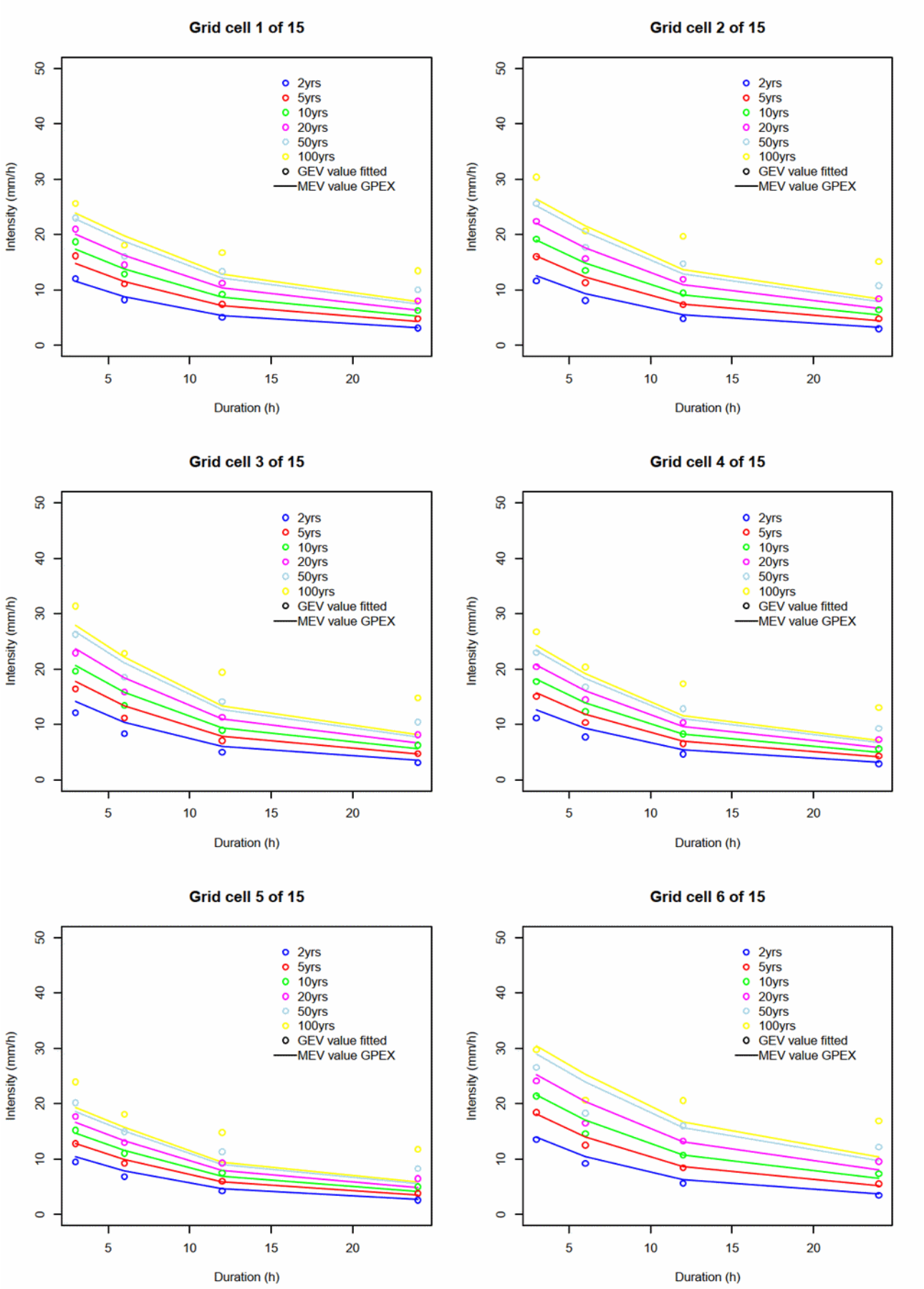
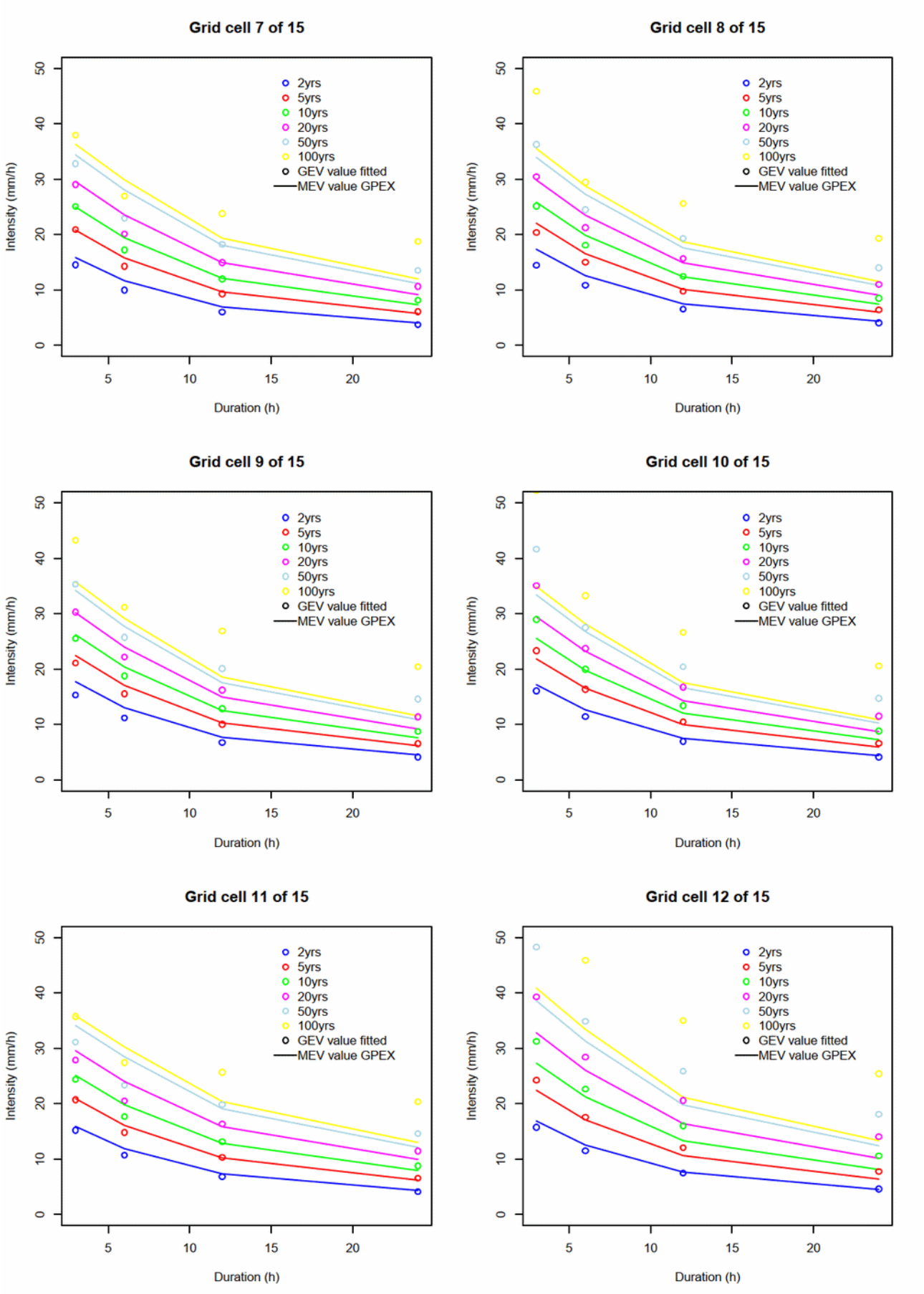


Figure 98: GEV (GPEX) estimates vs. MEV (GPEX) estimates

6.1.3 MEV (GPEX) estimates vs. GEV fitted from observed values in GPEX dataset:


Grid cell 1 of 15
Grid cell 2 of 15
Grid cell 3 of 15
Grid cell 4 of 15
Grid cell 5 of 15
Grid cell 6 of 15
2yrs
5yrs
10yrs
20yrs
50yrs
100yrs
GEV value fitted
MEV value GPEX
Intensity (mm/h)
Duration (h)
0
10
20
30
40
50
5
10
15
20


Grid cell 7 of 15
Grid cell 8 of 15
Grid cell 9 of 15
Grid cell 10 of 15
Grid cell 11 of 15
Grid cell 12 of 15
Intensity (mm/h)
Duration (h)
2yrs
5yrs
10yrs
20yrs
50yrs
100yrs
GEV value fitted
MEV value GPEX

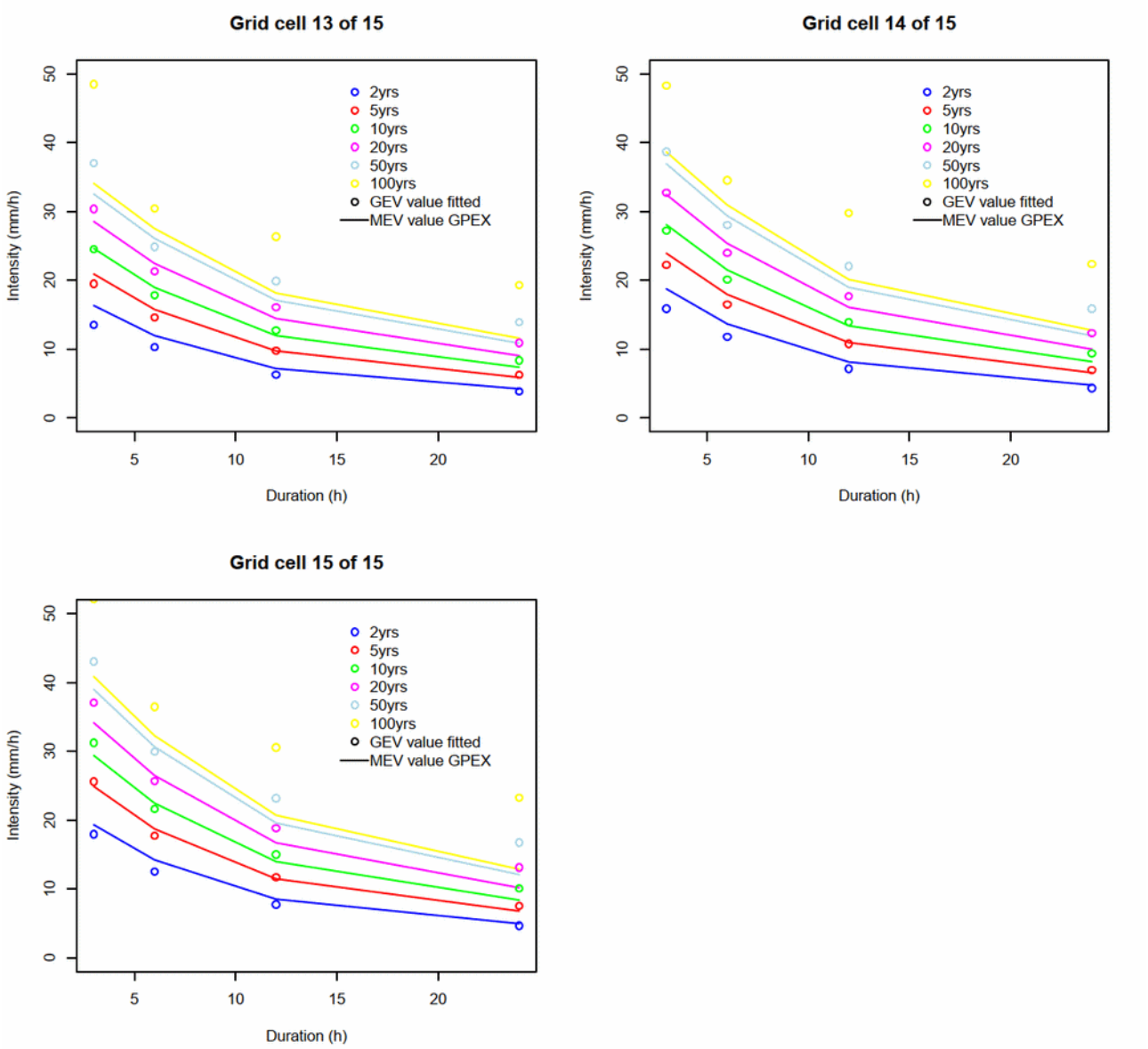


Figure 99: MEV (GPEX) estimates vs. GEV fitted from observed values in GPEX dataset

6.2. R Script used for the evaluation and of netCDF files and extraction of IDF Curves

```
library(ncdf4) library(RColorBrewer) library(rgdal) library(rgeos) library(extRemes) library(rnoaa)
library(readxl) library(raster)  ################################################ # Reading,
buffering and transforming shapefiles  ################################################## UG =
readOGR(dsn = "c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/BB_15_gridcells.shp")
crsmerc=CRS("+proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0") #UG = spTransform(UG, crsmerc)
#UG=UG(8:9,) # Departments Ouest and Sud-Est #https://de.wikipedia.org/wiki/D
%C3%A9partements_von_Haiti  AOI = readOGR(dsn =
"c:/UNIGIS_MSc/MASTERTHESIS/DATA/GIS/hti_boundaries_international_cnigs_polygon/AOI.shp")
AOI_buff=gBuffer(AOI,width=0.2) crsmerc=CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0") #AOI = spTransform(AOI, crsmerc) #AOI_buff = spTransform(AOI_buff, crsmerc)
#AOI_buff=gBuffer(AOI,width=0.1)  EZG = readOGR(dsn =
"c:/UNIGIS_MSc/MASTERTHESIS/DATA/GIS/hti_boundaries_international_cnigs_polygon/EZG_wgs84.shp")
EZG_buff=gBuffer(EZG,width=0.2) crsmerc=CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0") #EZG = spTransform(EZG, crsmerc) #EZG_buff = spTransform(EZG_buff, crsmerc)
plot(UG,border="black") plot(EZG,add=T,border="blue") plot(EZG_buff,add=T,border="lightblue")
plot(AOI,add=T, border="red") plot(AOI_buff,add=T,border="orange")  ############### # open netcdf
############### #Data are here
#https://opendap.4tu.nl/thredds/catalog/data2/fig/12764429/4/catalog.html  ncpath <-
"c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/data_old/" ncname <- "GPEX_extremes"  ncfname <-
paste(ncpath, ncname, ".nc", sep="") ncin <- nc_open(ncfname) print(ncin)  ncpath_val <-
"c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/data_old/" ncname_val <- "GPEX_parameters"  ncfname_val
<- paste(ncpath_val, ncname_val, ".nc", sep="") ncin_val <- nc_open(ncfname_val) print(ncin_val)  #
extract lat and lon lon <- ncvar_get(ncin,"lon") nlon <- dim(lon) #head(lon) lat <-
ncvar_get(ncin,"lat") nlat <- dim(lat) #head(lat)  lon_val <- ncvar_get(ncin_val,"lon") nlon_val <-
dim(lon_val) #head(lon) lat_val <- ncvar_get(ncin_val,"lat") nlat_val <- dim(lat_val) #head(lat)
################################################################## # Extract GEV estimates (15
points, 6 durations, 8 return periods)
################################################################## GEV_estimates=array(dim=c(15,6,8))
dname="gev" tmp_array <- ncvar_get(ncin,dname)  for (i in 1:8){ # return periods   print(i)   for
(ii in 1:6){ # durations     tmp_slice <- tmp_array(i,ii,,)     r <- raster(t(tmp_slice),
xmn=min(lon), xmx=max(lon), ymn=min(lat), ymx=max(lat),                  crs=CRS("+proj=longlat
+ellps=WGS84 +datum=WGS84 +no_defs+ towgs84=0,0,0"))     r2=raster::crop(r,UG)
GEV_estimates(,ii,i)=values(r2)   } }  #GEV_estimates=GEV_estimates(,,c(1,2,3,4,6,7)) # limit to
2,5,10,20,50,100,200 yrs
################################################################################ # Here I write out the
GEV quantiles from the paper as CSV for each grid cell
################################################################################ for (i in 1:15)
{   data_out=as.data.frame(round(sweep(GEV_estimates(i,,),1,c(3, 6, 12, 24, 48, 72),"/"),2))
rownames(data_out)=paste(c(3,6,12,24,48,72),"h",sep="")
colnames(data_out)=paste(c(2,5,10,20,39,50,100,200),"yrs",sep="")
write.csv(data_out,paste("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/n15_GEV_grid
cell_",toString(i),"_quantiles_intensities_mmh",".csv",sep=""),quote = F) }
################################################################## # Extract MEV estimates (15
points, 6 durations, 8 return periods)
################################################################## MEV_estimates=array(dim=c(15,6,8))
dname="mev" tmp_array <- ncvar_get(ncin,dname)  for (i in 1:8){ # return periods   print(i)   for
(ii in 1:6){ # durations     tmp_slice <- tmp_array(i,ii,,)     r <- raster(t(tmp_slice),
xmn=min(lon), xmx=max(lon), ymn=min(lat), ymx=max(lat),                  crs=CRS("+proj=longlat
+ellps=WGS84 +datum=WGS84 +no_defs+ towgs84=0,0,0"))     r2=raster::crop(r,UG)
MEV_estimates(,ii,i)=values(r2)   } }
################################################################################ # Here I write out the
MEV quantiles from the paper as CSV for each grid cell
################################################################################ for (i in 1:15)
{   data_out=as.data.frame(round(sweep(MEV_estimates(i,,),1,c(3, 6, 12, 24, 48, 72),"/"),2))
rownames(data_out)=paste(c(3,6,12,24,48,72),"h",sep="")
colnames(data_out)=paste(c(2,5,10,20,39,50,100,200),"yrs",sep="")
write.csv(data_out,paste("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/n15_MEV_grid
cell_",toString(i),"_quantiles_intensities_mmh",".csv",sep=""),quote = F) }
############################################################################################## #
Extract Annual maxima (38 years,15 points,8 durations) and calculate GEV quantile manually
##############################################################################################
amax=array(dim=c(38,15,8)) dname="annual_maximum" tmp_array <- ncvar_get(ncin_val,dname)  for (i in
1:38){ # years   print(i)   for (ii in 1:8){ # durations     tmp_slice <- tmp_array(i,ii,,)     r <-
raster(t(tmp_slice), xmn=min(lon_val), xmx=max(lon_val), ymn=min(lat_val), ymx=max(lat_val),
crs=CRS("+proj=longlat +ellps=WGS84 +datum=WGS84 +no_defs+ towgs84=0,0,0"))
r2=raster::crop(r,UG)     amax(i,,ii)=values(r2)   } }  # Here we fit the GEV ourselves to compare
the quantiles with the published values (code above) library(extRemes)
```

```
GEV_fit_estimates=array(dim=c(15,8,6)) for (i in 1:15){   print(i)   for (ii in 1:8){ # durations
gev_fit=extRemes::fevd(amax(!is.na(amax(,i,ii)),i,ii),#/dur_hour(ii),
                           type="GEV",method="Lmoments")
                           #type="Gumbel",method="MLE")
    gev_RP <- return.level(gev_fit, conf = 0.05, return.period= c(2,5,10,20,39,100))
    GEV_fit_estimates(i,ii,)=gev_RP
  }
}


#####################################################
# Compare IDF from GEV (GPEX) to IDF from GEV (fitted in R from observed values)
#####################################################
# Check if estimates and fitted GEVs are similar
dim(GEV_estimates)
dim(GEV_fit_estimates)

plot(GEV_estimates,GEV_fit_estimates,xlab="estimates from GPEX",ylab="estimates from fitted GEV to
amax")
abline(0,1)

#############################################
# Other plot as points and lines and as PDF #
#############################################
# set duration and return period values
dur_hour=c(3, 6, 12, 24, 48, 72, 120,240)
rp=c(2,5,10,20,39,100)
cols=c("blue","red","green","magenta","lightblue","yellow")

# create a PDF file to save the plots - page 1
pdf("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/GEV_stats_01.pdf", width=8.27,height=11.69) # size
of DIN-A4
par(mfrow=c(3,2))

# loop through each grid cell - cells 1-6 for page 1
for (st in 1:6){
  i=6
  plot(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),ylim=c(0,50),col=cols(6),
       xlab="Duration (h)",ylab="Intensity (mm/h)",main=paste("Grid cell ",toString(st)," of
15",sep=""))
  lines(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(6))

  # create a legend for the plot
  legend_text=c("2yrs","5yrs","10yrs","20yrs","50yrs","100yrs","GEV value fitted","GEV value GPEX")
  pos=legend(15,50,legend_text,           #15,68,legend_text,
             bty = "n",lty=c(rep(NA,6),1,1),cex=rep(1,8),pt.bg =
"white",pch=c(rep(1,6),1,NA),lwd=c(rep(NA,6),NA,1),
             bg="white",ncol=1,x.intersp=0.1,y.intersp=1,col=c(cols,1,1))

  # loop through each return period
  for (i in 5:1){
    points(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),
           xlab="Duration (h)",ylab="Intensity (mm/h)",col=cols(i))
    lines(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(i))
  }
}
dev.off() #saves the pdf


# create a PDF file to save the plots - page 2
pdf("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/GEV_stats_02.pdf", width=8.27,height=11.69) # size
of DIN-A4
par(mfrow=c(3,2))

# loop
 through each grid cell - cells 7-12 for page 2
for (st in 7:12){
  i=6
  plot(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),ylim=c(0,50),col=cols(6),
       xlab="Duration (h)",ylab="Intensity (mm/h)",main=paste("Grid cell ",toString(st)," of
15",sep=""))
  lines(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(6))

  # create a legend for the plot
```

```
  legend_text=c("2yrs","5yrs","10yrs","20yrs","50yrs","100yrs","GEV value fitted","GEV value GPEX")
  pos=legend(15,50,legend_text,           #15,68,legend_text,
             bty = "n",lty=c(rep(NA,6),1,1),cex=rep(1,8),pt.bg =
"white",pch=c(rep(1,6),1,NA),lwd=c(rep(NA,6),NA,1),
             bg="white",ncol=1,x.intersp=0.1,y.intersp=1,col=c(cols,1,1))

  # loop through each return period
  for (i in 5:1){
    points(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),
           xlab="Duration (h)",ylab="Intensity (mm/h)",col=cols(i))
    lines(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(i))
  }
}
dev.off() #saves the pdf


# create a PDF file to save the plots - page 3
pdf("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/GEV_stats_03.pdf", width=8.27,height=11.69) # size
of DIN-A4
par(mfrow=c(3,2))

# loop through each grid cell - cells 13-15 for page 3
for (st in 13:15){
  i=6
  plot(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),ylim=c(0,50),col=cols(6),
       xlab="Duration (h)",ylab="Intensity (mm/h)",main=paste("Grid cell ",toString(st)," of
15",sep=""))
  lines(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(6))

  # create a legend for the plot
  legend_text=c("2yrs","5yrs","10yrs","20yrs","50yrs","100yrs","GEV value fitted","GEV value GPEX")
  pos=legend(15,50,legend_text,           #15,68,legend_text,
             bty = "n",lty=c(rep(NA,6),1,1),cex=rep(1,8),pt.bg =
"white",pch=c(rep(1,6),1,NA),lwd=c(rep(NA,6),NA,1),
             bg="white",ncol=1,x.intersp=0.1,y.intersp=1,col=c(cols,1,1))

  # loop through each return period
  for (i in 5:1){
    points(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),
           xlab="Duration (h)",ylab="Intensity (mm/h)",col=cols(i))
    lines(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(i))
  }
}
dev.off() #saves the pdf


#################################################################################################################
#
# Compare IDF from GEV (GPEX) to IDF from MEV (GPEX) and IDF from GEV (fitted in R from obs values)
#
#################################################################################################################
#
# Check if GEV estimates and MEV estimates are similar / Check MEV estimates to fitted GEVs
dim(GEV_estimates)
dim(MEV_estimates)
dim(GEV_fit_estimates)

plot(GEV_estimates,MEV_estimates,xlab="GEV estimates GPEX",ylab="MEV estimates GPEX")
abline(0,1)

plot(MEV_estimates,GEV_fit_estimates,xlab="MEV estimates GPEX",ylab="estimates from fitted GEV to
amax")
abline(0,1)

###############################################
# IDF curves MEV to GEV and MEV to fitted GEV #
###############################################
# set duration and return period values
dur_hour=c(3, 6, 12, 24, 48, 72, 120,240)
rp=c(2,5,10,20,39,100)
cols=c("blue","red","green","magenta","lightblue","yellow")

# create a PDF file to save the plots - page 4
pdf("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/GEV_stats_04.pdf", width=8.27,height=11.69) # size
```

```
of DIN-A4
par(mfrow=c(3,2))

# loop through each grid cell - cells 1-6 for page 4
for (st in 1:6){
  i=6
  plot(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),ylim=c(0,50),col=cols(6),
       xlab="Duration (h)",ylab="Intensity (mm/h)",main=paste("Grid cell ",toString(st)," of
15",sep=""))
  lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(6))

  # create a legend for the plot
  legend_text=c("2yrs","5yrs","10yrs","20yrs","50yrs","100yrs","GEV value GPEX","MEV value GPEX")
  pos=legend(15,50,legend_text,           #15,68,legend_text,
             bty = "n",lty=c(rep(NA,6),1,1),cex=rep(1,8),pt.bg =
"white",pch=c(rep(1,6),1,NA),lwd=c(rep(NA,6),NA,1),
             bg="white",ncol=1,x.intersp=0.1,y.intersp=1,col=c(cols,1,1))

  # loop through each return period
  for (i in 5:1){
    points(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),
           xlab="Duration (h)",ylab="Intensity (mm/h)",col=cols(i))
    lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(i))
  }
}
dev.off() #saves the pdf


# create a PDF file to save the plots - page 5
pdf("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/GEV_stats_05.pdf", width=8.27,height=11.69) # size
of DIN-A4
par(mfrow=c(3,2))

# loop through each grid cell - cells 7-12 for page 5
for (st in 7:12){
  i=6
  plot(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),ylim=c(0,50),col=cols(6),
       xlab="Duration (h)",ylab="Intensity (mm/h)",main=paste("Grid cell ",toString(st)," of
15",sep=""))
  lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(6))

  # create a legend for the plot
  legend_text=c("2yrs","5yrs","10yrs","20yrs","50yrs","100yrs","GEV value GPEX","MEV value GPEX")
  pos=legend(15,50,legend_text,           #15,68,legend_text,
             bty = "n",lty=c(rep(NA,6),1,1),cex=rep(1,8),pt.bg =
"white",pch=c(rep(1,6),1,NA),lwd=c(rep(NA,6),NA,1),
             bg="white",ncol=1,x.intersp=0.1,y.intersp=1,col=c(cols,1,1))

  # loop through each return period
  for (i in 5:1){
    points(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),
           xlab="Duration (h)",ylab="Intensity (mm/h)",col=cols(i))
    lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(i))
  }
}
dev.off() #saves the pdf


# create a PDF file to save the plots - page 6
pdf("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/GEV_stats_06.pdf", width=8.27,height=11.69) # size
of DIN-A4
par(mfrow=c(3,2))

# loop through each grid cell - cells 13-15 for page 6
for (st in 13:15){
  i=6
  plot(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),ylim=c(0,50),col=cols(6),
       xlab="Duration (h)",ylab="Intensity (mm/h)",main=paste("Grid cell ",toString(st)," of
15",sep=""))
  lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(6))

  # create a legend for the plot
  legend_text=c("2yrs","5yrs","10yrs","20yrs","50yrs","100yrs","GEV value GPEX","MEV value GPEX")
  pos=legend(15,50,legend_text,           #15,68,legend_text,
```

```
             bty = "n",lty=c(rep(NA,6),1,1),cex=rep(1,8),pt.bg =
"white",pch=c(rep(1,6),1,NA),lwd=c(rep(NA,6),NA,1),
             bg="white",ncol=1,x.intersp=0.1,y.intersp=1,col=c(cols,1,1))

  # loop through each return period
  for (i in 5:1){
    points(c(3, 6, 12, 24),GEV_estimates(st,1:4,i)/dur_hour(1:4),
           xlab="Duration (h)",ylab="Intensity (mm/h)",col=cols(i))
    lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(i))
  }
}
dev.off() #saves the pdf

#############################################################################################################

# create a PDF file to save the plots - page 7
pdf("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/GEV_stats_07.pdf", width=8.27,height=11.69) # size
of DIN-A4
par(mfrow=c(3,2))

# loop through each grid cell - cells 1-6 for page 7
for (st in 1:6){
  i=6
  plot(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),ylim=c(0,50),col=cols(6),
       xlab="Duration (h)",ylab="Intensity (mm/h)",main=paste("Grid cell ",toString(st)," of
15",sep=""))
  lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(6))

  # create a legend for the plot
  legend_text=c("2yrs","5yrs","10yrs","20yrs","50yrs","100yrs","GEV value fitted","MEV value GPEX")
  pos=legend(15,50,legend_text,           #15,68,legend_text,
             bty = "n",lty=c(rep(NA,6),1,1),cex=rep(1,8),pt.bg =
"white",pch=c(rep(1,6),1,NA),lwd=c(rep(NA,6),NA,1),
             bg="white",ncol=1,x.intersp=0.1,y.intersp=1,col=c(cols,1,1))

  # loop through each return period
  for (i in 5:1){
    points(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),
           xlab="Duration (h)",ylab="Intensity (mm/h)",col=cols(i))
    lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(i))
  }
}
dev.off() #saves the pdf


# create a PDF file to save the plots - page 8
pdf("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/GEV_stats_08.pdf", width=8.27,height=11.69) # size
of DIN-A4
par(mfrow=c(3,2))

# loop through each grid cell - cells 7-12 for page 8
for (st in 7:12){
  i=6
  plot(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),ylim=c(0,50),col=cols(6),
       xlab="Duration (h)",ylab="Intensity (mm/h)",main=paste("Grid cell ",toString(st)," of
15",sep=""))
  lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(6))

  # create a legend for the plot
  legend_text=c("2yrs","5yrs","10yrs","20yrs","50yrs","100yrs","GEV value fitted","MEV value GPEX")
  pos=legend(15,50,legend_text,           #15,68,legend_text,
             bty = "n",lty=c(rep(NA,6),1,1),cex=rep(1,8),pt.bg =
"white",pch=c(rep(1,6),1,NA),lwd=c(rep(NA,6),NA,1),
             bg="white",ncol=1,x.intersp=0.1,y.intersp=1,col=c(cols,1,1))

  # loop through each return period
  for (i in 5:1){
    points(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),
           xlab="Duration (h)",ylab="Intensity (mm/h)",col=cols(i))
    lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(i))
  }
}
dev.off() #saves the pdf
```

```
# create a PDF file to save the plots - page 9
pdf("c:/UNIGIS_MSc/MASTERTHESIS/DATA/Regen/GPEX/GEV_stats_09.pdf", width=8.27,height=11.69) # size
of DIN-A4
par(mfrow=c(3,2))

# loop through each grid cell - cells 13-15 for page 9
for (st in 13:15){
  i=6
  plot(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),ylim=c(0,50),col=cols(6),
       xlab="Duration (h)",ylab="Intensity (mm/h)",main=paste("Grid cell ",toString(st)," of
15",sep=""))
  lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(6))


 # create a legend for the plot
 legend_text=c("2yrs","5yrs","10yrs","20yrs","50yrs","100yrs","GEV value fitted","MEV value GPEX")
 pos=legend(15,50,legend_text,           #15,68,legend_text,
 bty = "n",lty=c(rep(NA,6),1,1),cex=rep(1,8),pt.bg =
"white",pch=c(rep(1,6),1,NA),lwd=c(rep(NA,6),NA,1),
 bg="white",ncol=1,x.intersp=0.1,y.intersp=1,col=c(cols,1,1))
  # loop through each return period
 for (i in 5:1){
 points(c(3, 6, 12, 24),GEV_fit_estimates(st,1:4,i)/dur_hour(1:4),
 xlab="Duration (h)",ylab="Intensity (mm/h)",col=cols(i))
 lines(c(3, 6, 12, 24),MEV_estimates(st,1:4,i)/dur_hour(1:4),col=cols(i))
 }
}
dev.off() #saves the pdf
```

Baumann, W.W., Bunge, U., Frederich, O., Schatz, M., Thiele, F., 2006. Finite–Volumen–Methode in der Numerischen Thermofluiddynamik.

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., Dijk, A.I.J.M. van, McVicar, T.R., Adler, R.F., 2019. MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. Bull. Am. Meteorol. Soc. 100, 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

Bonnin, G.M., Martin, D., Lin, B., Parzybok, T., 2008. NOAA Atlas 14 Precipitation-Frequency Atlas of the United States Volume 3 Version 4.0: Puerto Rico and the U.S. Virgin Islands.

Bradbrook, K., 2006. JFLOW: a multiscale two-dimensional dynamic flood model. Water Environ. J. 20, 79–86. https://doi.org/10.1111/j.1747-6593.2005.00011.x

Breinl, K., Müller-Thomy, H., Blöschl, G., 2020. Space-time characteristics of areal reduction factors and rainfall processes. J. Hydrometeorol. 21, 671–689. https://doi.org/10.1175/JHM-D-19-0228.1

Cantir, C., 2017. “Savages in the midst”: Revolutionary Haiti in international society (1791-1838). J. Int. Relat. Dev. 20, 238–261. https://doi.org/10.1057/jird.2015.7

Dimitriadis, P., Tegos, A., Oikonomou, A., Pagana, V., Koukouvinos, A., Mamassis, N., Koutsoyiannis, D., Efstratiadis, A., 2016. Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping. J. Hydrol. 534, 478–492. https://doi.org/10.1016/j.jhydrol.2016.01.020

Dolisca, F., McDaniel, J.M., Teeter, L.D., Jolly, C.M., 2007. Land tenure, population pressure, and deforestation in Haiti: The case of Foret des Pins Reserve. J. For. Econ. 13, 277–289. https://doi.org/10.1016/j.jfe.2007.02.006

ERA5-Land hourly data from 1981 to present, 2019. https://doi.org/10.24381/CDS.E2161BAC

Ferrer, A., 2012. Haiti, Free Soil, and Antislavery in the Revolutionary Atlantic. Am. Hist. Rev. 117, 40–66. https://doi.org/10.1086/ahr.117.1.40

Gruendemann, G., Zorzetto, E., Beck, H., Schleiss, M., Giesen, N.V. de, 2020. GPEX: Global Precipitation EXtremes. https://doi.org/10.4121/uuid:12b5c941-cd54-45db-8d7b-fefaacecaa69

Gründemann, G.J., Zorzetto, E., Beck, H.E., Schleiss, M., Van de Giesen, N., Marani, M., van der Ent, R.J., 2021. Extreme Precipitation Return Levels for Multiple Durations on a Global Scale (preprint). Hydrology. https://doi.org/10.1002/essoar.10503814.3

Hannemann, J.-C., 2014. mediaTUM - Media and Publication Server [WWW Document]. Mediat. Univ. Tech. Univ. Münch. URL https://mediatum.ub.tum.de/1449706 (accessed 5.11.21).

Hedges, S.B., Cohen, W.B., Timyan, J., Yang, Z., 2018. Haiti's biodiversity threatened by nearly complete loss of primary forest. Proc. Natl. Acad. Sci. 115, 11850–11855. https://doi.org/10.1073/pnas.1809753115

Heimhuber, V., Hannemann, J.-C., Rieger, W., 2015. Flood Risk Management in Remote and Impoverished Areas—A Case Study of Onaville, Haiti. Water 7, 3832–3860. https://doi.org/10.3390/w7073832

Humer, G., Reithofer, A., Klar, R., Achleitner, S., 2015. Erweiterung eines 2D-Strömungsmodelles zur Berechnung von Sturzfluten. Hydrol. Wasserbewirtsch. 2015. https://doi.org/10.3243/kwe2015.03.002

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. J. Hydrometeorol. 5, 487–503. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Klar, R., Achleitner, S., Lumassegger, S., Aufleger, M., Hofer, M., 2014. Extension And Testing Of A 2D Hydrodynamic Model For Direct Rainfall Runoff Simulation.

Koutsoyiannis, D., Kozonis, D., Manetas, A., 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. J. Hydrol. 206, 118–135. https://doi.org/10.1016/S0022-1694(98)00097-3

Lamb, R., Crossley, M., Waller, S., 2009. A fast two-dimensional floodplain inundation model. Proc. Inst. Civ. Eng. - Water Manag. 162, 363–370. https://doi.org/10.1680/wama.2009.162.6.363

Landesanstalt für Umweltschutz Baden-Württemberg, 2002. Hydraulik naturnaher Fließgewässer Teil 2: Neue Berechnungsverfahren für naturnahe Gewässerstrukturen.

Langousis, A., 2005. The Areal Reduction Factor (ARF) : a multifractal analysis (Thesis). Massachusetts Institute of Technology.

Marra, F., Morin, E., Peleg, N., Mei, Y., Anagnostou, E.N., 2017. Intensity-duration-frequency curves from remote sensing rainfall estimates: Comparing satellite and weather radar over the eastern Mediterranean. Hydrol. Earth Syst. Sci. 21, 2389–2404. https://doi.org/10.5194/hess-21-2389-2017

Menne, M.J., Durre, I., Korzeniewski, B., 2012a. Global Historical Climatology Network - Daily (GHCN-Daily), Version 3.

Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E., Houston, T.G., 2012b. An Overview of the Global Historical Climatology Network-Daily Database. J. Atmospheric Ocean. Technol. 29, 897–910. https://doi.org/10.1175/JTECH-D-11-00103.1

Mora, S., Roumagnac, A., Asté, J.-P., Calais, E., Haase, J., Saborío, J., Marcello, M., Milcé, J.-E., Zahibo, N., 2010. COMPOSITION OF THE MULTI-MENACES-HA TEAM 63.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Obregón, L., 2018. Empire, Racial Capitalism and International Law: The Case of Manumitted Haiti and the Recognition Debt. Leiden J. Int. Law 31, 597–615. https://doi.org/10.1017/S0922156518000225

Pauleus, O., Aide, T.M., 2020. Haiti has more forest than previously reported: Land change 2000-2015. PeerJ 8. https://doi.org/10.7717/peerj.9919

Posner, A.J., Georgakakos, K.P., 2017. Quantifying the impact of community-scale flood mitigation. Int. J. Disaster Risk Reduct. 24, 189–208. https://doi.org/10.1016/j.ijdrr.2017.06.001

Ravazzani, G., Boscarello, L., Cislaghi, A., Mancini, M., 2019. Review of Time-of-Concentration Equations and a New Proposal in Italy. J. Hydrol. Eng. 24. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001818

Samela, C., Persiano, S., Bagli, S., Luzzi, V., Mazzoli, P., Humer, G., Reithofer, A., Essenfelder, A., Amadio, M., Mysiak, J., Castellarin, A., 2020. Safer_RAIN: A DEM-Based Hierarchical Filling-&-Spilling Algorithm for Pluvial Flood Hazard Assessment and Mapping across Large Urban Areas. Water 12, 1514. https://doi.org/10.3390/w12061514

Shaw, J., Kesserwani, G., Neal, J., Bates, P., Sharifian, M.K., 2021. LISFLOOD-FP 8.0: the new discontinuous Galerkin shallow-water solver for multi-core CPUs and GPUs. Geosci. Model Dev. 14, 3577–3602. https://doi.org/10.5194/gmd-14-3577-2021

Shrestha, A., 2013. Impact of Climate Change on Urban Flooding in Sukhumvit Area of Bangkok. https://doi.org/10.13140/RG.2.2.27839.00160

Smucker, G., Timyan, J., Ward, C., 2017. Participatory Watershed Management Planning Methodology [WWW Document]. URL https://www.profor.info/countries/haiti (accessed 10.30.22).

Svensson, C., Jones, D., 2010. Review of methods for deriving area reduction factors. J. Flood Risk Manag. 3, 232–245. https://doi.org/10.1111/j.1753-318X.2010.01075.x

Tarter, A., Freeman, K.K., Ward, C., Sander, K., Theus, K., Coello, B., Fawaz, Y., Miles, M., Ahmed, T.T.G., 2018. Charcoal in Haiti. World Bank, Washington, DC. https://doi.org/10.1596/31257

Torres-Sierra, H., 1996. Flood of January 5-6, 1992, in Puerto Rico (No. 95–374), U.S. GEOLOGICAL SURVEY Open-File Report. U.S. GEOLOGICAL SURVEY.

Vreugdenhil, C.B., 1994. Numerical Methods for Shallow-Water Flow, Water Science and Technology Library. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-015-8354-1

Wang, J., Gonzalez, H., Lindenbergh, R., Arias, P., Menenti, M., 2014. Geometric road runoff estimation from laser mobile mapping data, ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. https://doi.org/10.5194/isprsannals-II-5-385-2014

Winsemius, H.C., Aerts, J.C.J.H., van Beek, L.P.H., Bierkens, M.F.P., Bouwman, A., Jongman, B., Kwadijk, J.C.J., Ligtvoet, W., Lucas, P.L., van Vuuren, D.P., Ward, P.J., 2016. Global drivers of future river flood risk. Nat. Clim. Change 6, 381–385. https://doi.org/10.1038/nclimate2893

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F., Lin, R., 2017. Reprocessed, Bias-Corrected CMORPH Global High-Resolution Precipitation Estimates from 1998. J. Hydrometeorol. 18, 1617–1641.