Urban flooding in Oslo – Insurance claims and heavy rainfall

Hand-in date:
01.09.2022

Campus:

BI Oslo

Supervisor:

Wei-Ting Yang

Examination code:

GRA 19703 – Master Thesis

Program:

Master of Science in Business Analytics

Summary

Urban flooding events and damages to private and public buildings and infrastructure is expected to increase due to higher frequency of heavy rainfall. Even if all countries were to fulfill their declared reductions in emission of greenhouse gases in accordance with the Paris agreement and limit the increase in global mean temperature to 1.5°C, more extreme weather is likely to occur. Cities are especially vulnerable to flooding of infrastructure and buildings and basements from run-off after heavy rainfall. Adaptation measures could reduce these damages but are costly. Thus, there is a need to document the correlation between heavy rainfall and damage cost to show the benefit to society in terms of avoided damage costs from adaptation. This is of interest of both insurance companies and the affected businesses, municipalities and households.

This thesis adds to the scare literature on the correlation between precipitation and the damage costs from urban flooding by analyzing a unique dataset of daily insurance claim payments covering rainfall-related damages in Oslo over a 13-year period from 2008 to 2020 from the Norwegian insurance companies and good quality meteorological data from a set of weather stations in Oslo during the same period.

Data on insurance claims from business and households and the associated damage costs from the VASK data base of Finance Norway was used. Efforts were made to isolate urban flooding related damages from other water-related damages by applying combinations of coding of the damages. A sensitivity analysis of a less strict coding practice was also performed, in order to validate the damage coding.

Different rainfall indicators have been used and the results show that both for daily and weekly data of the number of insurance claims and the resulting damage costs, precipitation (in mm) that day/week and the preceding week have a significant (at the 5 % level) leading to increased number of insurance claims and damage costs. Precipitation the preceding month and the second and third week before the incident were included in the models, but precipitation more than one week before the incident was not significant.

Using the number of extreme rain days (defined as more than 20 mm precipitation per day) instead of the amount of precipitation per week, had a significant effect on both the number of insurance claims and damage costs; both per day and per week.

As damage costs from urban flooding are highest in August followed by September and October, seasonal dummy variables were also tested; showing that Summer (June-

August) and in some cases Autumn (September-November) had a significant effect on damage costs.

In conclusion, there is a clear correlation between rainfall and insurance claims, both in terms of numbers and costs. The explanatory power of the best models is, however, only around 20 % (in terms of adjusted R-square). This indicates that other factors explain a large part of the variation in insurance claims and costs. Thus, future research should collect data and investigate the effect of local conditions with regards to green areas and implementation of technical and nature-based adaptation measures in order to better explain the variation in future damager costs and predict the impacts of future adaptation measures.

Preface

My motivation for this thesis is the climate crisis, and the need to also adapt to climate change to reduce the damage costs from increased mean global temperature and extreme weather events. Since the costs of adaption can be high, there is a need to document the benefits of such measures in terms of avoided damages to buildings and infrastructure. One important component would be the avoided damage costs in terms of insurance companies' compensation to businesses and households due to urban flooding from heavy rainfall.

I would like to thank my supervisor Assistant Professor Dr. Wei-Ting Yang for her availability for supervision, constructive comments and guidance during the master thesis process.

I would also like to thank Finance Norway for access to their VASK database and especially Chief Actuary for non-life insurance Kari Mørk for very useful comments and help in interpreting the insurance damage cost data.

Oslo, August 31st 2022

Table of Contents

1 Introduction ... 7

1.2 Background ... 7

1.2 Problem statement and research questions ... 8

2. Literature review ... 9

3. Data description ... 12

3.1 Study site ... 12

3.2 Data sources ... 12

3.2.1 Insurance claim data -VASK ... 13

3.2.2 Meteorological data - MET ... 17

3.2.3 Population data – Statistics Norway ... 18

4. Results and discussion ... 20

4.1 Descriptive statistics ... 20

4.2. Description of variables ... 27

4.3 Regression analyses ... 29

4.4 Multiple regression models for daily data ... 29

4.5 Multiple regression models for weekly data ... 36

5. Conclusion ... 40

6. References ... 42

List of abbreviations

BAU = Business-As-Usual

CPI = Consumer Price Index

FN = Finance Norway (Finans Norge)

IDF-curves = Intensity-Duration-Frequency curves

m.a.s.l. = meter above sea level

MET = Norwegian Meteorological Institute

SSB = Statistics Norway (Statistisk Sentralbyrå)

VASK= Water Damage Statistics (Vannskadestatistikk at Finance Norway

WMO = World Meteorological Organization

1 Introduction

1.2 Background

The Intergovernmental Panel on Climate Change (IPCC) in their 6th Assessment Report found that there is a greater than 50% likelihood that global warming will reach or exceed 1.5°C in the near-term, even for the very low greenhouse gas emissions scenarios (IPCC 2021). They state that "Human-induced climate change, including more frequent and intense extreme weather events, has caused widespread adverse impacts and related losses and damages to nature and people, beyond natural climate variability" (IPCC 2022, Summary for Policy Makers SPM B.1.) and predicts that the further increase in global mean temperature will cause higher frequencies and duration of extreme weather events like heat waves, windstorms and heavy rainfall.

Heavy rainfall in urban areas can cause both fluvial and pluvial flooding. While a fluvial flood occurs when water levels in rivers rise and overtop their banks, pluvial floods occur when surface water accumulating from the result of intense rainfall saturates the urban drainage system, and the excess water cannot be absorbed. This study will investigate the cost of damage from the latter type of flooding, denoted by the term "urban flooding" throughout this study.

Insurance claims can be used to document the damages and costs caused by urban flooding, and insurance companies can contribute to the reduction of these damage costs by sharing their data with the municipalities so they can design and implement strategies for technical and nature-based measures to mitigate damages as well as adapt. This is in the interest of both the insurance companies, the municipalities, affected households and businesses and society at large. For example, if the insurance companies find that the cause of the damage is that the municipality has not maintained the drainage system well, they will claim the compensation paid to their customers back from the municipality.

Finance Norway`s (FN) Climate report 2021 shows that the largest pluvial damage costs can be found in the large cities in Norway like Oslo and Stavanger (Finance Norway 2021; figures 7, 10 and 11). While there are studies looking at the correlation between the number of insurance claims and intense rainfall in cities in Norway (Rohrbek et al 2016, 2018) and) and other European cities (Cortès et al 2018), there are only a few studies (e.g. Torgersen et al 2015) looking at also the total damage costs. Further, the insurance data available is often limited and uncertain in terms of defining the damages as a rainfall-related damage, and Torgersen et al (2015) used a qualitative scale for damage costs rather than the amounts. Thus, the insurance claims data need to be prepared in such a way that only damages related to pluvial floods (urban flooding), and not all other causes, are included in the analysis.

This thesis contributes to the scarce literature on the correlation between different rainfall indicators and urban flooding damage costs by analyzing a unique dataset of daily insurance claim payments covering rainfall-related damages in Oslo over a 13-year period from 2008 to 2020 from the Norwegian insurance companies and good quality meteorological data from a set of weather stations in Oslo during the same period.

1.2 Problem statement and research questions

The main aim of this thesis is to investigate the potential correlations between precipitation and the number of insurance claims as well as the damage costs from urban flooding in Oslo for the period 2008 – 2020.

The following four research questions will be explored:

1) Is there a correlation between rainfall data and the damage costs for insurance claims for urban flooding in Oslo, typically flooding of basements in built-up areas?
2) Is there a correlation between rainfall data and the number of insurance claims for urban flooding in Oslo?
3) What type of rainfall data – daily, weekly or monthly rainfall or discrete data in terms of number of “extreme rainfall” days (defined as having daily precipitation above a certain level) - can best explain the observed number of insurance claims and damage costs from urban flooding?
4) Do other variables affect the observed number of insurance claims and damage costs from urban flooding? Both the average daily temperature (i.e., increased temperature would lead to more precipitation being rain instead of snow) and the size of the population (i.e., with higher population, a higher number of households could be affected) could affect the insurance claims and damage costs.

2. Literature review

According to Lamond and Pennig-Rowsell (2014) a high level of insurance coverage demands that there is a functioning and sustainable insurance scheme; which they define as requiring that there is: i) an insurable risk that is quantifiable, distributed, and affordable, ii) an insurable population aware of risk, willing to insure and which can afford the necessary premiums, and iii) a solvent insurer that is willing and can afford to run the scheme and pay claims; and has arrangements in place to cover any abnormally large losses. They further note that all these conditions seem not to be met in many countries, especially not in the developing part of the world where insurance coverage is very low (see Lamond and Pennig-Rowsell 2014, table 1).

Norway, however, has compulsory flood insurance as part of property insurance policies, backed by a national pool (Norwegian Nature Perils Pool[1]), which is backed by reinsurance. Thus, Norway has high coverage of flood insurance for properties. Rohrbeck et al (2018) looked at the correlation between weather events like rainfall or snow melt and the number of water-related property insurance claims, but not the cost of the claims, in three Norwegian cities, Oslo, Bærum and Bergen. Rohrbeck et al (2016) also used Norwegian data. They looked at the correlation between the number of issued insurance claims data and multiple daily weather metrics, such as the amount of precipitation, for Oslo and ten neighboring municipalities on both sides of the Inner Oslofjord. Their insurance data provided the daily number of insurance claims due to precipitation, surface water, snow melt, undermined drainage, sewage back-flow or blocked pipes at municipality level from 1997 to 2006.

While Rohrbeck et al (2016, 2018) used data for the *number* of insurance claims for *multiple* cities in Norway for 1997-2006; Torgersen et al (2015) looked at the correlation between both the number and costs of insurance claims, and heavy precipitation in the city of Fredrikstad for the period 2006-2012. Figure 2.1 shows that insurance claims costs peak in late summer and early autumn (July, August and September). In their multivariate analysis they do not use the exact amounts but characterize the insurance claim costs on a scale from expensive to no cost at all.

[1] See https://www.naturskade.no/en/the-norwegian-natural-perils-pool/background/ and https://www.naturskade.no/en/naturskader-og-erstatning/compensation-schemes/

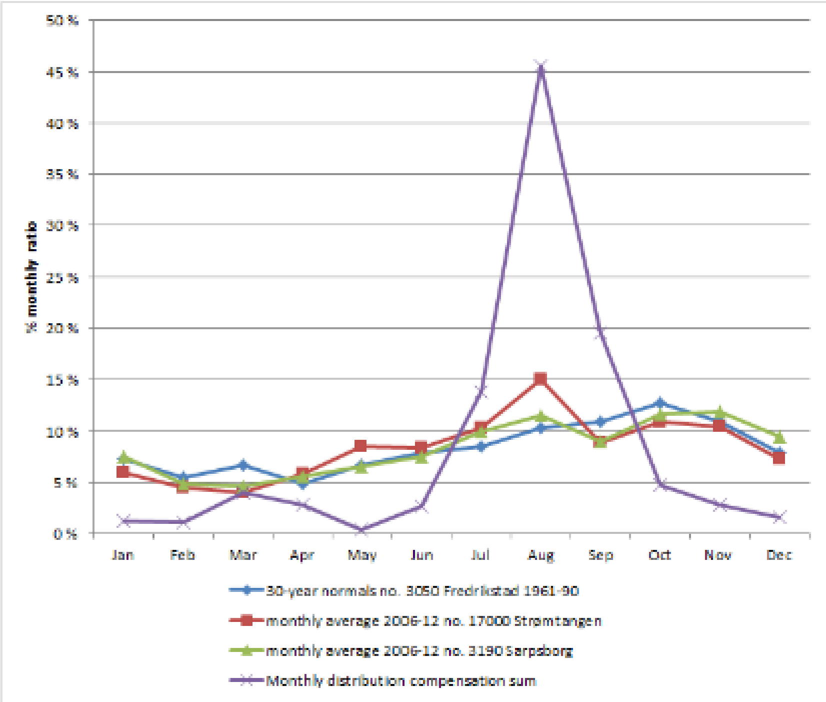


Figure 2.1. Monthly distribution of precipitation (30-year normal 1961-1990 (in blue) and monthly averages for two meteorological stations; Strømtangen (in red) and Sarpsborg (in green)) and insurance claims costs (in violet) for Fredrikstad during the period 2006-2012. <u>Source:</u> Torgersen et al (2015, figure 3)

Torgersen et al (2015) had access to insurance claims data for Fredrikstad for 2006-2012 from Finance Norway (FN) and found that for most days with high insurance claims amounts, the rain starts with relatively low intensity but that the intensity remains higher over longer time relative to the observations included for calculation of Intensity-Duration-Frequency curves (IDF-curves). An IDF-curve shows the probability that average rainfall intensity will occur in a specific region, where the calculated probability is based on statistical analysis of recorded rainfall data over a long period, typically 30 years. IDF-curves are available from the Norwegian Meteorological Institute (MET), and they locate patterns in the short-time duration rainfall and its impact of flooding and find that the measured progress of rain is characterized as either the long lasting / less intensive or short term/intensive rain. Very extreme short duration rainfall with little spatial distribution within a small area of Fredrikstad did not seem to be the main reason for the insurance claims for the period 2006 -2012. This is good news also for my analysis as the weather stations are not that widespread as to cover very local extreme precipitation, but at least the coverage seems to be better in the larger cities (which are at risk to urban flooding) than in the rural areas (which could be more subject to pluvial flooding[2] than urban flooding). Torgersen et al (2015) conclude that little precipitation the week before is a plausible explanation for why some days with heavy rain result in no claims. Although "sealed" areas (e.g., parking lots, streets, conventional roofs etc; as opposed to blue-green structures) dominate in the urban environment, they found a reduced risk of flooding when the ground was dry and unsaturated. Thus, the amount of precipitation over periods prior to extreme rainfall events should be considered as variables in addition to the rainfall the day of the extreme event to account for water runoff the extreme event caused. This thesis therefore include not only precipitation per day, but also for the preceding weeks.

Cortès et al (2018) note in their study of the correlation between fluvial flood-related insurance damage claims in Catalonia in Northeast Spain and heavy precipitation, that this kind of analysis had not been possible in the Mediterranean countries up until then because of limited quantitative data on flood related damages. However, with access to insurance claims data for Catalonia, and the fact that flood insurance is bundled into all building insurance in Spain with 100 % coverage (Lamond and Pennig-Rowsell 2014, table 1), they developed a model that can simulate the probability of a damaging event as a function of precipitation.

Cortès et al (2018) conclude that these correlations can be used to predict fluvial flood damage in future climate change scenarios as e.g., Wobus et al. (2014) do when they estimate the damage costs from flooding in the USA under a Business-As-Usual (BAU) climate change scenario. Cortès et al (2018) further note that complex relationships between climate variability, human activities and flood damage may limit the applicability of these findings to conditions that are very different from the current ones; and that more complex analyses including other factors such as soil physical characteristics (e.g., slope, soil characteristics, vegetation) and construction of fluvial defense structures could provide better predictions. This type of issues for fluvial floods can certainly also impact the correlations between heavy precipitation and pluvial floods in urban areas (i.e.; urban flooding) in Norway, as noted by Torgersen (2015), but demand very location-specific data for urban areas which are usually not available, and not for my case study area Oslo.

This literature review shows that there are few studies that have looked at the correlations between insurance claim costs and extreme weather events like flooding,

[2] Fluvial flooding refers to flooding caused by rivers overflowing

including urban flooding. This could be because it is difficult to get access to claims data from the insurance companies (both the number of claims and especially the damage costs), and/or that the insurance coverage is low for this type of extreme weather events in many countries.

3. Data description

3.1 Study site

The main reason for choosing the municipality of Oslo as the study area is that the capital of Norway is at risk for pluvial flooding, and the potential for damages is large as it is densely populated and a built-up area with a population of about 700,000 persons, is home to many businesses and regional and national government buildings. As the insurance claims statistics reflect damages only to households and businesses, this is what is analyzed here. Further, Oslo has a good network of weather stations with long time series of good quality data. Oslo is also located in an area where climate change scenarios show an increase in annual rainfall in the future (Hanssen-Bauer et al 2017).

This thesis considers data for the period from 2008 to 2020. Meteorological data go way back, but the insurance claims database VASK ended their trial period in 2007 and only from 2008 all the large insurance companies started to report to FN. I have not used data from 2021 and 2022 as the insurance claim costs are estimates provided by the insurance companies that later are confirmed or corrected when the actual damage costs become available and the reason for the damage is confirmed. In the latter case some claims could after some time be moved from the VASK database to FN's natural hazard database NASK[3]; which means the incident would no longer be classified as an urban flooding event. Thus, as data from 2021 and 2022 are still subject to changes, I use only the confirmed events and damage costs,

3.2 Data sources

The data used in this study consists of three different datasets. Weather data (precipitation and temperature) from the Norwegian Meteorological Institute (MET), insurance claims data from Finance Norway (FN) and population data from Statistics Norway (SSB). The data sources with links to the databases are listed in table 3.1.

[3] NASK : https://www.finansnorge.no/statistikk/skadeforsikring/naturskadestatistikk-nask/

Table 3.1. Data sources for precipitation, insurance claims/damage costs and population in Oslo for period 2008-2020

–Data	Source	Website
Precipitation, registered at meteorological stations in Oslo	MET	https://seklima.met.no/observations/ https://www.met.no/vaer-og-klima/kvaliteten-pa-observasjonene
Temperature	MET	https://seklima.met.no/observations/ https://www.met.no/vaer-og-klima/kvaliteten-pa-observasjonene
Insurance claims and Damage costs	FN - VASK	https://www.finansnorge.no/statistikk/skadeforsikring/vask/
Population in Oslo	SSB	https://www.ssb.no/statbank/table/07459/tableViewLayout1/

3.2.1 Insurance claim data -VASK

The insurance data were obtained from water damage statistics (VASK) provided by Finans Norge (FN). The basis for the data is insurance claims from households and businesses submitted to the insurance companies, which then report these to FN. From January 1st 2008 onwards all the largest insurance companies in Norway 2008 have delivered data to the VASK database. The data delivered covers 85 % of the Norwegian market. VASK uses weights to compensate for the underreporting. There is also a correction for slightly different routines in each insurance company and for damages that have occurred but are not yet reported. Thus, the incidents and damage costs in VASK should be able to give a correct picture of all water damages that have occurred in Norway since 2008. However, as the estimated damage costs based on the submitted claims might be corrected more than one year after they are reported and registered in VASK, data only until 2020 is used. A replacement cost approach is used assess damage costs.

The initiative to establish a common water damage statistics database came from the committee for actuarial damage (FAUS) in January 2005, and the month after the industry Board for Risk and Damage (BRS) decided to create the water damage statistics VASK database. The code system for the cause and type of damage was developed by a working group with representatives from the largest insurance claims companies. The purpose of the insurance claim database VASK is to explain and follow the development of water related damages over time, in order to provide a good foundation for damage prevention measures and dissemination of information and contribute to more water damage-proof products and installations. Further, FN and the insurance industry would like to use the database to support their arguments for changes in spatial planning and the Planning and Building Act, building detail sheets, guidelines etc.

In Norway, according to the VASK database, there are about 30 000 water related damage incidents annually with compensation payments of NOK 1,5 billion per year, but there are large fluctuations in water related damages from year to year in the period 2008-2020. 58 % (17400) of these incidents are urban flooding, creating damage costs of about and 700 million NOK annually on average.

In VASK only water damages that are sufficiently well coded are included. This might lead to the numerical material being somewhat smaller than FN's quarterly publication for claim statistics for land-based insurance.

With regards to coding principles in VASK, all insurance companies that report to VASK are provided with guidance on how to report and code each insurance claim. All water related insurance damages are to be coded at three levels. These are:

i) <u>Installation</u> - System that uses/transports water. Here the location of the damage is stated. Here it is stated whether the damage has occurred: outside or inside the building, whether the damage has occurred in a water supply pipe or in a drainage pipe, whether the part of the pipe where the damage has occurred is open or hidden, etc.

ii) <u>Source</u> – Where the leaking water comes from; rain, snow, melt water etc. Provides a more detailed description of the location of the damage or of the damage itself. Pipe material is specified here, whether it occurred in a metal pipe or plastic pipe, whether it occurred in a water connected machine such as a washing machine, or whether it occurred in a connection with a leaky wet room

iii) <u>Cause</u> - The actual cause of the damage. Here it is stated whether an assembly error has been made, whether the user has made a mistake, whether the damage is due to frost or kickback, etc.

For each water related damage installation have 11 different categories, while source and cause have 10 categories each. This leads to almost 1100 different code combinations, with some combination being invalid. There are no "uknown" values. Table 3.2. provides an overview over the codes that were selected and defined as damages most likely connected to urban flooding. Using these VASK codes for Oslo for the period 2008-2020, insurance claims and damage costs were reported on 3 017 of these 4 749 days. In total for these 3 017 days there were 10 439 insurance claims from households and businesses.

Table 3.2. Overview of type of damages and the codes in VASK most relevant to urban flooding

Installation		**Source**		**Cause**	
Code	Description	Code	Description	Code	Description
G	Outdoor- water- and sewer system	I	Precipitation/snow Melt/ground water	E	Old age
H	Water penetration from outside through foundation			G	Stop in sewer / Sewer back up
I	Water penetration from outside above foundation			I	Influence from outside
				J	Drainage system

Table 3.3. provides a detailed description of the codes used for the installation affected, source and cause of the damage in order to define the claim as an urban flooding incident.

Table 3.3. Detailed description of codes in VASK used to define urban flooding incidents**.**

Installation

Code	Description
G	Outdoor water and drainage systems - Intake pipe for water and outlet for drainage (plug pipe) outside the building including basins.
H	Water penetration from the outside through the ground - Water penetration through the parts of the building that are below the ground, i.e., through the foundation wall below ground level. Also includes infiltration from drainage pipes for surface water, drainage pipes from roof drains and drainage pipes.
I	Water penetration from outside above ground – Water penetration through the parts of the building that have an exterior facing the open air above the terrain, i.e., through the roof and exterior walls. Also includes infiltration from gutters, external and internal roof drainers.

Source

Code	Description
I	Precipitation. Meltwater. Groundwater – Penetration of precipitation, meltwater from snow and ice, floodwater / surface water in the terrain and groundwater.

Cause

Code	Description
E	Old age (normally >30 years) of installation – wear and tear due to age or salty or acidic water. Blackened rubber bearings, broken gaskets, leaking valves, corrosion etc. Entire or significant parts of the installation are worn out[4].
G	Stop in sewer. Sewer back up – Flooding in the building caused by water entering through the drainage network or due to closed drain, water trap or sanitary equipment.
I	Influence from outside – Sudden increase in pressure, nails or screws in water pipes, mechanical rubbing, exhaustion due to vibration or impact, the sprinkler head being run over by a truck etc., pipe breakage in plug pipes due to digging etc.
J	Drainage – Clogged drainage, extraordinary rainfall or snowmelt or larger amounts of surface water than the drainage is designed for.

[4] For more details on the term «ageing»; see "Levetid for sanitær-installasjoner i boliger" from Byggforsk ref no. 700.330. https://www.byggforsk.no/dokument/3112/levetider_for_sanitaerinstallasjoner_i_boliger

3.2.2 Meteorological data - MET

The meteorological data were obtained from The Norwegian Meteorological Institute (MET) and their site Norwegian Climate Service Center (Norsk Klimaservicesenter). The service has metrological data from all over Norway. MET gathers data both from the meteorological stations that they themselves own and stations owned by the municipalities. For the stations owned by municipalities, MET provide guidance to make sure the observations meet their required quality standards. Because the data is gathered from so many stations, MET has undertaken multiple measures to make sure their data have the best quality possible. The stations need to have a good location, the equipment need to have sufficient accuracy and they must be maintained and controlled properly. MET also controls the data in multiple steps when they receive it, both automatically and manually[5]

MET rates their stations and each element they observe on two different categories: i) exposure class and ii) performance class. Exposure class indicates the quality of the station's location for measuring the weather element on a scale from 1 to 4, where 1 meets all the World Meteorological Organization's (WMO) requirements for the location and 4 is a bad location. Performance class indicates the quality of the sensor in terms of measurement accuracy, calibration, and maintenance; on a scale from A to E, where *A* means the sensor meets all the requirements from WMO and *E* is unknown performance and is considered less valuable.

In Oslo there are 38 weather stations that gather data; see figure 3.1. In this study, data is gathered from eight of these. These are (altitude in m.a.s.l. in parenthesis) : Oslo – Blindern (94), Oslo – Disen (136), Oslo – Haugenstua (123), Oslo – Hovin (100), Oslo – Lambertseter (135), Oslo – Lilleaker (23), Oslo – Ljabruveien (92) and Tryvannshøgda (514). These eight weather stations were selected because the municipality of Oslo covers a large area (454 km^2), and it is important to select stations that can represent the entire area of the city. Precipitation is often local, and with the eight selected weather stations most of Oslo is covered. Weather station in locations that are densely populated were selected, as there will be a higher probability of an insurance claim occurring there; all other things being equal. As weather stations can be out of the operations for some periods of time, stations with continuous recordings for the period 2008-2020 and as few missing values as possible, were selected.

[5] See also their website: https://www.met.no/vaer-og-klima/kvaliteten-pa-observasjonene .

The quality of the weather stations used in this study are *at least* in category “2 C” (i.e. exposure class 2 and performance class C. This seems to be in the upper half with regards to quality of the weather stations. Exposure class 2 means that the quality of the location is unknown but assumed to be good. Performance class C means the sensor type is assumed to meet the requirements of WMO but there is a lack of control measurement, routines for calibration, or maintenance.

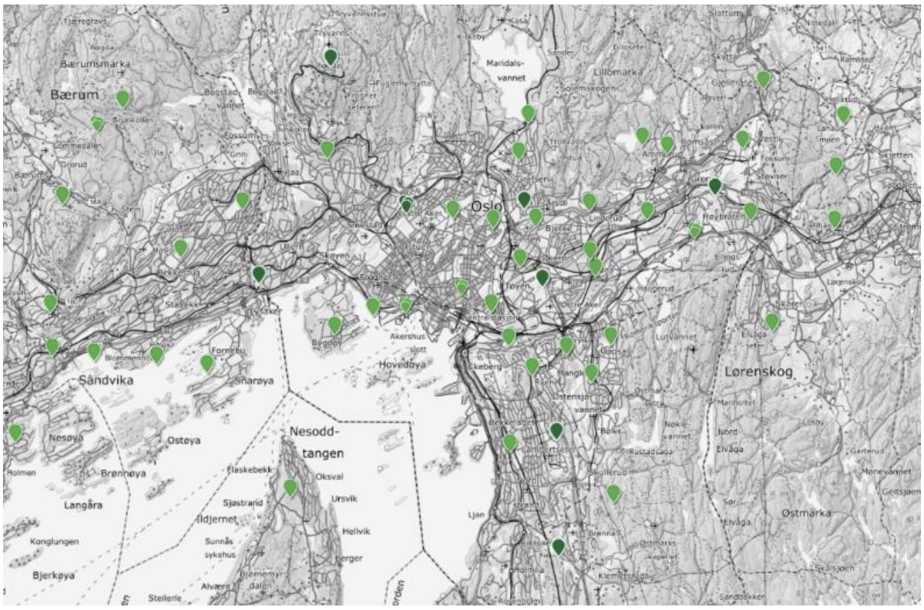


Figure 3.1. Weather stations in Oslo municipality shown in green. The dark green ones have been selected for data collection for this study. These eight weather stations were selected based on their consistent recordings since 2008, the high quality of their equipment to gather precise data and their geographical spread in order to get a representative picture of the rainfall and temperature in the municipality of Oslo. Source: https://seklima.met.no/stations/

3.2.3 Population data – Statistics Norway

The population data were obtained from Statistics Norway (SSB, 2022). SSB is the main authority for the preparation and dissemination of official statistics in Norway and produces approximately 85% of the official Norwegian statistics (https://www.ssb.no/). SSB do admit some errors in their population data under the collection and processing of the data but regards the error to be insignificant. There will be some dropout errors due to persons staying in Norway while not being registered as residents. SSB mainly uses the Central Population Register as their statistical base for population. Some of the disadvantage with the Central Population Register is that too many are registered as residents; but also that certain groups, such as unmarried students are registered as residents with their parent(s). SSB revealed in a survey in 1990 that the registered residence was incorrect for 5.5 percent of the population. This is relevant for this study as there are many students in the municipality of Oslo.

There are also people that move without reporting their departure, and immigrants that leave without reporting their move to the population register. Conversely, there are several people arriving in Norway that live here illegally in Norway and most of them live in Oslo. Since the two factors pull in opposite directions, SSB assumes that these effects largely cancel each other out (SSB, 2022)

4. Results and discussion

4.1 Descriptive statistics

Tables 4.1 and 4.2 report descriptive statistics of the dataset for the period 2008-2020 for daily data (N=4749 days) and weekly data (N= 678 weeks), respectively. In 3017 days of the 4749 days (i.e., 63,53 %) insurance claims due to urban flooding were registered. During these 3017 days 10439 claims were recorded, i.e., on average 3.46 claims per day with claims recorded. Table 4.1 shows that counting also the days with zero claims, the average number of daily claims is 2.14; varying from zero to 362 claims in one day. The average damage costs per day is 135,922 2021-NOK; varying from zero to more than 75 million 2021-NOK[6]. This clearly illustrates the large variation in both number of claims and damage costs per day related to urban flooding in Oslo over this 13-year period.

Table 4.1. Descriptive statistics of daily data for 2008-2020 (N= 4749 days)

Variables	Mean	Std	Min	Max
Precipitation day (mm)	4.91	8.42	0	88.60
Precipitation week (mm)	34.26	32.24	0	207.40
Precipitation month (mm)	145.93	87.69	7.20	499.00
Extreme rainfall	0.07	0.25	0	1
Population (no. of people)	631876	41071	548617	693494
Temperature days above 0	0.76	0.43	0	1
Damage cost (2021 - NOK)	135922	1349084	0	75223420
Claims (no. of claims per day)	2.14	8.77	0	362

Table 4.2. Descriptive statistics of the weekly data for 2008-2020 (N= 678 weeks)

Variables	Mean	Std	Min	Max
Precipitation week t (mm)	34.11	32.22	0	162.80
Precipitation week t-1 (mm)	34.15	32.30	0	162.80
Precipitation week t-2 (mm)	34.05	32.10	0	162.80
Precipitation week t-3 (mm)	33.94	32.01	0	162.80
Extreme rainfall days week t	0.46	0.81	0	4
Population (no. of people)	632 214	40 918	560 484	693 494
Temperature – no. of days above 0 in week t	5.33	2.58	0	7
Damage cost week (2021- NOK)	948 245	4 335 045	0	92 768 068
Claims week	14.91	31.50	0	467

[6] All insurance claim costs (damage costs) for the period 2008-2020 have been converted to 2021-NOK using the Consumer Price Index (CPI); see SSB https://www.ssb.no/priser-og-prisindekser/konsumpriser/statistikk/konsumprisindeksen

Figures 4.1 shows the percentage distribution of damage costs over the months of the year as well as precipitation and the number of days with extreme precipitation (defined here as 20 mm per day as an example); on average for all 13 years for the period 2008-2020. Figure 4.2 shows the same as figure 4.1 but with number of insurance claims instead of the damage costs (note that the y-axis is somewhat different in the two figures). These figures clearly show that the month of August has the highest percentage of both damage costs and the number of insurance claims, followed by the months of September and October; and that this pattern follows the same patterns as monthly precipitation and monthly number of extreme precipitation days (defined here as 20 mm per day).

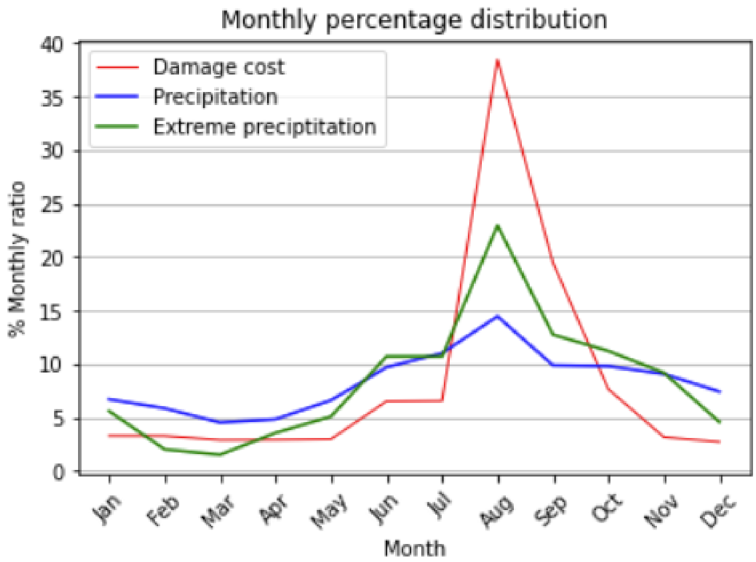


Figure 4.1. Distribution of damage costs (i.e., insurance claim amounts in 2021-NOK), precipitation and number of days with extreme precipitation (defined here as 20 mm per day) for each month of the year; average for the period 2008-2020.

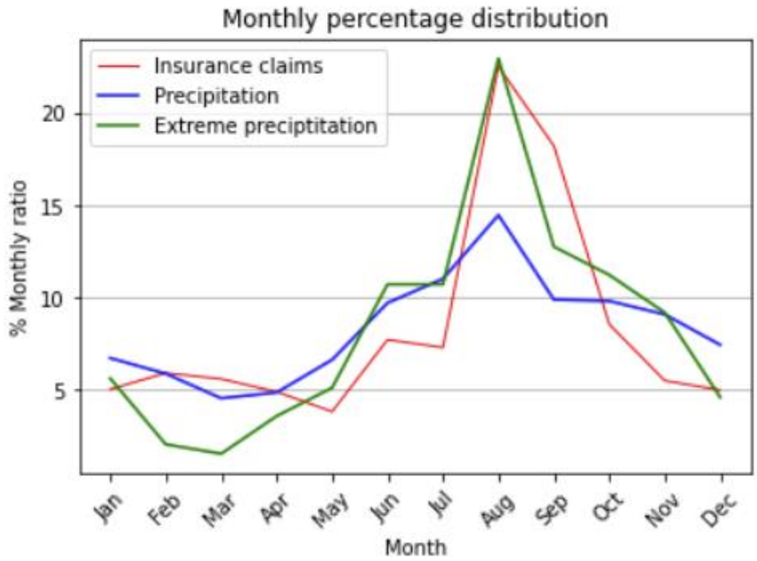


Figure 4.2. Distribution of number of insurance claims, precipitation and number of days with extreme precipitation (defined here as 20 mm per day as an example, but other definitions also tested) for each month of the year; average for the period 2008-2020

While figures 4.1 and 4.2 show percentage distribution over each of the twelve months averaged over the 13-year period, figures 4.3 - 4.6 show the damage costs per month (in 10 million NOK), monthly precipitation, number of claims per month and number of extreme rainfall days (> 20 mm) per month, respectively; for *each* of the years 2008 - 2020: Note than in figure 4.3, 2016 is not included as in August 6th 2016 there was very high daily rainfall of 39.1 mm[7], causing very high damage costs of 75 223 420 NOK from 362 insurance claims just for this day. This extreme observation makes the total damage costs for August 2016 very high compared to other years and months and the figure difficult to read. This figure, including the data for 2016, can be found in table A-3 in the appendix.

Comparing figures 4.3 and 4.4 we see that there the same common pattern as in figures 4.1 and 4.2 with both damage costs per month and monthly rainfall peaking in the autumn months of August, September and October; and a smaller peak in June. The relative annual variation over the months seems to be higher for monthly rainfall than monthly damage costs. Further, there seems to no clear increasing trend in neither rainfall nor damage costs over time, but then this is a relatively short period of time which easily can

[7] See https://www.aftenposten.no/oslo/i/V5Rvr/enorme-nedboersmengder-over-oslo-omraadet

be affected by natural variation in the weather and is too short a period to see a climate-induced increase in rainfall and damage costs.

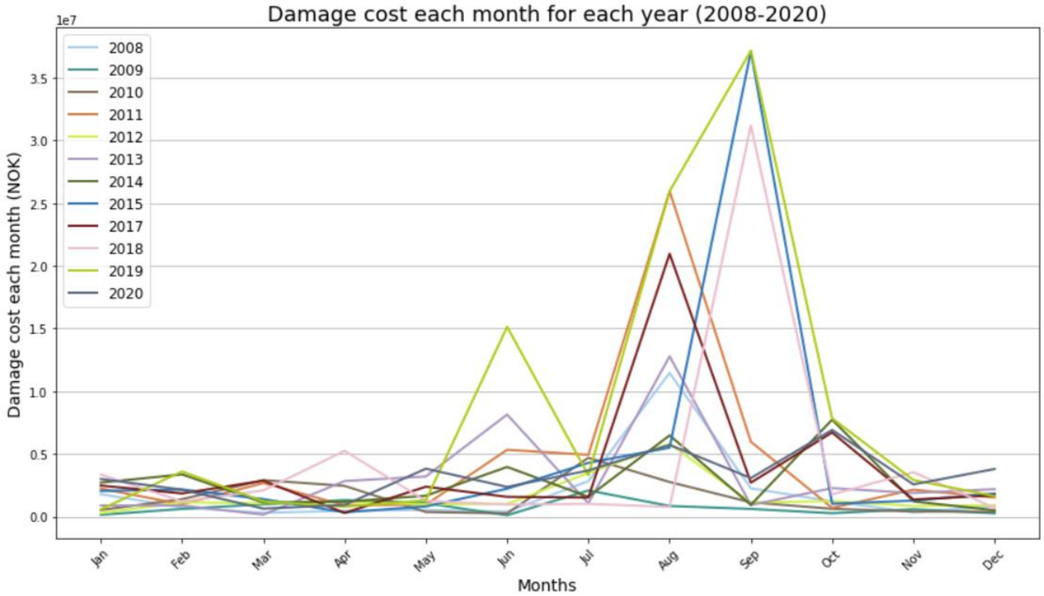


Figure 4.3. Monthly damage costs (in 10 million 2021-NOK) for each year for the period 2008 -2020; except 2016 (due to an extreme event on August 6th 2016). Figure A-3 in the Appendix includes data also for 2016.

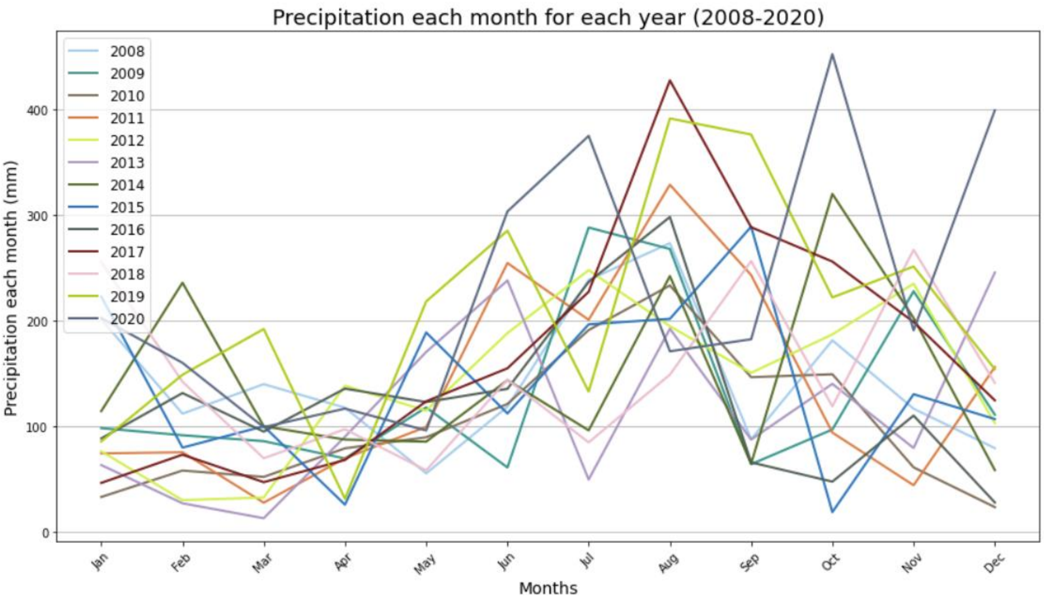


Figure 4.4. Precipitation each month (in mm) each year for the period 2008 -2020.

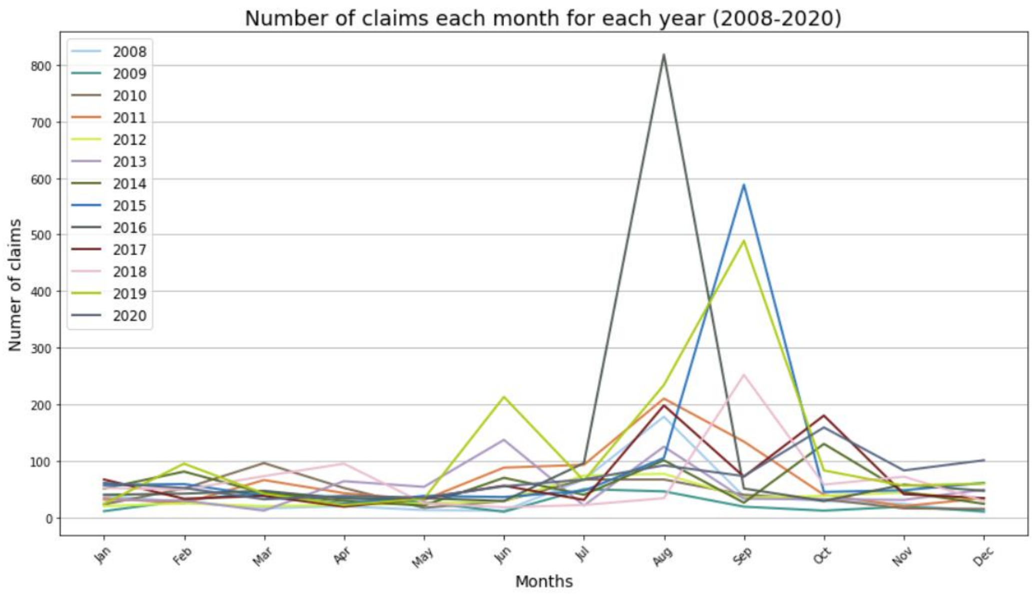


Figure 4.5. Number of insurance claims each month each year for the period 2008 - 2020.

Figure 4.5 shows that the number of insurance claims each month for each year seems to have less variation between years than the damage costs in figure 4.3 (no need to exclude 2016 for illustrating this) and shows the same pattern as in figure 4.,3 but even clearer.

Figure 4.6 shows extreme rainfall days each month for each year and there is a large variation from year to year and between months both in summer and autumn (and also winter) but in many years August, September October still seems to stand out as having the highest number of extreme rainfall days.

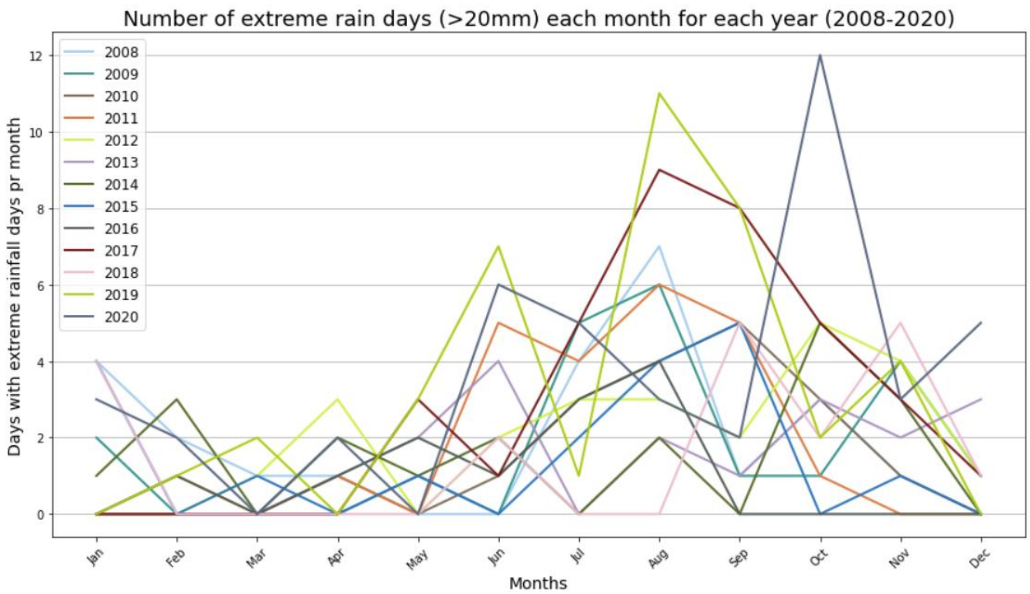


Figure 4.6. Number of extreme rainfall days (> 20 mm) each month each year for the period 2008 -2020.

As there seems to be no generally agreed definition of an extreme rainfall day, two definitions in addition to the “above 20 mm” definition used for Oslo in this study, are illustrated in figure 4.7. While the number of days using the strictest definition (i.e., above

40 mm daily precipitation), seems to be quite constant over time; the two other definitions (above 20 and 25 mm) both indicate an increasing trend the last four years; from 2016 to 2020.

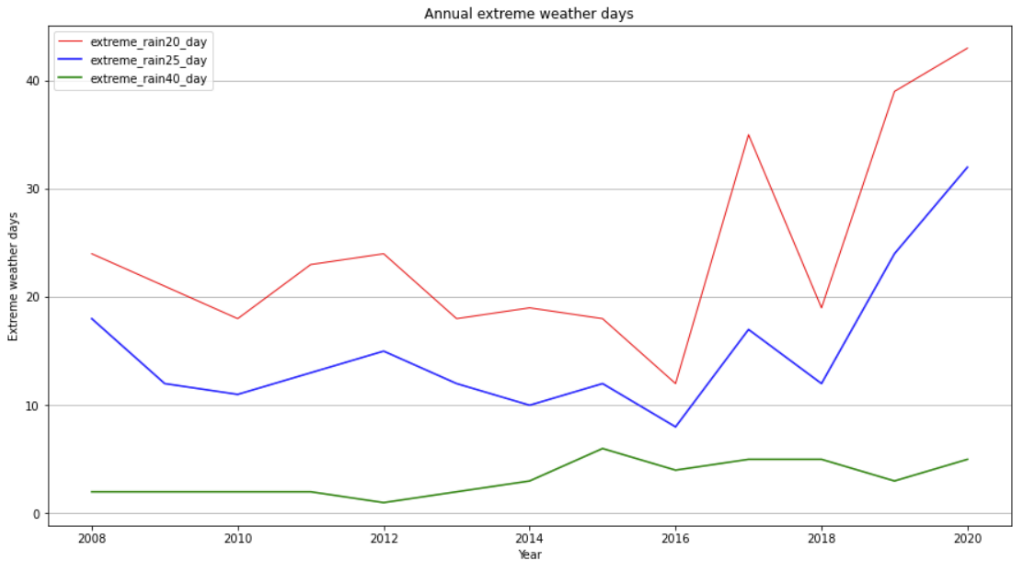


Figure 4.7 Annual number of extreme rainfall days for the period 2008 -2020 at three different definitions; above 20, 25 and 40 mm per day.

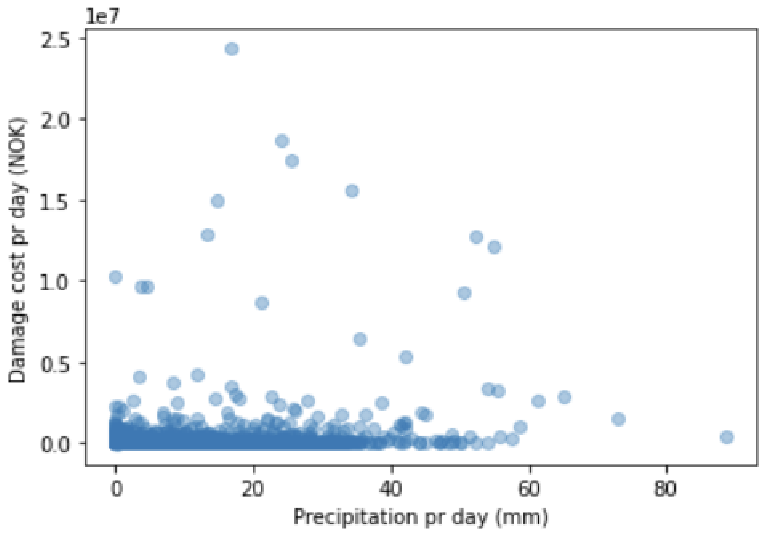


Figure 4.8. Plot of daily data on damage costs (in 10 million NOK/day) and precipitation (mm/day) in Oslo 2008-2020**.** All observations except the extreme rainfall day August 6th 2016 (N= 4748 days)

To further illustrate the data, figure 4.8 and 4.9 show damage costs plotted against precipitation for daily data (N=4748) and weekly data (N= 677); respectively. Again, the extreme observation of August 6th 2016 (and week of this event; week 31 in 2016) is left out in order to avoid compressing the other observations and illustrate better the distribution. Figures A-4 and A-5 in the appendix shows the plots with all observations (including this extreme rainfall day and corresponding week)

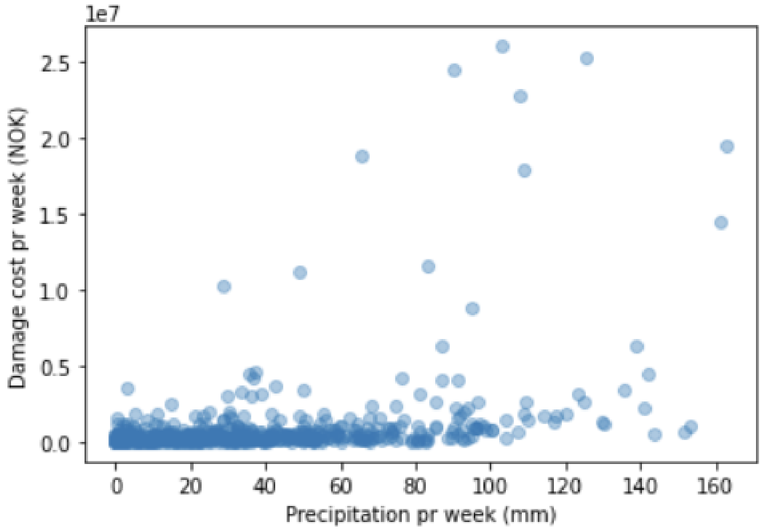


Figure 4.9. Plot of weekly data on damage costs (in 10 million NOK/week) and precipitation (mm/week) in Oslo 2008-2020**.** All observations except the week of the extreme rainfall day August 6th 2016; week 31 (N= 678 weeks)

Finally in this descriptive statistics section, figure 4.10 shows the development of annual damage costs and annual precipitation in Oslo for each year from 2008 to 2020. Both curves seem relatively flat until 2016 (with the exception of the hike in damage costs in 2016). Then the annual precipitation seems to increase (especially after 2018) while the annual damage costs show much more variation and even decline in 2020 compared to 2019. It is premature to conclude from this that climate adaption and damage mitigation measure that the Oslo municipality, businesses and households have implemented in recent years; and longer time series and analyses are needed to test whether this recent trend will continue.

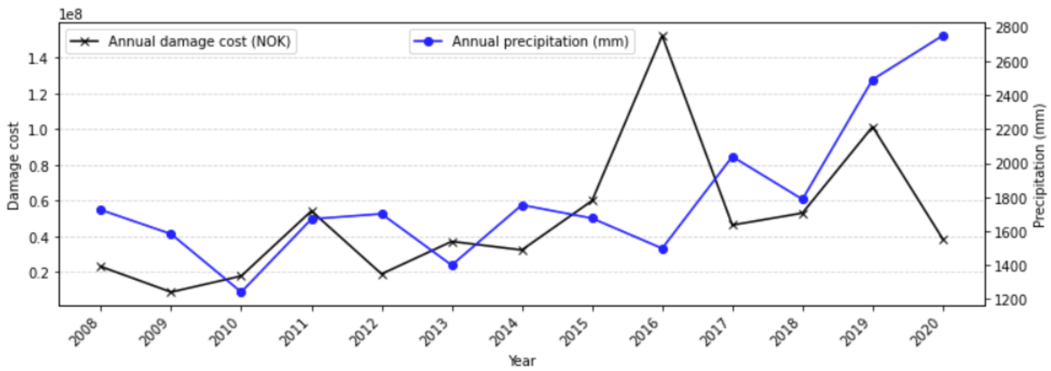


Figure 4.10. Annual damage costs (in 100 million 2021-NOK; in black) and annual precipitation (in mm; in blue) in Oslo for each of the years 2008 -2020.

4.2. Description of variables

Table 4.3 lists and describes the variables used in the regression analyses and their unit of measurement.

Table 4.3. Description of variables

Variables	Unit	Description
Precipitation day	mm	The maximum precipitation recorded at any of the eight weather station a given day.
Precipitation week	mm	Total precipitation the seven days preceding the given day, using the daily precipitation from the weather station (out of the eight) that had the highest observation.
Precipitation month	mm	Total precipitation the 30 days preceding the given day, using the daily precipitation from the weather station (out of the eight) that had the highest observation.
Population	No. of people	No. of people in Oslo Municipality. Annual data, at the end of the year.
Damage cost	2021-NOK per day	Insurance claims cost per day. Estimated cost related to the insurance claim. Includes appraiser, and compensations to the insurance customer (craftsman, new parts etc.).
Claims	No. of claims	Number of insurance claims per day
Damage cost week	2021-NOK per week	Insurance claim costs per week. Estimated cost related to the insurance claim. Includes appraiser, and compensations to the insurance customer (craftsman, new parts etc.).
Claims week	No. of claims	Number of insurance claims per week
Precipitation week t	mm	Total precipitation in a week, the given week. The maximum precipitation recorded at any of the eight weather stations a given day; aggregated for all days of the given week.

Precipitation week t-1	mm	Total precipitation in a week; one week prior to the given week. The maximum precipitation recorded at any of the eight weather stations a given day; aggregated for all days of the week prior to the given week.
Precipitation week t-2	mm	Total precipitation in a week, two weeks prior to the given week. The maximum precipitation recorded at any of the eight weather stations a given day; aggregated for all days of the week, two weeks prior to the given week.
Precipitation week t-3	mm	Total precipitation in a week, three weeks prior to the given week. The maximum precipitation recorded at any of the eight weather stations a given day; aggregated for all days of the week, three weeks prior to the given week.
Spring*	0-1	Dummy variable, defined as 1 if the month is March, April, or May and 0 otherwise
Summer*	0-1	Dummy variable, defined as 1 if the month is June, July, or August and 0 otherwise
Autumn*	0-1	Dummy variable, defined as 1 if the month is September, October, or November and 0 otherwise
Extreme rainfall days week t		Number of days in a week with extreme rainfall; defined here as more than 20 mm in a 24h period (00-24).
Population	No. of people	Population in Oslo, annual data
Temperature	No. of days above 0 in week t	Number of days in the given week with a temperature arithmetic mean for a 24h period 00-24 being above 0.

Note: * Seasons were defined according to meteorological seasons; see Season Definition: When Do They Start? (timeanddate.com) (Winter, defined as December, January February, is the hidden/reference category)

The correlation matrices for all variables used in the regressions are reproduced in table A-1 and A-2 in the Appendix (Chapter 7). Highly correlated independent variables are not included in the same regression, and only independent variables having r = 0.4 or less were included in the same regression model.

4.3 Regression analyses

Ordinary Least Square (OLS) regressions models are run in the form of:

$$Y_i = \beta_0 + \beta_1 D_{ii} + \beta_2 W_i + \beta_3 M_i + e_i \qquad (4.1)$$

where Y_i is the total insurance claim amounts (referred to as damage costs; in NOK) for day *i*, D_i is the rainfall (in mm) for day *i*, W_i is the rainfall the preceding seven days (week), M_i is the rainfall the preceding 30 days (month), and e_i is the error term. Regression models are also run for Y_i being the number of insurance claims for day *i*. to check whether the three precipitation variables D, W and M are better in predicting the *number* of claims than the total claims *costs*. As temperature can also play a role (in terms of precipitation in the form of snow causing less flooding) regression models with arithmetic mean daily temperature T_i (defined as a dummy with above zero temperatures defined as 1 and 0 otherwise) are also run. Further, as increase in the population of Oslo over these 13 years could also affect the number of insurance claims and thus the damage costs, the annual population of Oslo has also been included as a variable in the model.

The results of all the models using daily data on precipitation and insurance claims and damage costs are presented in chapter 4.4. In chapter 4.5. regression models are run for Y_i being weekly damage costs; and for Y_i being the number of insurance claims in week *i*, A significance level of 5 % will be applied in the interpretation of the regression models, but 1% and 10% significance levels are also shown in the tables.

4.4 Multiple regression models for daily data

Table 4.3B. shows that daily precipitation has a significant effect on damage costs and so does rainfall the week before the day in model 2, but not rainfall the month before the day as model 3 shows. As expected, and in accordance with previous studies, damage costs increase with the amount of rainfall, both per day and week. Out of the three models, model 2 seems to perform the best. The predictive power in terms of adjusted R-squared is, however, low in all models; clearly indicating that there are other factors affecting the daily damage costs.

Table 4.3B. Damage costs (defined according to the codes in table 3.2) per day and impact of daily, weekly, and monthly precipitation for 2008 - 2020 (N =4749 days)

	Model 1	Model 2	Model 3
Intercept	55.46	- 68 660.00**	- 89 600.00**
	(22 300.00)	(28 800.00)	(37 700.00)
Precipitation day	27 680.00***	25 370.00***	25 220***
	(2 291.30)	(2 367.95)	(2 374.13)
Precipitation week	-	2 335.84***	1 932.49**
		(618.21)	(755.31)
Precipitation month	-	-	243.15
			(282.04)
Adj. R-squared	0.030	0.032	0.032

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

The number of insurance claims and damage costs vary over the year with the highest frequency during late summer and autumn. Thus, models with variables to account for this seasonal fluctuation should also be run. This can be done by converting the year into four 3-months periods and construct three dummy variables for the four seasons (Winter, spring, Summer and Autumn). As the meteorological seasons means that the months July-October, when the highest number of insurance claims occur, overlap with two of these seasons, separate models for a subset of the data has also been run; see table A-2 in the Appendix. Comparing table A-2 with models in the full dataset in table 4.3. we find that the same precipitation variables are significant and positive; i.e. daily precipitation and precipitation the preceding week, and Adjusted R-square is only slightly higher; increasing from 3 to about 5 %. Thus, we continue using the full dataset.

Adding annual data on the population in Oslo as a variable to model 3 in table 4.3., shows that population was not significant at the 5% level, the adj. R-squared remained unchanged at 0.032, and only daily precipitation was significant of the three precipitation variables.

In table 4.4 the same models are run as in table 4.3, but with a less strict definition of rainfall-related damage; i.e. including all damages with Source = I but for all codes for Installations and Causes; and not only the codes listed in table 3.2) resulting in an increase from 3017 to 3270 days with reported insurance claims and an increase in the total number of claims from 10 439 to 12 189. The same variables as in table 4 are significant, and with the same correct, expected sign, showing that when the precipitation per day and the preceding week increase, the damage costs per day increase.

Table 4.4. Damage costs per day (wider definition; i.e, all damages coded as Source= I; but for all codes of Installation and Cause and not only the ones listed in table 3.2.) and impact of daily, weekly, and monthly precipitation for 2008 - 2020 (N =4749 days)

	Model 5	Model 6	Model 7
Intercept	9 361.39	- 73 180**	- 102 700**
	(24 300)	(31 300)	(41 000)
Precipitation day	31 310***	28 540***	28 330***
	(2 494.11)	(2 576.70)	(2 583.29)
Precipitation week	-	2 805.89***	2 238.16***
		(672.71)	(843.61)
Precipitation month	-	-	342.25
			(306.88)
Adj. R-squared	0.032	0.035	0.035

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

Precipitation the preceding month is still not significant in table 4.4. Thus Model 6 (equivalent to model 2 in table 4.3) still seems to perform the best, and Adjusted R-squared have increased just marginally from 0.032 to 0.035. Thus, this test of the sensitivity of the models to the definition of rainfall-related damages shows that the definition of claims /damage costs in table 3.2 performs well. Thus, in all other model runs, urban flooding insurance claims /damage costs will be defined in accordance with the VASK codes provided in table 3.2

Table 4.5 reports the same models as in table 4.3 but separate for insurance claims from households and businesses. Out of the 4749 days in these 13 years, 2 588 days had claims from households (with damage costs) and a total of 8 046 claims. For businesses the respective numbers were 1339 days and 2277 claims. Again, daily and weekly rainfall are significant in all models (at the 5 % level), while monthly precipitation is not; and adjusted R-square is again low but slightly higher for households. Thus, there seems to be no differences between households and businesses in terms of the significant effect of rainfall variables on daily damage costs.

Table 4.5. Damage cost and impact of daily, weekly, and monthly precipitation for households (Models 1-3) and business (Models 4-6) (N =4749)

	Households			Businesses		
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-11 360	70 420***	-90 950***	11 420	1 801.03**	1 367**
	(20 400)	(26 300)	(34 400)	(6 759)	(8 720)	(11 400)
Precip. day	22 840***	20 860***	20 710***	4 834***	4 512***	4 508***
	(2 094)	(2 165)	(2 170)	(694)	(717)	(720)
Precip. week	-	2008***	1612**	-	327**	319**
		(565)	(709)		(187)	(235)
Precip. month	-	-	238.30	-	-	5
			(258)			(85)
Adj. R-squared	0.024	0.027	0.027	0.010	0.010	0.010

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

Table 4.6. Damage costs per day and impact of extreme rain, defined at 40mm pr day (N =4749 days)

	(1)	(2)	(3)
Intercept	122 700***	-979	-39 450
	(19 600)	(28 300)	(37 700)
Extreme rain day	1 494 000***	1 368 000***	1 365 000***
	(208 000)	(208 000)	(208 000)
Precipitation week	-	3643***	2898***
		(605)	(773)
Precipitation month	-	-	439
			(283)
Adj. R-squared	0.011	0.018	0.018

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

Tables 4.6 – 4.8 show models with the “Extreme rain day” variable instead of daily precipitation. “Extreme rain day” is defined as the day having extreme rain (defined as a dummy with more than 40, 25, and 20 daily precipitation equal to 1; 0 otherwise; in tables 4.6, 4.7. and 4.8, respectively). Note there is no formal definition of an extreme rain day, and in Norway this would vary geographically due to different local conditions and could be as low as 20 mm in the driest areas but more than seven times this in the wettest areas in Western Norway (DSB 2016; p. 9)). As urban areas have built-up areas and less green areas that can absorb the rain, 20 mm is used here to define an extreme rain day throughout this thesis, expect in tables 4.6-54.8 where the sensitivity of the results to different definitions of an extreme rain day is explored.

For all definitions of “Extreme rain day”, the variable is significant, and has the expected positive impact on daily damage costs. Precipitation the preceding week also has a significant, positive effect on damage costs per day (see models 2 and 3) for all definitions of an extreme rain day. Adjusted R-squared is still low, and in general lower than for models using the continuous variable of daily precipitation (in mm). Thus, using extreme rainfall days as a dummy variable does not seem to increase the explanatory power.

Table 4.7. Damage costs and impact of extreme rain, defined at 25mm pr day, and weekly and monthly precipitation (N =4749)

	(1)	(2)	(3)
Intercept	93 290***	-10 740	-39 200
	(19 800)	(28 200)	(37 500)
Extreme rain day	1 033 000***	954 600***	949 000***
	(97 300)	(98 200)	(98 300)
Precipitation week	-	3131***	2583***
		(606)	(770)
Precipitation month	-	-	325
			(282)
Adj. R-squared	0.023	0.028	0.028

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

Table 4.8. Damage costs and impact of extreme rain, defined at 20mm pr day, and weekly and monthly precipitation (N =4749)

	(1)	(2)	(3)
Intercept	83 690***	-16 710	-40 330
	(20 000)	(28 200)	(37 500)
Extreme rain day	792 500***	722 600***	717 000***
	(78 100)	(79 100)	(79 300)
Precipitation week	-	3065***	2613***
		(609)	(771)
Precipitation month	-	-	270
			(283)
Adj. R-squared	0.021	0.026	0.026

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

In tables 4.9 and 4.10, mean daily temperature is added as a variable in the models, as a continuous variable (in degrees C) in table 4.9, and in table 4.10 as a dummy variable (above zero or not; thought to reflect that the risk of urban flooding and damages could be higher if the precipitation is rain rather than snow). The results show that in most of the models for the two temperature variables is the temperature variable not significant. The only exception is model 1 in table 4.9, where the mean daily temperature has a significant and positive effect on daily damage costs.

Table 4.9. Damage costs and impact of daily, weekly, and monthly precipitation and the arithmetic mean daily temperature. *Temperature day* is calculated as the arithmetic mean for a 24h period 00-24. (N =4749 days).

	(1)	(2)	(3)
Intercept	-36 810	89 950***	-101 600***
	(26 200)	(30 600)	(37 800)
Temperature day	6765***	5182	4947
	(2515)	(2557)	(2598)
Precipitation day	26 890***	25 000***	24 930***
	(2309)	(2374)	(2379)
Precipitation week	-	2100***	1867**
		(629)	(776)
Precipitation month	-	-	147
			(286)
Adj. R-squared	0.031	0.033	0.033

Table 4.10. Damage costs and impact of daily, weekly, and monthly precipitation and whether the arithmetic mean daily temperature was above zero. T*emperature day above zero* is calculated as the arithmetic mean for a 24h period 00-24 is above zero equals 1; 0 otherwise (N =4749 days)

	(1)	(2)	(3)
Intercept	40 930	-88 110**	-105 100**
	(40 200)	(42 200)	(47 200)
Temperature day	56 240	29 250	25 370
above zero	(45 800)	(46 300)	(46 600)
Precipitation day	27 280***	25 230***	25 110***
	(2313)	(2379)	(2383)
Precipitation week	-	2272***	1904**
		(626)	(777)
Precipitation month	-	-	227
			(284)
Adj. R-squared	0.030	0.032	0.032

In tables 4.11 and 4.12 variables to account for seasonal variations in rainfall have been added to the models. Three dummy variables (Spring, Summer and Autumn; see table 4.3. for definitions) have been constructed by converting the year into four 3-months periods according to the meteorological definition of seasons and leaving Winter out as the hidden/reference category. The results show that out of these three seasonal dummy variables, only *Summer* was significant. Thus, *Summer* has a significant positive impact on daily damage costs; showing rainfall occurring in the summer significantly increase the damage costs compared to the same rainfall occurring in winter. Looking at the number of insurance claims instead of the damage costs, table 4.12 shows that now both *Summer* and *Autumn* become significant (at the 5 % level). This indicates that the seasons better explains the variation in the daily number claims than the damage costs per day.

Table 4.11. Damage costs and impact of daily, weekly, and monthly precipitation and the seasons as dummy variables (N =4749 days)

	(1)	(2)	(3)
Intercept	-54 470	-100 420**	-115 100**
	(39 800)	(42 800)	(49 300)
Precipitation day	26 570***	24 930***	24 880***
	(2325)	(2380)	(2382)
Precipitation week	-	2033***	1 838**
		(643)	(777)
Precipitation month	-	-	135
			(303)
Spring	19 120	29 080	32 020
	(54 600)	(54 600)	(55 000)
Summer	150 500***	116 700**	112 100**
	(55 000)	(56 000)	(56 900)
Autumn	69 010	45 500	39 820
	(54 900)	(55 300)	(56 800)
Adj. R-squared	0.032	0.033	0.032

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

Table 4.12. Number of insurance claims per day and impact of daily, weekly, and monthly precipitation and the seasons as dummy variables (N =4749 days)

	(1)	(2)	(3)

Intercept	0.358	-0.339	-0.725**
	(0.254)	(0.272)	(0.313)
Precipitation day	0.260***	0.233***	0.231***
	(0.015)	(0.015)	(0.015)
Precipitation week	-	0.029***	0.022***
		(0.004)	(0.005)
Precipitation month	-	-	0.005**
			(0.002)
Spring	0.056	0.195	0.299
	(0.348)	(0.347)	(0.349)
Summer	1.070***	0.600*	0.432
	(0.351)	(0.355)	(0.361)
Autumn	0.957***	0.628*	0.427
	(0.350)	(0.351)	(0.360)
Adj. R-squared	0.067	0.077	0.078

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

Table 4.13 reports models using only the three precipitation variables in trying to explain the variation in the number of insurance claims per day. The results show that all three precipitation variables (i.e., rainfall daily, weekly and per month) have a significant impact on the number of insurance claims.

Table 4.13. Insurance claims per day and impact of daily, weekly, and monthly precipitation for 2008 - 2020 (N =4749 days)

	Model 1	Model 2	Model 3
Intercept	0.8326***	- 0.0638	- 0.5174**
	(0.143)	(0.183)	(0.239)
Precipitation day	0.2656***	0.2355***	0.2323***
	(0.015)	(0.015)	(0.015)
Precipitation week	-	0.0305***	0.0217***
		(0.004)	(0.005)
Precipitation month	-	-	0.0053***
			(0.002)
Adj. R-squared	0.065	0.076	0.078

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

4.5 Multiple regression models for weekly data

Tables 4.14 – 4.17 report the results from using stepwise regression to explain the variation in the weekly damage costs (tables 4.14 and 4.15) and number of insurance claims (tables 4.16 og 4.17); starting with all variables in Model 1 and deleting the variable with the smallest correlation, ending up with a model with only significant independent variables; e.g. model 8 in table 4.14.

As the amount precipitation (in mm) in a week and number of extreme rainfall days are highly correlated; r= 0.85, see table A-2 in the Appendix, the regression models include just one of these two variables. Thus, tables 4.14 and 4.15 show the results for weekly damage costs, using precipitation in week *t* and extreme rainfall days in week *t,* respectively. In the same way, tables 4.16 and 4.17 report results for the stepwise regression for the weekly number of insurance claims using precipitation in week *t* and extreme rainfall days in week *t*, respectively.

Table 4.14. Stepwise regression of *damage costs per week* and the impact of *precipitation* in week *t*, precipitation in the previous week (t-1), an in the second (t-2) and third week (t-3) prior to week *t*, and annual population data for Oslo. Seasons, and number of temperature days above zero in week *t* (N = 678 weeks)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	-3.2e+6 (2.5e+6)	-3.3e+6 (2.5e+6)	-3.1e+6 (2.5e+6)	-3.1e+6 (3.5e+6)	-3.1e+6 (2.5e+6)	-7.8e+5*** (2.8e+5)	-9.1e+5*** (2.6e+5)	-8.6e+5*** (2.6e+5)
Precipitation week t	4.0e+4*** (5361)	4.0e+4*** (5346)	4.0e+4*** (5354)	4.0e+4*** (5257)	3.9e+4*** (5223)	4.0e+04*** (5179)	3.4e+04*** (5157)	4.1e+04*** (5073)
Precipitation week t-1	1.2e+4** (5359)	1.2e+4** (5352)	1.3e+4** (5356)	1.2e+4** (5311)	1.2e+4** (5276)	1.2e+4** (5249)	1.1e+4** (5121)	1.2e+4** (5061)
Precipitation week t-2	-7139 (5360)	-6806 (5279)	-6515 (5259)	-6724 (5234)	-6984 (5179)	-6471 (5151)	-	-
Precipitation week t-3	1945 (5311)	-	-	-	-	-	-	-
Spring	4.7e+5 (5.7e+5)	4.6e+5 (5.7e+5)	2.0e+5 (4.2e+5)	1.4e+5 (4.0e+5)	-	-	-	-
Summer	1.0e+6 (5.3e+5)	1.0e+5* (6.3e+5)	7.1e+5* (4.3e+5)	6.3e+5* (3.9e+5)	5.9e+5 (3.7e+5)	5.7e+5 (3.7e+5)	5.2e+5 (3.7e+5)	-
Autumn	3.5e+5 (5.7e+5)	3.8e+5 (5.6e+5)	-	-	-	-	-	-
Population	3.81 (4)	3.96 (3.97)	3.72 (3.95)	3.69 (3.95)	3.76 (3.84)	-	-	-
Temperature days above 0 week t	-6.4e+4 (8.7e+4)	-6.4e+4 (8.7e+4)	-3.2e+4 (7.1e+4)	-	-	-	-	-
Adj. R-squared	0.111	0.112	0.113	0.114	0.115	0.115	0.114	0.113

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

The results from the stepwise regression in table 4.14 show that precipitation in week *t* as well as the previous week (*t-1*) have a significant, positive impact on damage costs per week. This corresponds well with results from the regressions based on daily data, presented in chapter 4.4. None of the other independent variables are significant (at the 5 % level). Compared to the models based on daily data in chapter 3.4, the explanatory power is much higher using weekly data, with adjusted R-square around 11 %. Thus, 11% of the variation in the damage costs can be explained by these two precipitation variables.

Using the number of extreme rainfall days in week *t* (instead of precipitation in week *t*), table 4.15 shows this variable and precipitation the previous week (t-1) are still the only significant variables to explain the variation in damage costs per week. The explanatory power in terms of adjusted R-square stays higher than with the use of daily data but drops by about one percentage point compared to table 4.14.

Table 4.15. Stepwise regression of *damage costs per week* and the impact of number of *extreme rainfall days* (> 20 mm) in week *t*, precipitation in the previous week (t-1), and in the second (t-2) and third week (t-3) prior to week *t*, Seasons, Population and the number of temperature days above zero in week *t* (N = 678 weeks).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	-3.8e+6	-3.7e+6	-3.8e+6	-3.8e+6	-3.8e+6	-3.6e+6	-3.1e+5
	(2.5e+6)	(2.5e+6)	(2.5e+6)	(2.5e+6)	(2.5e+06)	(2.5e+6)	(2.4e+5)
Extreme rainfall days week t	1.4e+6***	1.4e+6***	1.4e+6***	1.4e+6***	1.4e+6***	1.4e+06***	1.4e+6***
	(2.1e+5)	(2.1e+5)	(2.1e+5)	(2.1e+5)	(2.1e+5)	(2.0e+5)	(2.0e+5)
Precipitation week t-1	1.4e+4***	1.4e+4***	1.4e+4***	1.4e+4***	1.4e+4***	1.2e+4**	1.3e+4**
	(5376)	(5371)	(5362)	(5203)	(5288)	(5176)	(5127)
Precipitation week t-2	-7520	-7420	-7001	-7336	-7392	-	-
	(5402)	(5389)	(5301)	(5231)	(5224)		
Precipitation week t-3	2139	2337	-	-	-	-	-
	(5349)	(5310)					
Spring	3.2e+5	2.0e+5	1.7e+5	-	-	-	-
	(5.8e+5)	(4.3e+5)	(4.3e+5)				
Summer	1.0e+6	8.6e+5**	8.6e+5**	8.0e+5**	7.5e+5**	6.9e+05*	-
	(6.3e+5)	(4.3e+5)	(4.3e+5)	(4.0e+5)	(3.7e+5)	(3.7e+5)	
Autumn	1.8e+5	-	-	-	-	-	-
	(5.7e+5)						
Population	5.82	5.69	5.86	5.91	5.86	5.25	-
	(4)	(3.98)	(3.95)	(3.95)	(3.94)	(3.92)	
Temperature days above 0 week t	-4.8e+4	-3.3e+4	-3.1e+4	-2.1e+4	-	-	-
	(8.7e+4)	(7.2e+4)	(7.2e+4)	(6.8e+04)			
Adj. R-squared	0.098	0.099	0.100	0.101	0.102	0.101	0.100

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

Tables 4.16 and 4.17 shows a similar pattern as tables 4.14 and 4.15 in terms of significant precipitation variables, with both precipitation in week t and number of extreme rainfall days in week *t* having a significant positive impact (in table 4.16 and 4.17, respectively) on the number of insurance claims per week; and in both regressions precipitation the previous week (t-1) is significant. These results correspond well with previous Norwegian studies looking at the correlation between number of insurance claims and precipitation (Rohrbek et al 2016, 2018 and Torgersen et al 2015; see chapter 2) and shows that the more rain there is the preceding week (t-1), soils will be saturated and when there is more rain in week t the number of insurance claims in week t will increase.

Table 4.16. Stepwise regression of number of *insurance claims per week* and the impact of *precipitatio*n in week *t*, precipitation in the previous week (t-1), an in the second (t-2) and third week (t-3) prior to week *t*, and annual population data for Oslo. Seasons, and number of temperature days above zero in week *t* are dummy variables. A (N = 678 weeks)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	-31.77* (17.14)	-31.31* (17.1)	-31.18* (17.1)	-29.79* (16.89)	-29.58* (16.88)	-30.33* (16.84)	-28.99* (16.79)	3.24* (1.787)
Precipitation week t	0.366*** (0.037)	0.366*** (0.037)	0.363*** (0.036)	0.365*** (0.036)	0.363*** (0.036)	0.366*** (0.036)	0.373*** (0.035)	0.38*** (0.035)
Precipitation week t-1	0.141*** (0.037)	0.139*** (0.036)	0.137*** (0.036)	0.139*** (0.036)	0.137** (0.036)	014*** (0.035)	0.146*** (0.035)	0.151*** (0.035)
Precipitation week t-2	-0.016 (0.037)	-	-	-	-	-	-	-
Precipitation week t-3	0.029 (0.036)	0.026 (0.036)	0.026 (0.036)	0.029 (0.035)	0.026 (0.035)	-	-	-
Spring	3.359 (3.935)	3.459 (3.926)	2.273 (3.193)	1.382 (2.753)	-	-	-	-
Summer	5.337 (4.322)	5.217 (4.311)	3.759 (3.269)	2.706 (2.654)	2.333 (2.546)	2.526 (2.533)	-	-
Autumn	3.017 (3.9)	2.912 (3.89)	1.773 (3.212)	-	-	-	-	-
Population	4.3e-5 (2.8e-5)	4.2e-5 (2.7e-5)	4.1e-5 (2.7e-5)	4.0e-5 (2.7e-5)	4.1e-5 (2.7e-5)	4.3e-5 (2.7e-5)	4.1e-5 (2.7e-5)	-
Temperature days above 0 week t	-0.301 (0.595)	-0.309 (0.594)	-	-	-	-	-	-
Adj. R-squared	0.205	0.205	0.206	0.207	0.208	0.209	0.209	0.207

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

When using the number of insurance claims per week in tables 4.16 and 4.17 instead of weekly damage costs as the dependent variables, the explanatory power of the models nearly doubles to an adjusted R-square of 19.0 and 20.7 % (models 7 and 8 in tables 4.17 and 4.16, respectively); and highest when using precipitation per week (table 4.16) instead of number of extreme rainfall days (table 4.17). This might be explained by damage costs from the same amount of rainfall or the same number of extreme rainfall varying more than the number of insurance claims from the same precipitation.

In table 4.17 *Population* is also significant (at the 5 % level) and positive, indicating that increased population will increase the weekly number of insurance claims. This could be due to both more households that can be potentially affect and higher population leading to more densely built-up areas with less green areas that can absorb the rain. Note that *Population* does not have a significant effect on the weekly damage costs.

Table 4.17. Stepwise regression of *number of insurance claims per week* and the impact of the number of *extreme rainfall days* (> 20 mm) in week *t*, precipitation in the previous week (t-1), and in the second (t-2) and third week (t-3) prior to week *t*, Seasons, Population and the number of temperature days above zero in week *t* (N = 678 weeks).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	37.24**	-37.15**	-36.49**	-36.36**	-35.87**	-36.64**	-28.99*
	(17.3)	(17.28)	(17.07)	(17.06)	(17.03)	(16.99)	(16.96)
Extreme rainfall days week t	13.15***	13.10***	13.15***	13.11***	13.05***	13.17***	13.49***
	(1.453)	(1.44)	(1.425)	(1.419)	(1.414)	(1.402)	(1.389)
Precipitation week t-1	0.155***	0.153***	0.155***	0.154***	0.15***	0.153***	0.163***
	(0.037)	(0.037)	(0.037)	(0.037)	(0.036)	(0.036)	(0.035)
Precipitation week t-2	-0.02	-0.02	-0.02	-0.021	-	-	-
	(0.037)	(0.037)	(0.037)	(0.037)			
Precipitation week t-3	0.03	0.03	0.031	0.03	0.036	-	-
	(0.037)	(0.037)	(0.037)	(0.036)	(0.035)		
Spring	2.028	1.392	0.985	-	-	-	-
	(3.969)	(3.232)	(2.801)				
Summer	5.429	4.643	4.148	3.888	3.741	3.935	-
	(4.37)	(3.321)	(2.68)	(2.57)	(2.556)	(2.541)	
Autumn	1.434	0.827	-	-	-	-	-
	(3.946)	(3.28)					
Population	6.1e-5**	6.1e-5**	6.0e-5**	6.1e-5**	6.0e-5**	6.2e-5**	6.0e-5**
	(2.8e-5)	(2.8e-5)	(2.7e-5)	(2.7e-5)	(2.7e-5)	(2.7e-5)	(2.7e-5)
Temperature days above 0 week t	-0.166	-	-	-	-	-	
	(0.6)						
Adj. R-squared	0.187	0.188	0.189	0.190	0.191	0.191	0.190

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

5. Conclusion

This thesis set out to explore the following four research question:

i) Is there a correlation between rainfall data and the damage costs for insurance claims for urban flooding in Oslo, typically flooding of basements in built-up areas?

ii) Is there a correlation between rainfall data and the number of insurance claims for urban flooding in Oslo?

iii) What type of rainfall data – daily, weekly or monthly rainfall or discrete data in terms of number of "extreme rainfall" days (defined as having daily precipitation above a certain level) - can predict the observed number of insurance claims and damage costs from urban flooding?

iv) Do other variables affect the observed number of insurance claims and damage costs from urban flooding? Both the average daily temperature (i.e., increased temperature would lead to more precipitation being rain instead of snow) and the size of the population (i.e., with higher population, a higher number of households could be affected) could affect the insurance claims and damage costs.

The results show that there are large fluctuations in annual precipitation and annual number of extreme rainfall days over the 13-year period analysed. However, there is a clear correlation between rainfall and insurance claims, both in terms of numbers and costs.

Different rainfall indicators have been used and the results show that both for daily and weekly data of the number of insurance claims and the resulting damage costs, precipitation (in mm) that day/week and the preceding week have a significant (at the 5 % level) leading to increased number of insurance claims and damage costs. Precipitation the preceding month and the second and third week before the incident were included in the models, but precipitation more than one week before the incident was not significant.

Among the other variables, neither temperature nor population was not significant in most models.

The explanatory power of the best models is, however, only around 20 % (in terms of adjusted R-square). This indicates that other factors explain a large part of the variation in insurance claims and costs. Thus, future research should collect data and investigate the effect of local conditions with regards to green areas and implementation of technical and nature-based adaptation measures in order to better explain the variation in future damager costs and predict the impacts of future adaptation measures.

6. References

Cettner, A., Ashley, R. M., Hedström, A. & Viklander, M. 2012. Sustainable development and urban stormwater practice. *Urban Water Journal* 11:3, 185-197, DOI: 10.1080/1573062X.2013.768683

Chatterton, J., Viavattene, C., Morris, J., Penning-Rowsell, E., & Tapsell, S. 2010. The cost of the summer 2007 floods in England. Department for Environment, Food and Rural Affairs (DEFRA), UK. https://assets.publishing.service.gov.uk/media/602e9870e90e07660dec0b0a/The_Costs_of_the_Summer_2007_Floods_in_England_technical_report.pdf (Accessed August 20 2022)

Cortès, M, M.Turco, M. Llasat-Botija & M.C. Llasat 2018. The relationship between precipitation and insurance data for flood damages in a region of the Mediterranean (Northeast Spain). N*atural Hazzards and Earth System Science* 18; 857–868 https://nhess.copernicus.org/articles/18/857/2018/nhess-18-857-2018.pdf

DSB (2016): Risk analysis of urban flooding. Crisis scenarios 2016 of serious incidents that could harm Norway (In Norwegian). Report HR 2349 December 2016. The Norwegian Directorate for Civil Protection (DSB), Oslo. 68 p. https://www.dsb.no/globalassets/dokumenter/rapporter/delrapport_-regnflom_2016.pdf (Accessed July 10 2022)

Finance Norway (2021): Climate Report 2021 (In Norwegian). Finance Norway, Oslo. https://www.finansnorge.no/siteassets/skadeforsikring/klimarapporten/finans-norges-klimarapport-2021_enkeltsider.pdf (Accessed June 8 2022)

Hanssen-Bauer, I., E.J. Førland, I. Haddeland, H. Hisdal, D. Lawrence, S. Mayer, A. Nesje, J.E.Ø. Nilsen, S. Sandven, A.B. Sandø, A. Sorteberg og B. Ådlandsvik (2017): Climate in Norway 2100. Norwegian Centre for Climate Services (NCCS) Report no 1/2017. klimarapport-2100-engelsk-web-0160517.pdf

IPCC 2021: Climate change 2021. The Physical Science Basis. Sixth Assessment Report (AR 6) of the Intergovernmental Panel for Climate Change (IPCC). Report from Working Group I. https://www.ipcc.ch/report/ar6/wg1/

IPCC 2022: Climate Change 2022. Impacts, Adaptation and Vulnerability. Sixth Assessment Report (AR 6) of the Intergovernmental Panel for Climate Change (IPCC). Report from working Group II. https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/

Lamond, J. & E. Pennig-Rowsell 2014. The robustness of flood insurance regimes given changing risk resulting from climate change *Climate Risk Management* 2 (2014); 1–10 https://doi.org/10.1016/j.crm.2014.03.001

Rambøll 2019. Ekstremnedbør Oslo. Skadeomfang og kostnader. (Extreme rainfall in Oslo. Damage assessment and costs). Utredning , Oslo kommune, Klimaetaten. Mars 2019. https://www.miljodirektoratet.no/ansvarsomrader/klima/for-myndigheter/klimatilpasning/klimatilpasning-prosjekter/2018/studie-av-skadeomfang-og-kostnad-ved-ekstremnedbor/# (accessed August 1 2022)

Rasmussen, J. 2014: Klimatilpasning i København – Skybrud [Adaption to climatechange in Copenhagen – cloudburst]. In Norwegian. *In*: NTVA, ed. Teknisk naturvitenskapelig kunnskap i samfunnssikkerhetsarbeid, 2014 Trondheim.

Rohrbeck, C., D. A. Costain & A. Frigessi 2016. Bayesian spatial monotonic multiple regression *Biometrika*, 103 (1); 1–16.

Rohrbeck, C, E.F. Eastoe, A. Frigessi & J.A. Tawn, 2018. Extreme value modelling of water-related insurance claims, *Annals of Applied Statistics*, 12 (1); 246-282. https://projecteuclid.org/journals/annals-of-applied-statistics/volume-12/issue-1/Extreme-value-modelling-of-water-related-insurance-claims/10.1214/17-AOAS1081.pdf

Semandeni-Davies, A., Hernebring, C.; Svensson, G. & Gustafsson, L.-G. 2008. The impacts of climate change and urbanisation on drainage in Helsingborg, Sweden: Suburban stormwater. *Journal of Hydrology*, 350**,** 114-125.

SSB 2022: Population data for the municipality of Oslo 2008-2020, Statistics Norway (Statistisk Sentralbyrå SSB), Oslo. https://www.ssb.no/statbank/table/07459/tableViewLayout1/ (accessed august 2 2022)

Torgersen, G,. J.T. Bjerkholt, K. Kvaal & O.G. Lindholm 2015: Correlation between extreme rainfall and insurance claims due to urban flooding – case study Fredrikstad, Norway. *Journal of Urban and Environmental Engineering*, 9 (2);127-138

Willems, P. (2012). Impacts of climate change on rainfall extremes and urban drainage systems. London: IWA Publishing (cited from Torgersen et al 2015)

Wobus, C., Lawson, M., Jones, R., Smith, J., and Martinich, J. 2014. Estimating monetary damages from flooding in the United States under a changing climate, *Journal of Flood Risk Managment*, 7, 217–229. https://onlinelibrary.wiley.com/doi/full/10.1111/jfr3.12043

1. Appendix

Table A-1. Average monthly precipitation and damage costs in Oslo by month for the period 2008-2020.

Precipitation and damage cost each month in Oslo 2008 - 2020		
Month	Precipitation in mm	Damage cost in 2021-NOK
January	1 566	21 332 712
February	1 369	21 253 171
March	1 058	18 752 528
April	1 131	18 882 658
May	1 544	19 255 779
June	2 263	42 171 189
July	2 568	42 487 989
August	3 373	248 262 883
September	2 305	125 647 346
October	2 287	49 384 137
November	2 118	20 467 307
December	1 734	17 596 373

There are seasonal variations in damage costs, and table A-1 shows clearly that August is the month with highest rainfall and highest damage costs, followed by September and then October and July which have about the same average precipitation and damage costs (although the two months of June and July has somewhat higher average precipitation and October somewhat higher damage costs). As these months covers parts of two meteorological seasons, regression models for the data from these three months only have been run for the period 2008-2020. Table A-2 presents the results and shows that this neither results in more variables being (at the 5 % level) nor better model fit than models with observations for all months.

Table A-2. Damage costs and impact of daily, weekly, and monthly precipitation in the months of July, August, and September 2008 – 2020 (N =1196 days)

	(1)	(2)	(3)
Intercept	-24 190	-199 300**	-129 000**
	(87 900)	(123 000)	(191 000)
Precipitation day	54 010***	51 260***	51 360***
	(6 860.35)	(6984.78)	(6990.79)
Precipitation week	-	3991.35***	4600.51***
		(1971.21)	(2351.75)
Precipitation month	-	-	-473.11
			(995.39)
Adj. R-squared	0.049	0.051	0.050

*** Significant at the 1 percent level.
** Significant at the 5 precent level.
* Significant at the 10 percent level.

Figure A-1. Correlation matrix for variables used in the regressions using daily data.

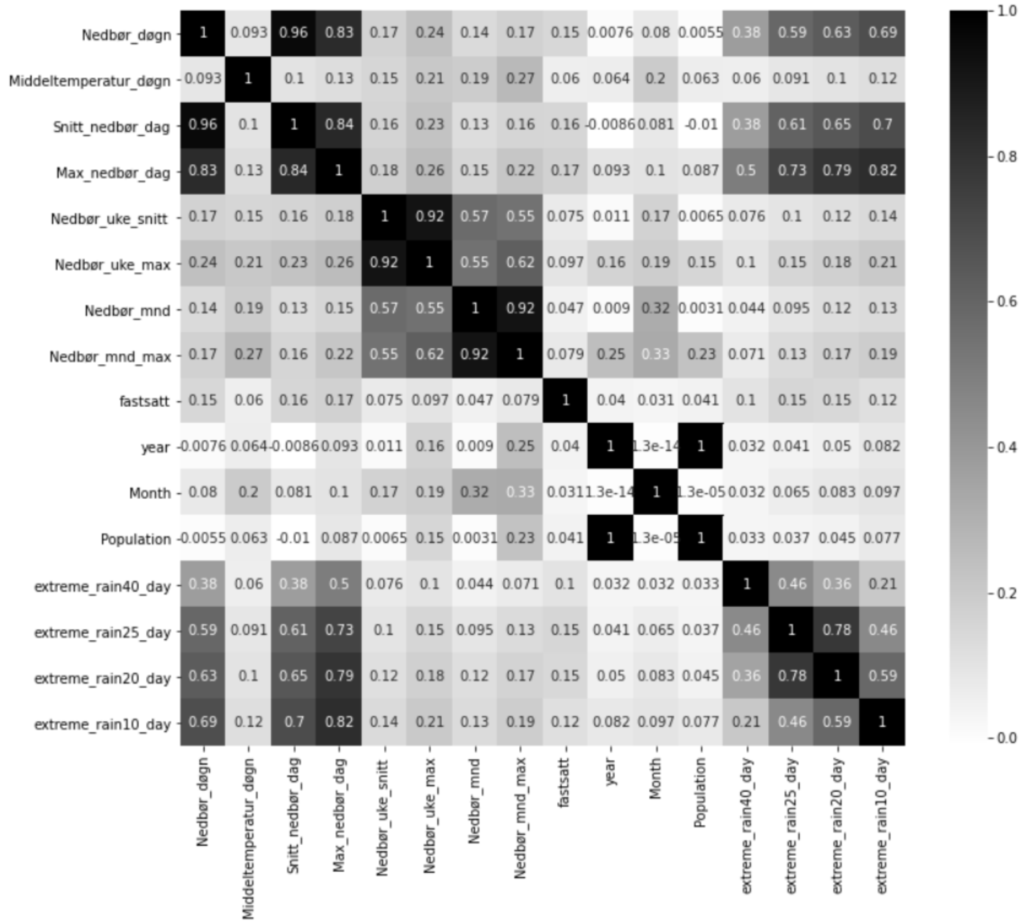
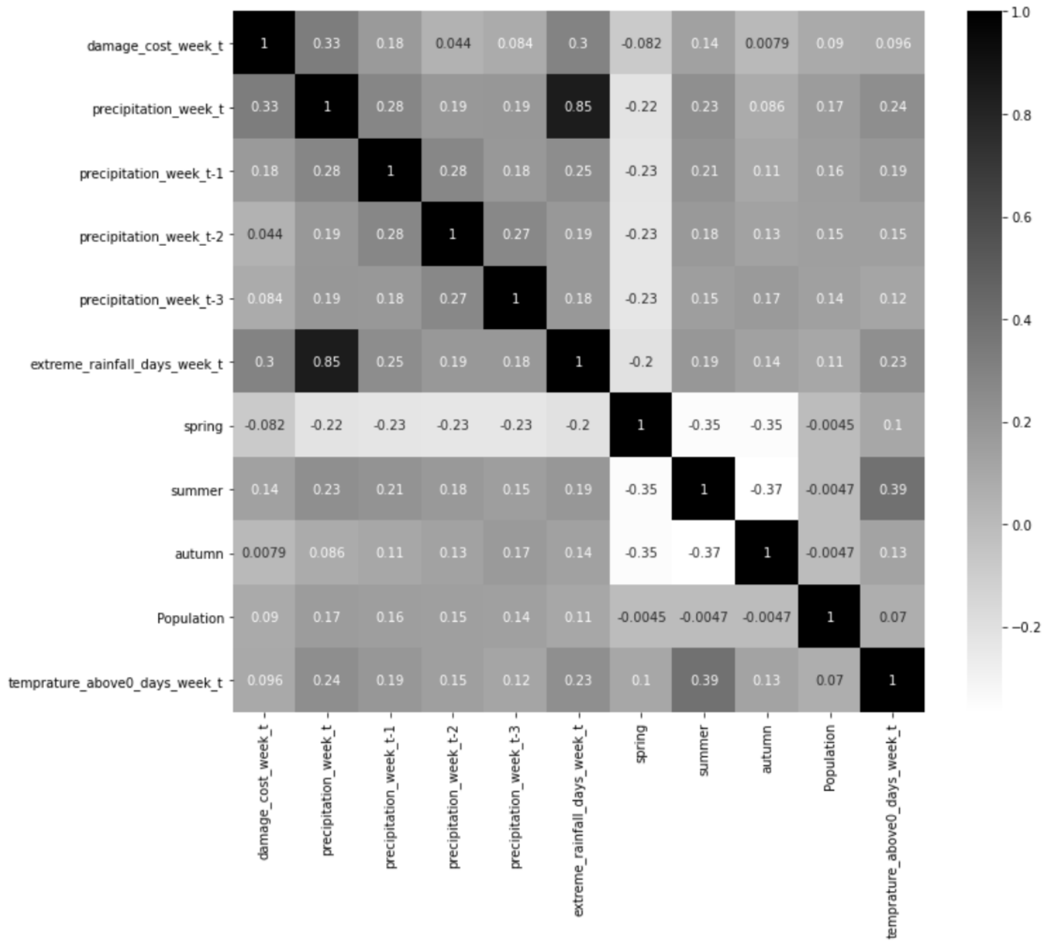


Figure A-2. Correlation matrix for variables used in the regressions using weekly data.



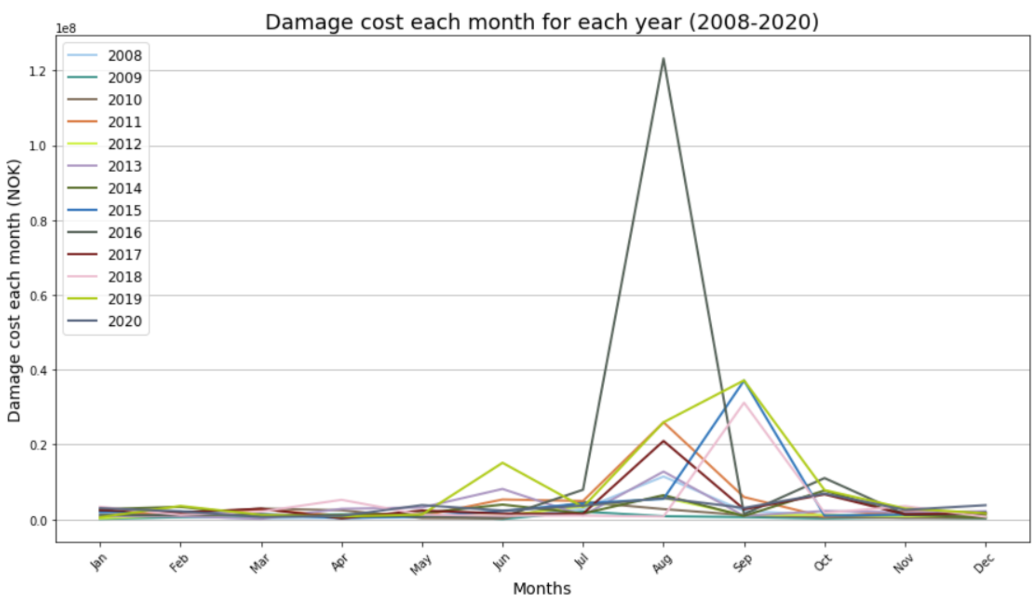


Figure A-3. Monthly damage costs (in 100 million 2021-NOK) for each year for the period 2008 -2020 (including 2016 with the extreme event on August 6th 2016)

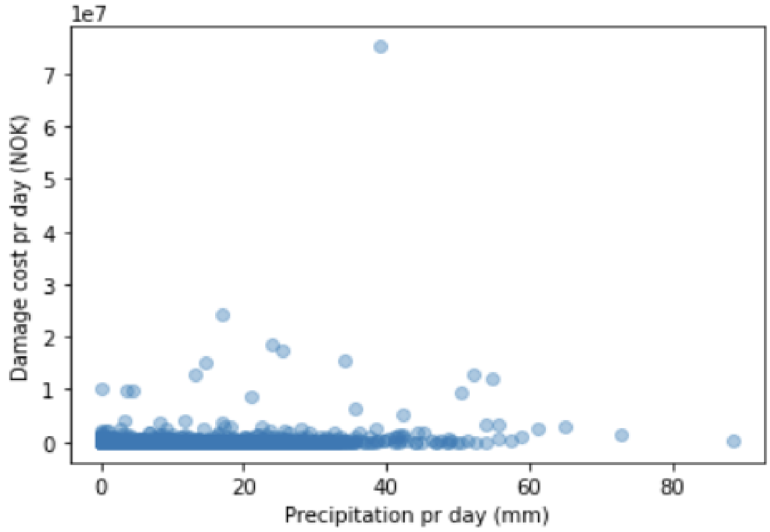


Figure A-4. Plot of daily data on damage costs (in 10 million NOK/day) and precipitation (mm/day)**.** All observations (N= 4749 days)

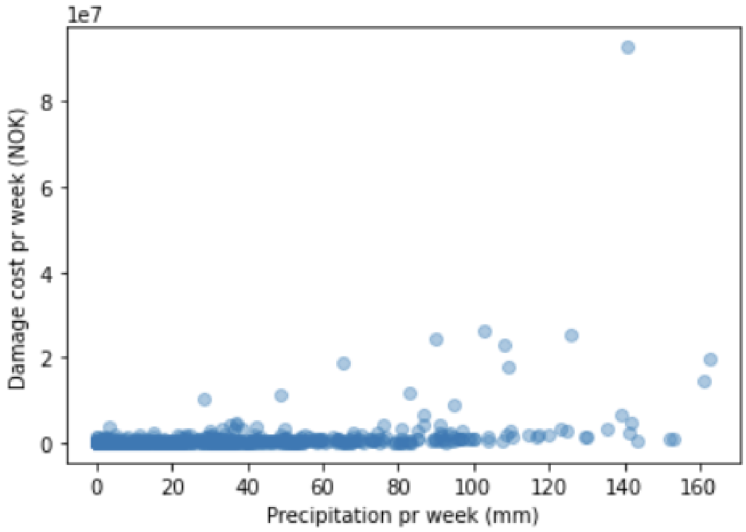


Figure A-5. Plot of weekly data on damage costs (in 10 million NOK/week) and precipitation (mm/week)**.** All observations (N = 678 weeks)