Norwegian University of Life Sciences

Master's Thesis Spring 2022 (30 ECTS)
Faculty of Science and Technology (REALTEK)

Statistical Analysis of Precipitation and Expected Annual Damage from Urban Floods

Nathan Amuri Bahati
Vann-og miljøteknikk

ACKNOWLEDGEMENT

Accomplishing a master's degree on full time and on the other side working work on full time hasn't been an easy task, and without the support of people around me, then this thesis wouldn't be in place. First and foremost, I would like to thank my wife Cecilia Chomba Amuri for being understanding, supportive, and motivating by always mentioning that short sentence to me "What you are doing is important, just do it and I will take care of our kids". She drove me all the way from the time I read the first word when starting this course, to the time I wrote the last word in this thesis.

I would also thank my thesis supervisor Vegard Nilsen of faculty of science and technology at Norwegian University of Life Sciences for facilitating the accomplishment of this thesis on full time studying progression, while I was also working full time basis. He took my situation into consideration and was always there to answer my e-mails at any day of the week. He continuously permitted this thesis to be my own work, his discipline and high work ethics steered med in the right direction whenever I needed it.

ABSTRACT

Expected damage costs of a pluvial flood is obtained from the product of probability of failure and cost associated with the failure, for which intensity duration frequency (IDF) relationships play a critical role in determining the precipitation amount for specific return periods which are in turn functions of precipitation durations. The construction of intensity duration frequency goes back to as early as 1932(Bernad 1932). The curves play a very important role in designing of urban infrastructures such as bridges, roads, culverts, and dams. Basically, these curves are constructed by considering different durations regardless of their relations to each other. The main objective of this thesis is to investigate the extent at which extreme rainfall events of different durations are related to each other (overlaps) and assess whether the results obtained from such analysis would be of practical relevance in determination of expected annual damage (EAD). This problem formulation was answered by studying rain records of Blindern precipitation station in Oslo. The study was accomplished in four main stages. The first stage being the extraction of first fifty-four independent extreme values of each duration from a single minute up to 1440 minutes duration, and the number of events in each duration corresponded to the timespan of the data series. The independent extreme values selection was secured by a time space function, the spacing function provided the possibility of increasing extreme value's ability to overlap with other precipitation intervals, and those events which overlapped with the extreme value intervals were assigned a zero-precipitation values, so that they wouldn't be considered when the next extreme value was being chosen in the duration. This procedure guaranteed that the selected extreme values (within the same duration) were independent of each other, and each duration had 54 independent extreme values. The second stage involved finding the extent at which some of the selected 54 events of a given duration overlap with 54 extremes events of all other durations form a single minute up to 1440 minutes duration. Third stage involved merging all intervals obtained in stage one into independent merged events. The final stage involved the extraction of events with highest return periods within the merged independent events by interpolating their return periods from the original station's IDF curves, thus Blindern's IDF curves, and visualize this in the same original IDF diagram.

The results of each of the above-mentioned stages pointed toward one direction, thus precipitation events have the tendencies to overlap, particularly events with higher duration, and that these events cannot be described regardless of their overlapping abilities. All stages correlated, for instance, the analysis performed in the second stage indicated that all events, independent of their duration, have high ability to overlap with other events. Events with short durations overlap mostly with events which have higher duration than their own durations, while the opposite happened with events of long durations. Visualizing of number independent merged event obtained in the third stage as function of maximum events duration, showed little variation of number independent merged events obtained, this events characteristic reflects precipitations event's abilities to overlap. The same event's tendency was again manifested when independence merged events was again visualized as function of the spacing function described in the first stage of the analysis. The incorporation of events with highest return periods obtained from final analysis into the original Blindern's IDF curves showed that most of these events were distributed around two years return period. It also showed that most of independent events were events with low durations, this result correlated with the result obtained from visualizing the overlapping probability of low duration events (with all other events), for which non- continuous lines were obtained on the overlapping probability for very low duration events. The results obtained from the above-mentioned analysis, describing dependence of precipitation events across different durations could be of practical relevance in way that, it can be used in calculations of catchment's summative expected annual damage (EAD) across event's durations, for which the current method is solely based on a single event duration.

TABLE OF CONTENTS

LIST OF FIGURES

1 INTRODUCTION

Natural hazards, such as floods and landslides often happen in Northern Europe, in addition to general wear are these hazards main cause of damage to infrastructure (Arnbjerg-Nielsen and Fleischer 2009). Their occurrence is expected to increase with climate change, particularly in small catchments (Madsen, Arnbjerg-Nielsen et al. 2009). To minimize the risk cause by named hazards, intensive research has been going throughout the globe with focused-on climate change and flood management. Another environmental concern is the increase in urbanization around the globe, which in turn increases the quantity and frequency of pluvial flood (Madsen, Arnbjerg-Nielsen et al. 2009). Therefore, knowledge of climate change and flood management is becoming more and more prominent(Arnbjerg-Nielsen 2011).

Calculation of expected annual damage (EAD) caused by urban flooding is of great interest for both water engineers and other stakeholders. It's a measure which is being used to indicate an extent at which a given area can be vulnerable to flood and in determination of the reward to be gained by implementing climate adaption measures. Zhou (Olsen, Zhou et al. 2015) outlined a foundation for identifying climate change adaption options to urban flood risk, for which the main element has been the calculation of the expected annual damage (EAD). These calculations are based on the information about the current and future extreme precipitation records (IDF). The approximation of EAD is associated with some assumptions and criteria to be fulfilled to trigger the use its equation, among other is event's statistical dependence which is considered as a criterion used to determine the number of event's terms to be included in the EAD equation. This thesis analyzes the statistical dependence of precipitation events and explore the possibilities of extending the EAD equation to more than one term based on criteria that events are statistically dependent, thus they overlap.

1.1Background and motivation

Rainfall-runoff models of different levels of sophistication are used to design and assess urban drainage systems, including flood damages. These models include among others, the rational formula, unit hydrograph, and PQRUT. The rainfall input to such models is often based in some ways on IDF statistics. For instance, the determination of rain intensity values to be used in the rational formula has it trace from IDF, for which the intensity value of a given precipitation event is determined from the curves when the concentration time of the catchment is known. Therefore, we can conclude that rain intensities values of these models are extracted from IDF curves, thus it is also possible to relate rainfall and runoff to frequency of precipitation occurrence. Furthermore, consider the assumption that, a point precipitation represent a small urban catchment, and flooding is insensitive to the temporal distribution of rainfall within individual events, and that runoff or flooding is insensitive to antecedent conditions in the catchment; then it is possible to conclude the following; for a given precipitation event duration d, it is therefore possible to obtain damage function $DF_d(T)$ that returns the catchment flood damages (e.g in NOK) that happen during a specific rainfall event with return period T. The expected annual damage which relate the intensity of the precipitation to the resulting damage, EAD_d, for a giver duration can be computed from the following equation (Rosbjerg 2017):

$$EAD_d = \int_0^\infty \frac{DF_d(T)}{T^2} dT \qquad (1.1)$$

Where:

- EAD_d = Expected annual damage for an event with duration d
- T = Return period for a given event
- D = Damage function of a precipitation event
- F_d(T) = The relation between the average rain event intensity and its return period for event duration d represented by IDF curves.

As mentioned earlier, that construction of IDF curves is based on times series for point precipitation. The method to its construction is always as follow; for a given duration, one extract extremes precipitation events or annual exceedance series and perform a frequency analysis for each of selected durations to determine their return periods (Anita Verpe Dyrrdal 2022). Thereafter, the final step being the drawing of the curves which are average precipitation intensity values as functions of both return periods and precipitation durations. One of the prerequisites for performing such calculations is to ensure that events are independent of each other, and this condition trigger the application of some probability distribution, for instance Poisson distribution ((Chin, Mazumdar et al. 2000), (Donna Wilson 2011)). Consider the current method called METmethod which is based on extraction of annual exceedance series of each duration, and the previous method also based on a specific threshold determined from the length of event's durations, both methods does not consider the description of precipitation events across different durations. Therefore, because of this IDF curves extracted from these two methods cannot be sufficient to determine the summative EAD of events with different durations. There is a need of an extended version of EAD equation. The equation (1.1) is based on the way IDF curves are constructed, thus current IDF curves are built without taking event's ability to overlap into consideration, thus the equation (1.1) can only be used to determine the EAD of single independent duration. In other words, the equation above must be extended with more of the same terms which stand for different event durations, and simultaneously it should include a parameter which denotes the contribution of each duration to EAD for it to be eligible for the summative EAD computation across durations, and that equation can be written as follow:

$$EAD = \alpha_1 \int_0^{\infty} \frac{DF_{d1}(T)}{T^2} dT + \alpha_2 \int_0^{\infty} \frac{DF_{d2}(T)}{T^2} dT + \cdots + \alpha_n \int_0^{\infty} \frac{DF_{dn}(T)}{T^2} dT \qquad (1.2)$$

Alpha (α_1) terms in equation (1.2) represent the contribution weight of each duration, therefore considering that there is some degree of statistical dependence (degree of overlapping) between events of different durations. In such a way that, if events were totally independent (not overlapping), therefore all alpha values will be one, and if they were totally dependent (complete overlap), all alphas would be zero except for one duration, thus the equation would have only a single term, and finally if they partially overlap, then alpha values would be represented by the degree of overlapping. This thesis analyzes precipitation event's abilities to overlap, and the result of this will be used to trigger a prospect of computing a summative EAD of events with different durations. In addition, it's also analyzed the extent at which independent precipitation events propagate throughout the original IDF curves.

1.2 Objective

The thesis presents sensitivity analysis of precipitation event's ability to overlap and consider the practical relevance of such analysis in computation of summative EAD of different event durations. In addition, it also considers what effect does non-consideration of such analysis has on the computation of the summative EAD across different precipitation durations, and finally the extent at which independent merged event extracted from this analysis will be distributed throughout the original IDF curves.

1.3 Problem formulation

- What extent overlapping/correlation of recorded extremes events of different durations exists in the data?
- Is the extent of overlapping of practical relevance, i.e is there a danger of underestimating risks by accounting for only a single duration?
- How do the recorded independent extreme events propagate in existing IDF diagrams?

1.4 Literature study

The distribution, movement and properties for natural surface water is being studied within a discipline of engineering hydrology called surface-water hydrology. The discipline includes modelling of precipitation events and estimating the quality and quantity of the resulting surface runoff. Furthermore, the quality and quantity estimations are used to design water drainage and treatment systems. This thesis is limited only to some procedures used to estimate water quantity, therefore the literature study cover theory behind precipitation measurement, return period calculations by the newly introduced METmethod and short description about climate surcharge.

1.4.1 Overview of Nordic climate

Nordic hydrology is mainly characterized by cool climate, heavy rainfall, and geology of Quaternary ice ages characteristics. Norway has a significant topography, and warm sea temperatures due to warm current from a gulf stream. The Nordic region is in the middle of the westerly wind belt along the polar front and arctic front, this result to a large quantity of moist air in autumn and winter. Areas with low altitude receive moderate quantity of precipitation, but variation in altitude is very influential to precipitation quantity due to raising and cooling of the air. Figure 2 below denotes the accumulation of precipitation depths of different regions of the country within a period of two decades, with the western coastal area having the highest records. Relationship between temperature and precipitation is vital in hydrological contest. High precipitation and low temperature provide good conditions for bogs development, this is a common landscape feature in Nordic countries, and very important in hydrology. For material balance of glaciers, winter precipitation and summer temperature are vital. Thus, high winter precipitation causes large accumulation, and high summer temperatures cause large glacier melting. Lakes in Nordic countries play a very important role as runoff leveling elements. Thus, with lakes in catchment areas, the flood is always delayed reaching rivers and other flood infrastructure systems.

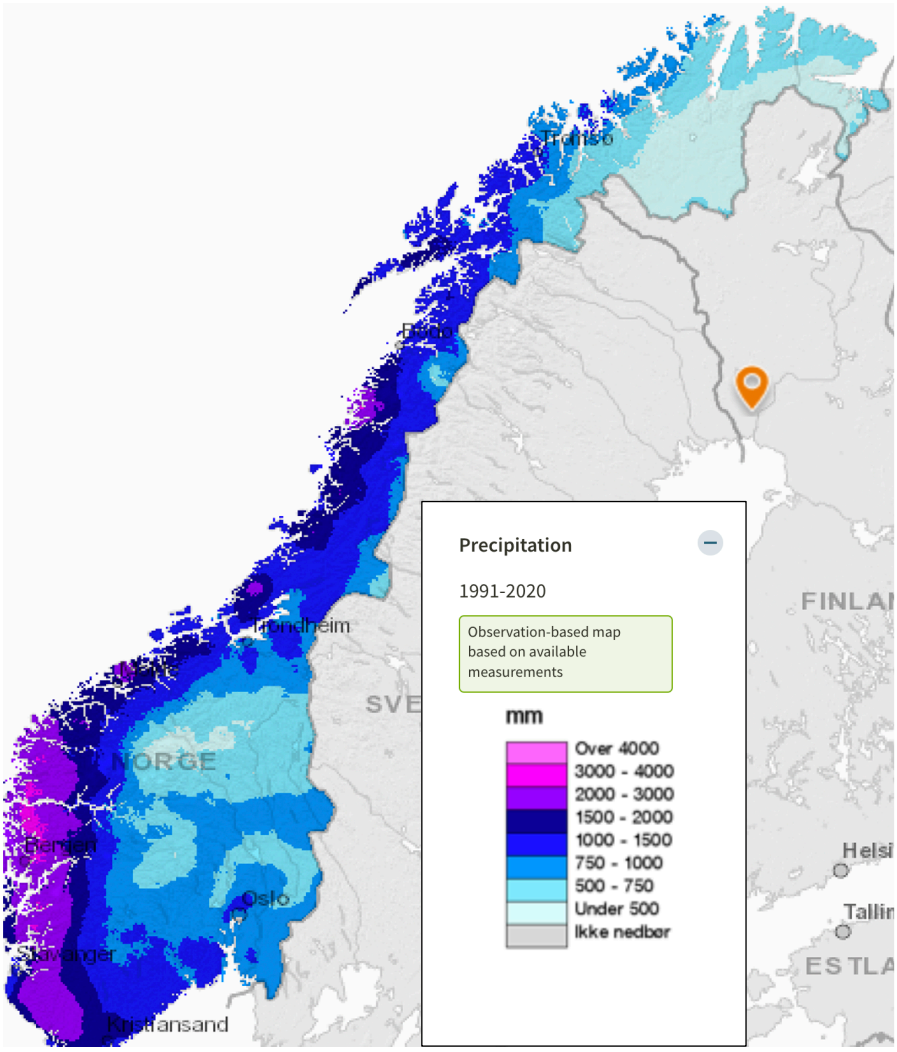


Figure 1: Cumulative precipitation depth distribution of Norway(seNorge.no)

1.4.2 Precipitation measurements and records

Precipitation records have been measured for several decades by different methods, with the traditional method of measurement being gauges. In hydrology, the amount of precipitation is described by the volume of precipitation descending per unit area and is given as a depth of water. Precipitation measurements by rain gages are point measurements of rainfall and only stand for a limited area surrounding the rain gauge.

1.4.1.1 Weighing rain gauge

Gauges for measuring precipitation can be classified either as a nonrecording(manual) or recording, with most modern ones being considered as recording. Precipitation can be measured in three different ways using the rain gauges. Figure 1 shows the three gauges, with (a) being a weighing gauge. The weighing gauge is based on the mass of precipitation being continuously monitored. Since a liter of water has a mass of one kilogram, then any difference in mass can be easily transformed into a difference in precipitation volume, therefore rainfall depth. Modern weighing gauges are electronic-

weighing and analog-recording instruments, meanwhile older ones are mechanical- weighing and analog-recording equipment. Uncertainties in rain intensities measured by these two different types of gauges has been studied, and the result showed significant discrepancies for precipitation duration of 5 minutes(Keefer, Unkrich et al. 2008). Certain types of weighing gauges use a pen on rotating drum, or by utilizing a shaking wire attached to a data logger. This type of gauge has some advantages over the tipping gauge, that it does not underestimate precipitation with higher intensities, can monitor other forms of precipitation including rain, hail, and snow. Disadvantages with them are that they are expensive and usually have higher operational cost because of frequently maintenance.

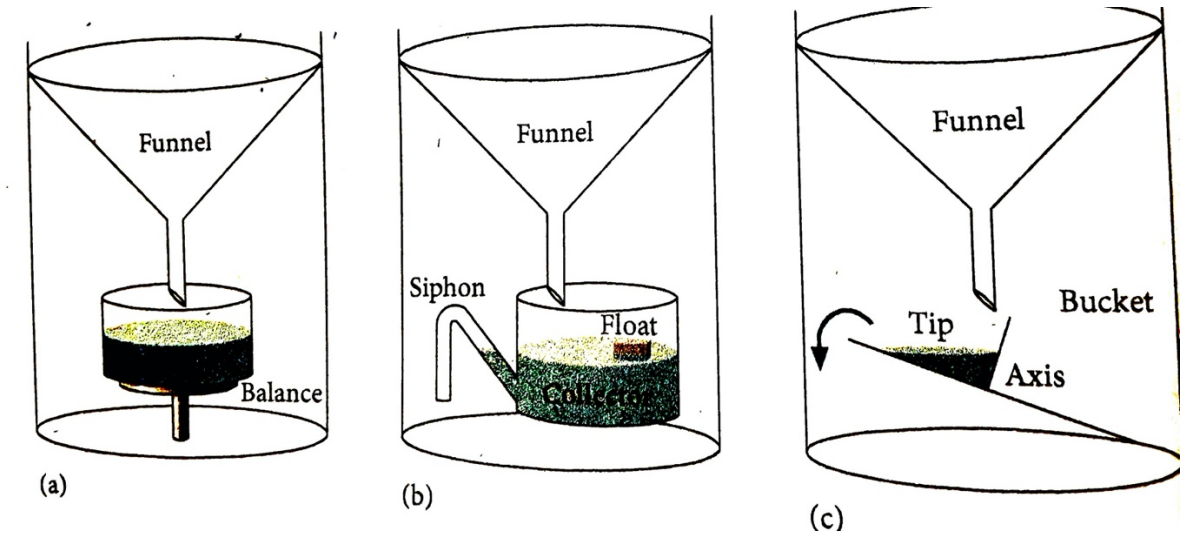


Figure 2. Weighing(a), float and siphon(b), and tipping bucket gauge (Hendriks 2010)

1.4.1.2 Tipping bucket rain gauge

The original tipping-bucket precipitation gauge was discovered by Sir Christopher Wren around 1662, and covers almost half of world precipitation record gages (Rahimi, Holt et al. 2003). A tipping bucket rain gauge is made up of a funnel, magnet and two small buckets. The rain passes through the funnel, and dropping into the two small buckets, balanced on a pivot. The two buckets are positioned by a magnet until they have been filled to the calibrated quantity (normally 0.1 or 0.2mm of rain). Immediately the buckets are filled to the mentioned amount the magnet will release its hold, leading to tipping of the bucket. As soon as the buckets tips, it initiates a reed switch (or sensor), sending signals to the display or weather station. The display monitors the frequency at which the switch is being initiated. By monitoring the timing of the tips, the average rainfall intensity each time the bucket tips, as well as rainfall depths over longer periods of time, can be determined. Common problems with tipping-buckets rain gages are blockages, wetting and evaporation losses usually of quantity 0.05 mm per rain event(Berndtsson and Niemczynowicz 1988). The tipping process takes half a second, and during this period there is always a missing of rain records (Marsalek 1981). Rain blown by wind, position of the gauge relative to other barriers and dynamic rainfall effects (Habib, Meselhe et al. 2008).

1.4.1.3 Float and siphon rain gauge

In this gauge the height of water in the cylindrical water collector is continuously monitored and recorded using the float on water surface. The Siphon emptied the water immediately the maximum water level has been reached; this corresponds to the bend position of the siphon. By this procedure water depth and intensities are being estimated.

1.4.4 Return period

A traditional application of probability theory in water management engineering includes the calculation of an exceedance probability for hydrological events (P_e). A given flood management system modelled to control a specific hydrological even is expected to fail with the probability equivalent to probability of exceedance (P_e), of the modelled event. And the probability of failure is known as the risk of failure, and the opposite, thus the probability that the system will not collapse is known as the reliability of the system. In hydrological contest the exceedance probabilities of events are normally determined in a state that a given hydrological event is exceeded in any given year. The average number of years between exceedances is known as the return period (T). Thus, a given event can be an exceedance value of different events, provided that the size of the event is bigger than other named events. The relationship between the exceedance probability and return period is represented as follow:

$$T = \frac{1}{P_e} \qquad (1.3)$$

Or

$$T = \frac{1}{P(X \geq x_T)} \qquad (1.4)$$

Where:

- T = Return period for a given event
- $P(X \geq x_T)$ = Probability of exceedance
- x_T = Given event with return period T

In hydrological sense it is more traditional to describe a precipitation event by its return period rather than its exceedance probability. For instance, floodplains are normally delineated for the “hundred-year flood”, which is an exceedance probability of one percent in any given year. In Norway most of flood management systems are design at return period of 10 years, and this selection is based on the frequency of the events, for which events with high return period rarely happens. Generally, the return period of a design rainfall should be selected based on economic efficiency, for which guidelines of selection are provided by the government institutions. Table below shows a guideline for selection of return period suggested by the Norwegian state highways authority, use for roads construction. In this case the criteria for selecting an appropriate return period being the yearly traffic, and the presence of detour on the road. Many are not familiar with the term “return period”, therefore there is always

a misconception about this term. Many believes that the term stands for a regular period for which a specific flood magnitude occurs, in contrast the term stands for probability of occurrence for a given flood which is always very dependent of environmental factors, such as climate and catchment characteristics.

Table 1: Safety class for roads (Vegvesen 2018)

Safety class	Year-round traffic	Return period T	
		With detour option	Without detour option
V1	0-500	50 years	100 years
V2	500-4000	100 years	200 years
V3	>4000	200 years	200 years

The Selection criteria of return periods may vary, as mentioned the main objective of selection being economic efficiency. Figure 3 shows a diagram of a drainage system, designed with different return periods and different risk level being one of the main criteria for selection. In selecting the return period for a given project, local drainage regulations and guidelines should be followed. As mentioned earlier, an implicit assumption in modelling drainage systems for a given return period of a precipitation event is that the precipitation is resulting to a runoff which is equivalent to the return period of the design rainfall. With a regard to this assumption, the risk of failure of the drainage system is equivalent to the exceedance probability of the design precipitation.

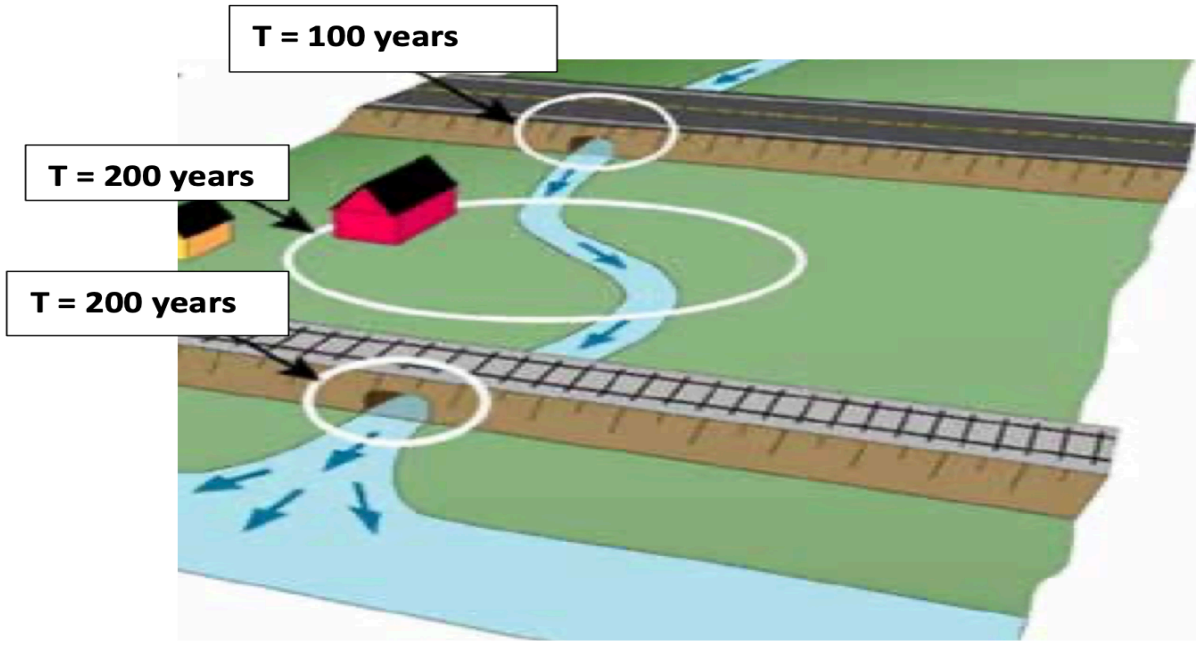


Figure 3: Different return period's dimensioning along the same way (Vegvesen 2018)

The return period, T, also known as the average recurrence interval (ARI) is normally considered as a criterion for choosing design events for drainage systems. In a situation where water resource will be available in a certain period of the years, thus less than a year, then it might be appropriate to estimate the return periods of the design events with reference to their occurrence within a specific season of the year (McCuen and Beighley 2003). Natural channels are being assumed to have a surface runoff having a return period of approximately 2 years (He and Wilkerson 2011), and the procedure behind this assumption plays a critical role in determination of flows corresponding to a specific return period, therefore, being very important in the restoration of natural channels. However as briefly mentioned, the expression "return period" can be misinterpreted, many who have little, or no knowledge of extreme value statistics may think that the expression implies that a specific flood size is only exceeded at a regular interval, or that it stands for a fixed period until the nest occurrence. It is important to bear in mind that the likelihood of a flood event may be different because of natural variability such as environmental changes, such as changes in land and climate, therefore flood sizes can not be exceeded at a regular interval.

1.4.5 IDF curves.

Precipitation records are rarely used directly in flood approximations, but rather the statistical analysis of rainfall records is normally utilized. Precipitation statistics are normally presented in form of intensity-duration -frequency also known as intensity frequency duration (IDF) curves (Gyasi-Agyei 2005), which denotes the relationship between average rainfall intensity and time(=duration), and the average intensity being also a function of a specific probability of occurrence. In Norway, the Norwegian meteorological institute has been responsible for extraction and calculation of IDF diagrams for all Norwegian precipitation stations with timeseries which span ten years back in time. The previous method which was used to derive IDF curves until January 2018 was based on a procedure which was established in 1975 and was characterized with uncertainties in its results. Particularly on events with short durations and long return periods. Therefore, there was a need of introducing another method called MET method with a goal to quantify uncertainties in IDF values. Both previous and current method adapt a generalized extreme value (GEV) distribution to the recorded rainfall (Coles, Bawa et al. 2001). The GEV distribution describe the distribution of maximum values per unit of time, e.g., the distribution of annual-maximum rainfall (AMR) for different durations which is converging toward the GEV distribution G(x) when the timeseries increases. Thus, there are three parameters which control the GEV distribution (see equation 1.5), and they are these which needs to be estimated. The GEV distribution with the inclusive of these parameters can be described as follows:

$$G(x) = exp\{-[1 + \xi(\frac{\chi - \mu}{\delta})]^{-1/\xi}\} \; for \; 1 + \xi(\frac{\chi - \mu}{\delta}) > 0, \qquad (1.5)$$

Where:

- μ = Location.
- δ = Scale parameter.
- ξ = Shape parameter.

1.4.5.1 The previous method

The previous method was previously updated in 2017, and as mentioned before, the curves were based on the procedure which was established in 1975. The method adapted the Gumbel distribution, thus the GEV distribution with zero shape parameter(Coles, Bawa et al. 2001) for extreme precipitations of all durations and return periods presented in figure 4. With the length of timeseries being the number of years included in such calculations, then extreme precipitation values in the dimensioned number of years were being obtained from sliding time window, thus the extreme value could be extracted at any random time for which the event had happened, and there was no requirement that the station had been in operation in a specific period. Some of the criteria which had to be met were that the event had to take place in 80% and 75% of its duration for 1-30- and 45-360-minute's events respectively. There were no equivalent criteria for longer events rather than precipitation depth record had to reach 15 and 20 mm in 720 and 1440 minutes respectively. In addition, they had to be a certain time interval between two events so that extreme events could be considered as being independent to each other. Some of these intervals were, an hour for 1-3 minutes events and two hours for five, three hours for ten and six hours for fifteen-twenty minutes events. Even though an effort was being done to secure events were independent of each other in such calculations, inconsistent between durations could still exist in IDF curves, this is because dimensioned precipitation values were obtained from separated durations, that the estimated events for a given duration could be less than the previous events. In. such situation, the event was being removed to secure ascending curves on IDF diagrams, and this method was known as the Gumbel-Reduced Variate or Gumbe-RV((Gumbel 1954), (Gumbel 1958))

1.4.5.2 The current method (METmethod)

The current and new method was launched in March 2022, it implements sliding time window on timeseries of one minute time interval to extract annual maxima of each event duration. In this case the criteria for approving annual maxima values are that the precipitation event covers 80% of its duration, and that the event has to happen in a period between May and September, this is so because this time of the year is considered as a period in which low duration and high precipitation events occurs frequently in Norway (Anita Verpe Dyrrdal 2022). Three hours duration is the most important event's duration to be considered by this method. The method implements the same durations and return periods presented in figure 4 as well.

The extremes data analysis is based on Bayesiansk inference for estimating all probability distribution of parameter set θ which is turn a function of all GEV-parameters as $\theta(\mu, \delta, \xi)$, the method is known as GEV-Bayesiansk (GEV-Bay) (Lutz, Grinde et al. 2020). It is different from other methods because it estimates each of the GEV parameters more than once and present them in a distribution format. The inference is as well based on Bayes's theorem which states that; the conditional probability of an event is based on the knowledge of occurrence of previous events. This knowledge can be extracted from precipitation timeseries, and trigger the use of the following equation in calculation of probability of parameter set θ:

$$P(\theta/x) = \frac{(L(x/\theta)\mathrm{P}(\theta))}{P(x)} \quad (1.6)$$

Where:

- $P(\theta/x) =$ *is the probability function for* θ *given observation* x
- $\mathrm{P}(\theta) =$ Posterior probability
- $L(x/\theta) =$ Probability function
- $\mathrm{P}(\theta) =$ prior distribution for θ

Given that $P(x)$ is a constant then $P(\theta/x)$ can be written as a product of $L(x/\theta)$ and $\mathrm{P}(\theta)$ as follow:

$$P(\theta/x) \propto L(x/\theta)\,\mathrm{P}(\theta) \quad (1.7)$$

This method calculates the dimensioned precipitation values separately as well, thus there is a possibility of obtaining inconsistence between different durations in IDF curves. The inconsistencies are being corrected by the method described by Roksvåg (Roksvåg, Lutz et al. 2021).

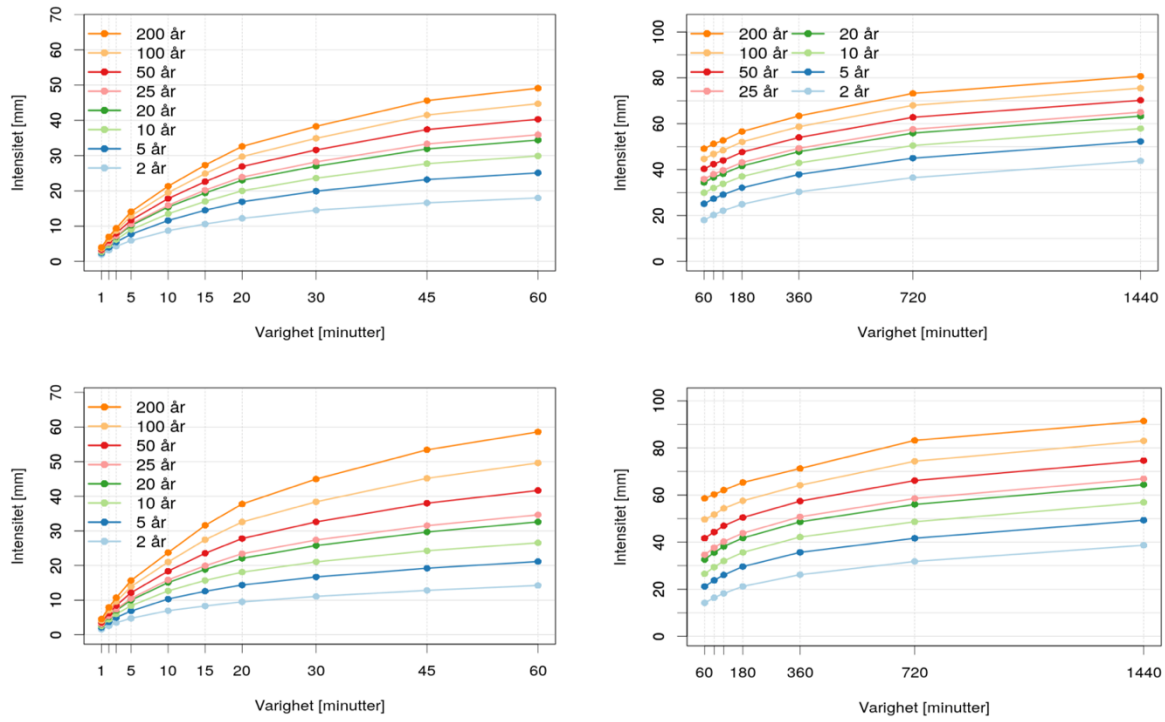


Figure 4: IDF curves for Oslo-Blindern for 1-60 minutes (left) and 60-1140 minutes (right). With return periods of the top two diagrams estimated with Gumbel-RV and two bottoms' diagrams estimated with GEV-Bay(Anita Verpe Dyrrdal 2022).

The main objectives of introducing a new method of extraction of IDF curves was primarily the inconsistency of Gumbel-RV. The Gumbel-RV which is being used in Gumbel distribution is well known for its ability to underestimate the highest precipitation values ((Papalexiou and Koutsoyiannis 2013),(Koutsoyiannis 2004)), and the second reason being the Bayesiansk statistic which is being applied in GEV-Bay provides a possibility of direct and intuitive quantifying of uncertainties in IDF values (Anita Verpe Dyrrdal 2022). Figure 4 represents curves extracted from the previous method (Gumbel-RV) and the new method (GEV-Bay) for Oslo-Blindern precipitation station. In general, one can observe that the forms of these curves are identical, and the difference is the distribution between

return periods for which GEV-Bay has higher distribution. One can also observe that low return periods are relative lower and higher return periods are higher in GEV-Bay than Gumbel-RV.

1.4.5 Climate surcharge

Another parameter which is related to rainfall statistic is the climate surcharge, also known as climate factor, when is applied in a rain runoff model. To avoid an increased risk of infrastructure and flood systems damages, the Norwegian environment agency (Miljødirektoratet) recommend imposing a climate surcharge on the dimensioned values of precipitation or flood, particularly when planning long-term infrastructures. The climate surcharge denotes how much of the present dimensioned extreme values of rainfall and floods should be increased to consider future climate changes. For instance, the climate factor of 40% and 50% is equivalent to climate factor of 1.40 and 1.50 when the values are used in dimensioning of extreme rainfalls and floods. Table 2 below denotes a guideline for climate surcharges selection retrieved from Norwegian Climate Service center, and the values are related to dimensioned return period and the duration of an event. Rainfall with high intensity and return period need high climate surcharge and vice versa.

Table 2: Climate surcharge selection based on event duration and dimensioning period (Klimaservicesenter.no)

Event duration		***Dimensioning return period*** < 50	***Dimensioning return period*** ≥ 50
	$\leq$ 1 hours	40 %	50 %
	$>$ 1-3 hours	40 %	40 %
	>3-24 hours	30 %	30 %

The climate surcharge reflects the expected effect of climate change for a given century of high emissions of greenhouse gases. Their determination is independent of uncertainties, and dimensioning of short-term infrastructures does not need the inclusion of climate surcharges. It is recommended to use climate surcharges for precipitation and flood designed from IDF curves.

2 STATISTICAL INDEPENDENCE METHOD.

The Statistical independent method is based on extracting the annual exceedance series from a record of N number of years and implement the statistical analysis to determine their overlapping properties, and finally merged events which overlapped into single independent events. This chapter presents the overview of both the data which was being used and the description of the method.

2.1 Data and preliminary processing

Blindern precipitation station was commissioned in 1968, it has data record that spun back to 1968 up to date. The station has one of Norway's longest precipitations record with time resolution of one minute. The records consist of date and time at which a rain event occurred, and in addition the intensity of precipitation which stand for number of tipping on the gauge happened in a single minute (see figure 6a. below).

A single tip corresponds to 0.2 mm precipitation depth, which was later calibrated to quantity of 0.1 mm in year 2000. Even though this adjustment was made, it will never affect the rain intensity because intensity is a measurement of average precipitation in a specific duration. The highest one-minute's rain intensity of 4.1 mm was recorded in middle 1980, and since then the records has been below that value (see figure 5 below).

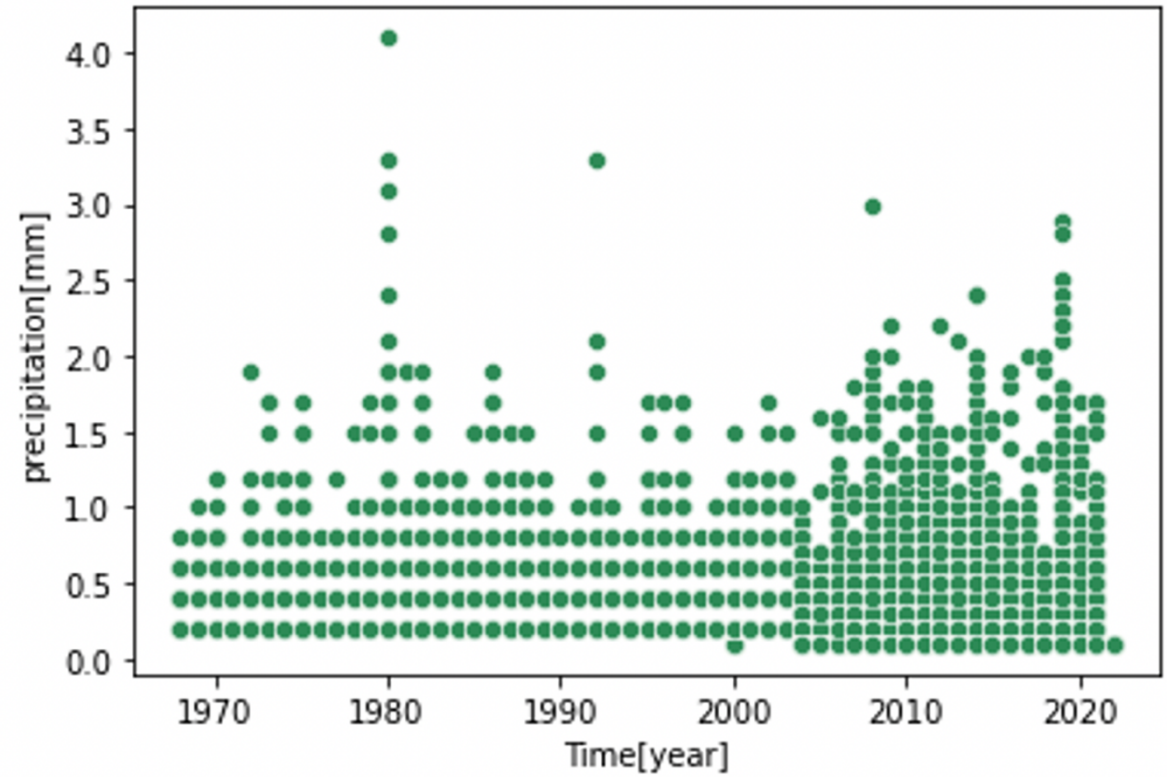


Figure 5. Rainfall records for Blindern precipitation station with 1 min time resolution and precipitation depth of 0.2 mm and 0.1 per tip before and after the year 2000 respectively.

The first procedure was to transform the original data (figure 6a) into a CSV file, in which the file was then imported into Python and by using Panda module the data was then transformed into a Panda data frame to facilitate further editing (see figure 6b below). Panda is a python module package which facilitate data analysis. From the figure (6b) one can observe that both registered and non-registered precipitation records were in overalls 554360 records. The zero records are non-registered records which stands for non-registration tipping of the gauge at that specific time. The next procedure was to remove all non-registered records (see figure 6c) for which one can observe from the figure that the overall number of records has been reduced from 554360 to 199108 after this stage. In addition, the time at which events happened had to be as well transformed into date-time format using datetime module in Python and placed them in index column (see figure 6c). The datetime module in Python provide the possibility to work in dates and time.

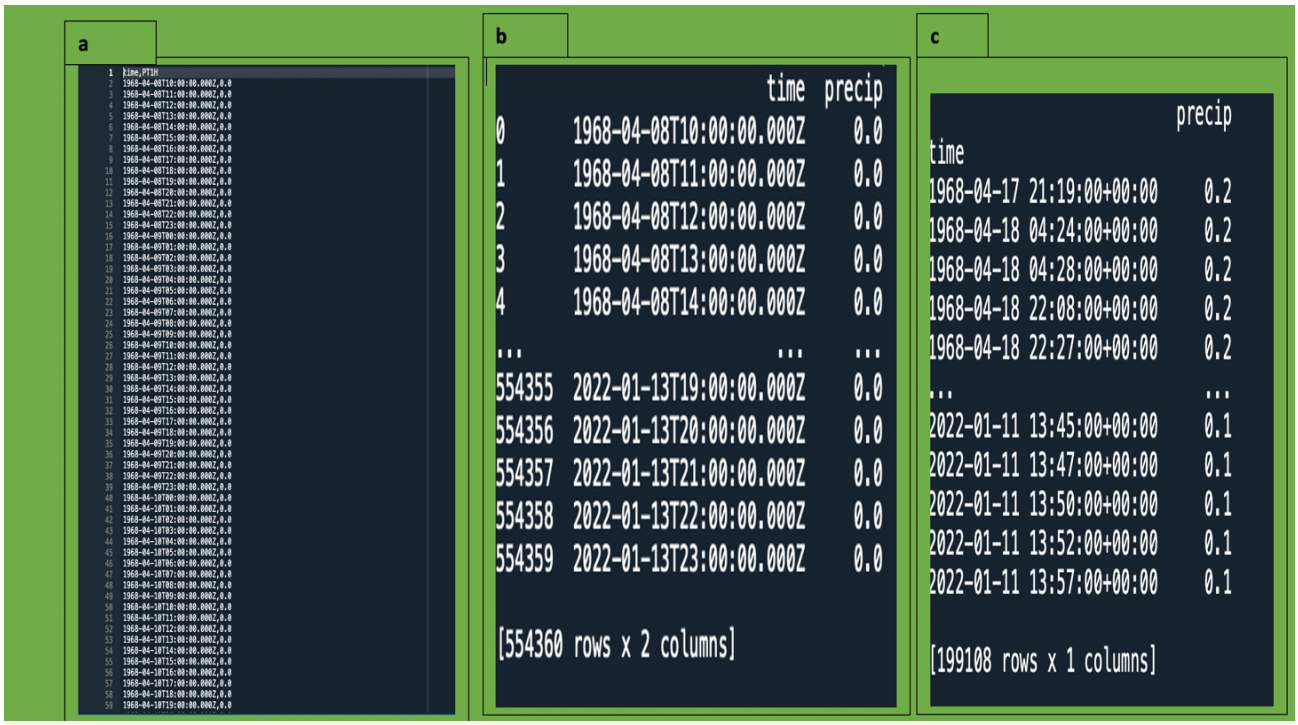


Figure 6: Data presentation of blindern precipitation station. Diagram (a) represent the original timeseries from the station with 1 min time resolution (b) Represent the output data from Panda data frame after logging from timeseries (a), (c) represent the data in (b) after the zero precipitation values has been removed and the date-time at which the events happened are placed in their own column as index.

2.2 Extraction of extreme values

Generally, the number of extreme values in annual exceedance series used in static analysis approach is always based on the length of timeseries in years, thus the number of extreme values should always be equivalent to the number of years in the timeseries (Donna Wilson 2011). In this situation, Blindern station had 54 years data records, based on this it was decided that each duration had to have 54 extreme values of precipitation events (see figure 7). The durations interval which was chosen to be worked on were from 1 minute to 1440 minutes duration, and from each of these durations, first 54 extreme precipitation events had to be extracted (see figure 7). All the 54 extremes precipitations events on each duration were extracted using rolling function in Python. The rolling function in Python provide the features of rolling window calculations. To simplify the calculation and workload, it was decided that the rolling function could only return the date-time at which extremes events occurred (ended), rather than precipitation values. The date-time at which those extreme events occurred were extracted by first tracking the 54 extreme values of each duration and find their corresponding date-time indexes and saved them for further use (see figure 6c).

	1min	2min	3min	4min	5min	...	1436min	1437min	1438min	1439min	1440min
0	4.1	7.4	10.5	13.3	15.7	...	72.8	72.8	72.8	72.8	72.8
1	3.3	5.4	6.8	8.9	10.3	...	64.7	64.7	64.7	64.7	64.7
2	3.0	5.1	6.6	8.7	10.1	...	62.2	62.2	62.2	62.2	62.2
3	2.9	4.8	6.6	8.3	9.8	...	58.8	58.8	58.8	58.8	58.8
4	2.8	4.6	6.3	7.6	8.8	...	57.2	57.2	57.2	57.2	57.2
5	2.8	4.3	5.7	7.4	8.2	...	57.2	57.2	57.2	57.2	57.2
6	2.5	4.2	5.4	7.0	8.0	...	56.7	56.7	56.7	56.8	56.8
7	2.4	4.0	5.3	6.9	8.0	...	56.1	56.1	56.1	56.1	56.1
8	2.4	3.8	5.3	6.8	7.7	...	55.4	55.4	55.4	55.4	55.4
9	2.4	3.7	5.3	6.5	7.5	...	55.3	55.3	55.3	55.3	55.3
10	2.3	3.6	5.1	6.5	7.4	...	55.0	55.0	55.0	55.0	55.0
11	2.2	3.6	5.0	6.1	7.0	...	51.2	51.2	51.2	51.2	51.2
12	2.2	3.6	5.0	6.1	7.0	...	50.1	50.1	50.1	50.1	50.1
13	2.2	3.6	5.0	6.0	6.7	...	49.3	49.3	49.3	49.3	49.3
14	2.1	3.6	5.0	6.0	6.6	...	48.6	48.6	48.6	48.6	48.6
15	2.1	3.6	4.9	5.9	6.6	...	48.3	48.3	48.3	48.3	48.3
16	2.1	3.5	4.8	5.7	6.6	...	46.8	46.8	46.8	46.8	46.9
17	2.0	3.5	4.7	5.6	6.6	...	46.2	46.2	46.2	46.2	46.2
18	2.0	3.4	4.7	5.6	6.4	...	45.9	45.9	45.9	45.9	45.9
19	2.0	3.4	4.6	5.5	6.3	...	45.9	45.9	45.9	45.9	45.9
20	2.0	3.4	4.5	5.5	6.2	...	45.4	45.4	45.4	45.4	45.4
21	2.0	3.4	4.4	5.5	6.2	...	45.0	45.1	45.1	45.2	45.3
22	1.9	3.3	4.4	5.4	6.1	...	44.3	44.3	44.3	44.3	44.3
23	1.9	3.3	4.4	5.3	6.0	...	44.0	44.0	44.0	44.0	44.2
24	1.9	3.3	4.4	5.3	6.0	...	43.7	43.8	43.8	43.9	43.9
25	1.9	3.3	4.4	5.3	5.9	...	43.7	43.7	43.7	43.7	43.7
26	1.9	3.3	4.3	5.1	5.9	...	42.5	42.5	42.5	42.5	42.5
27	1.9	3.3	4.3	5.1	5.9	...	41.0	41.0	41.0	41.0	41.0
28	1.9	3.2	4.3	5.0	5.8	...	41.0	41.0	41.0	41.0	41.0
29	1.9	3.2	4.2	5.0	5.8	...	40.6	40.7	40.7	40.8	40.8
30	1.9	3.2	4.2	5.0	5.7	...	40.4	40.4	40.4	40.4	40.4
31	1.9	3.1	4.2	5.0	5.7	...	40.2	40.2	40.2	40.2	40.2
32	1.9	3.1	4.2	4.9	5.7	...	39.8	39.8	39.8	39.8	39.8
33	1.9	3.1	4.2	4.9	5.7	...	39.4	39.4	39.4	39.4	39.4
34	1.8	3.1	4.1	4.9	5.6	...	39.3	39.3	39.3	39.3	39.3
35	1.8	3.1	4.1	4.8	5.5	...	39.3	39.3	39.3	39.3	39.3
36	1.8	3.1	4.1	4.8	5.5	...	38.6	38.6	38.6	38.8	38.8
37	1.8	3.0	4.0	4.8	5.5	...	38.3	38.3	38.3	38.3	38.3
38	1.8	3.0	4.0	4.7	5.4	...	38.2	38.2	38.2	38.2	38.2
39	1.8	3.0	3.9	4.7	5.3	...	37.8	37.8	37.8	37.8	37.8
40	1.8	3.0	3.9	4.7	5.3	...	37.6	37.6	37.6	37.6	37.6
41	1.7	3.0	3.9	4.6	5.2	...	37.4	37.4	37.4	37.4	37.4
42	1.7	2.9	3.9	4.6	5.2	...	37.4	37.4	37.4	37.4	37.4
43	1.7	2.9	3.9	4.5	5.1	...	37.2	37.3	37.3	37.3	37.3
44	1.7	2.9	3.8	4.5	5.1	...	36.8	36.8	37.0	37.0	37.0
45	1.7	2.9	3.7	4.5	5.1	...	36.8	36.8	36.8	36.8	36.8
46	1.7	2.9	3.7	4.5	5.1	...	36.8	36.8	36.8	36.8	36.8
47	1.7	2.8	3.7	4.5	5.1	...	36.8	36.8	36.8	36.8	36.8
48	1.7	2.7	3.7	4.4	5.0	...	36.7	36.7	36.7	36.7	36.7
49	1.7	2.7	3.7	4.4	5.0	...	36.5	36.5	36.5	36.5	36.5
50	1.7	2.7	3.6	4.4	5.0	...	36.4	36.5	36.5	36.5	36.5
51	1.7	2.7	3.5	4.4	5.0	...	36.3	36.3	36.3	36.3	36.3
52	1.7	2.7	3.5	4.4	4.9	...	36.3	36.3	36.3	36.3	36.3
53	1.7	2.7	3.5	4.4	4.9	...	36.2	36.2	36.2	36.2	36.2

[54 rows x 1440 columns]

Figure 7 : Annual exceedance series for each duration from 1min to 1440 min extracted from rolling function in Python. The row on top stands for event's durations and the columns for precipitation depth arranged in ascending order.

In addition to the above description of extracting the extremes precipitation values, there was also a need of establishing a spacing function. The spacing function was used to determine the minimum time space between a selected extreme value of a given event and the rest of rolled the events. As mentioned earlier, spacing function could allow the possibility of varying the selection of independent

extreme values into stricter condition of not overlapping by increasing time spaces between them and the rest of events. The variation was being achieved during the extraction of extreme values by which all time intervals which overlapped with time interval at which an extreme event occurred plus and minus the size of spacing were assigned a zero precipitation values, so that those zero events wouldn't be considered during the extraction of the proceeding extreme value in the same rolled time window calculations. For instance, consider one is extracting first 54 extreme values (in descending order) of five minute duration by rolling window calculations, and the spacing value is assigned a value of two minutes .When the first or highest extreme value is extracted, then all rolled time window events which overlaps with that extreme value interval plus-minus the size of spacing ($5 \mp 2min$) are then assigned zero precipitation values so that they wouldn't be considered when extracting second extreme value on the same duration, and so forth until the 54 independent extreme values are selected . Consider the above description of the spacing function, one can expect that events will overlap frequently when the spacing has been increased**.** It is also important to emphasize that the spacing function work on extremes values selection, thus when the highest extreme value of a given time window has been extracted, then the spacing function assign zero values to all other values which have time intervals that overlaps with the time interval of the selected highest extreme value. In this way, the function secures 54 independent extreme values selection on each duration.

2.3 Merging overlapping events

The third part of statistical independent method involved sorting extreme event durations which were extracted in section 2.2 and merged them into single independent events. This was achieved by extracting an initial interval after sorting process, and every interval (in chronological order) was being verified against the previous one, starting with the initial one. Furthermore "if expression" normally known as Boolean expression was used to check if the intervals overlapped or not. If two intervals overlapped, then they were merged and saved as a single interval, in addition the initial interval was as well substituted by the merged interval, if they didn't overlap then the new interval was saved a single independent interval. The procedure above and spacing value of zero minutes could produce 159 independent merged intervals shown in figure 8, and the dark area represent their corresponding lengths in minutes arranged in ascending order, with shortest events having a length of one minute and longest having 2524 minutes duration. The codes in appendix 2 represent the methods described in section 2.1-2.3 of this thesis.

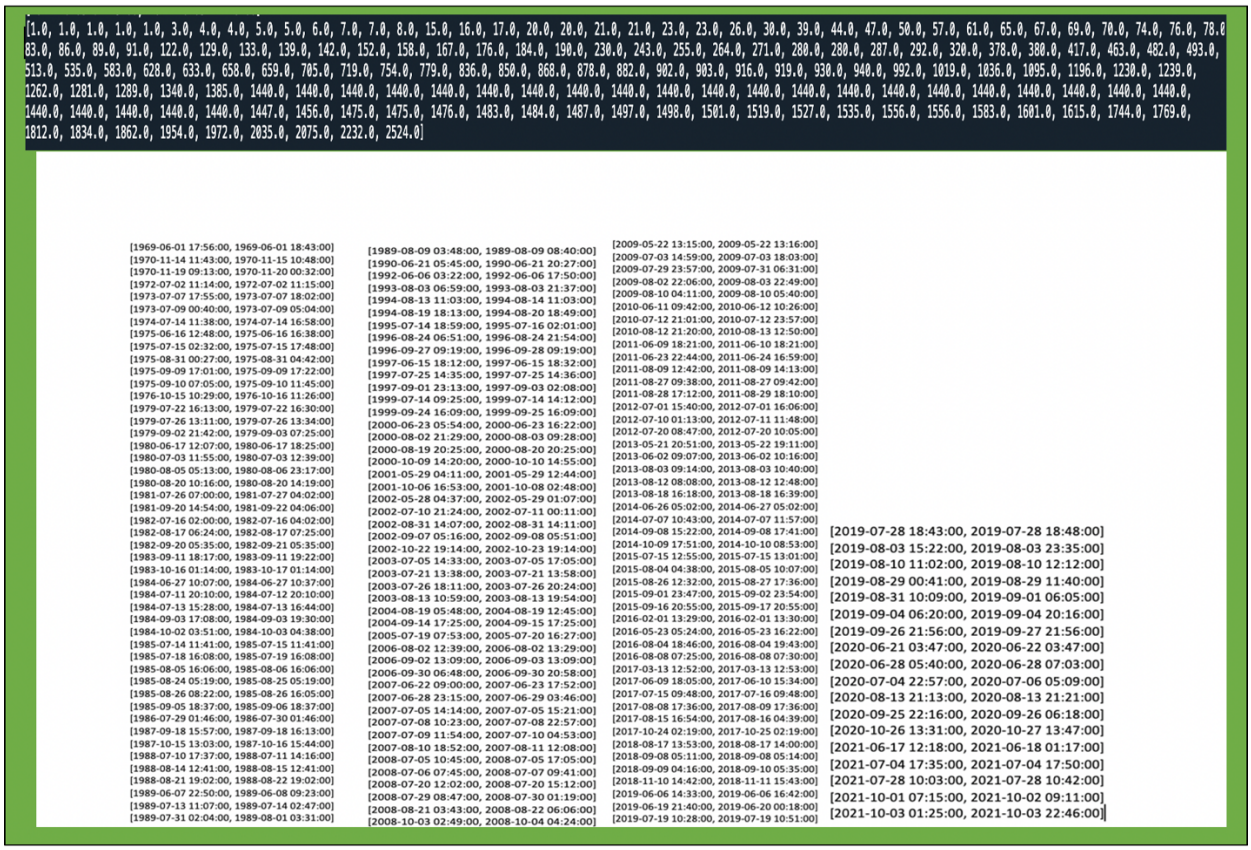


Figure 8: Time intervals for 159 Independent merged events obtained from the analysis with their corresponding length in minutes on top.

2.4 Analysis of the number of non-overlapping events

Analysis of the data obtained by the methods described in section 2.2 and 2.3 is presented in this chapter. To map the extend at which extreme events of given duration overlap with extreme events of all other durations (1-1440min), each extreme interval of that given duration was verified against all extreme intervals of other durations. For instance, if the overlapping probability of 5-minutes intervals were to be verified against all other durations events, then all 54 intervals of 5 minutes duration had to be verified against all 54 durations of 1-1440 duration and return the overall probability values on each duration of the 1-1440 min duration (see appendix 3). The result of this analysis was visualized in figure 10 in the result section. The second non-overlapping event's analysis was to verify the cumulative number of merged independent events obtained when the maximum number of event's duration was varied. This was achieved by varying the maximum number of durations for the analysis described in section 2.2 and 2.3 and saving the cumulative number of merged independent event returned by each duration (1-1440 min), and finally visualized this result as shown in figure 14 in the result section of the thesis. The analysis was well extended to the maximum duration of 2000 minutes as shown in the subplot of figure 14.

2.5 Illustration of non-overlapping independent events in the IDF diagram

The next analysis involved incorporating the independent merged events into the original Blindern's IDF diagram. This was achieved by using the data frame from the original data series (figure 6b) and resample them into one-minute intervals. The resampling function in Panda filled the missing (or unregistered) precipitation values with zero values by default, as result a complete data frame with a complete date-time and their corresponding precipitation values was created. Finally, merged independent time intervals were used to extract their own corresponding data fame from the complete data frame which was created from resampling function. Each of data frame of the independence merged was rolled into their own maximum window, finally used to visualize their own propagation in the original Blindern's IDF diagram as shown in figure 15. Codes used to generate this analysis are shown in appendix 4. The diagram created in figure was not that presentable, therefore there was a need taking another approach, and this led to the consideration the final stage described below.

The final analysis involved incorporating the events with highest return period from each of the merged independent events into the original Blindern's IDF diagram as shown in figure 16. This was achieved by rolling each of the data frame obtained from independent merged event to their own maximum windows, which were their respective time lengths presented in figure 8. The time intervals and their corresponding precipitation values obtained from rolling of independent merged intervals data frames, were then passed into a linear interpolation function which was established from Blindern's IDF data in Python shown in appendix 1. The linear interpolation function in Python is an established function based on data incorporated in it (in this case the data was Blindern's IDF data shown in appendix 1), it then establishes a function which can be used for further interpolation of other data of the same type as input data. The linear interpolation function calculated return periods of the rolled events within each of the independent merged event, and from this the highest return period event was then retrieved and visualized in the original Blidern IDF diagram as shown in figure 16. Codes used to generate this analysis are shown in appendix 5.

3 DISCUSSION AND RESULTS

The chapter presents the results obtained from the statistical independence method described in section 2.1 – 2.5. These results include visualizing extreme values (see figure 7) from the first stage analysis described in section 2.2, visualizing the overlapping probability of specific events, these include low duration events, average duration, and high duration events (see figure 10), graphs of cumulative independent merged events as function of spacing and maximum event's durations as shown in figure 13 and 14 respectively. Finally, the incorporation of events with highest return periods from independent merged events into the original Blindern's IDF diagram as shown in figure 15 and 16.

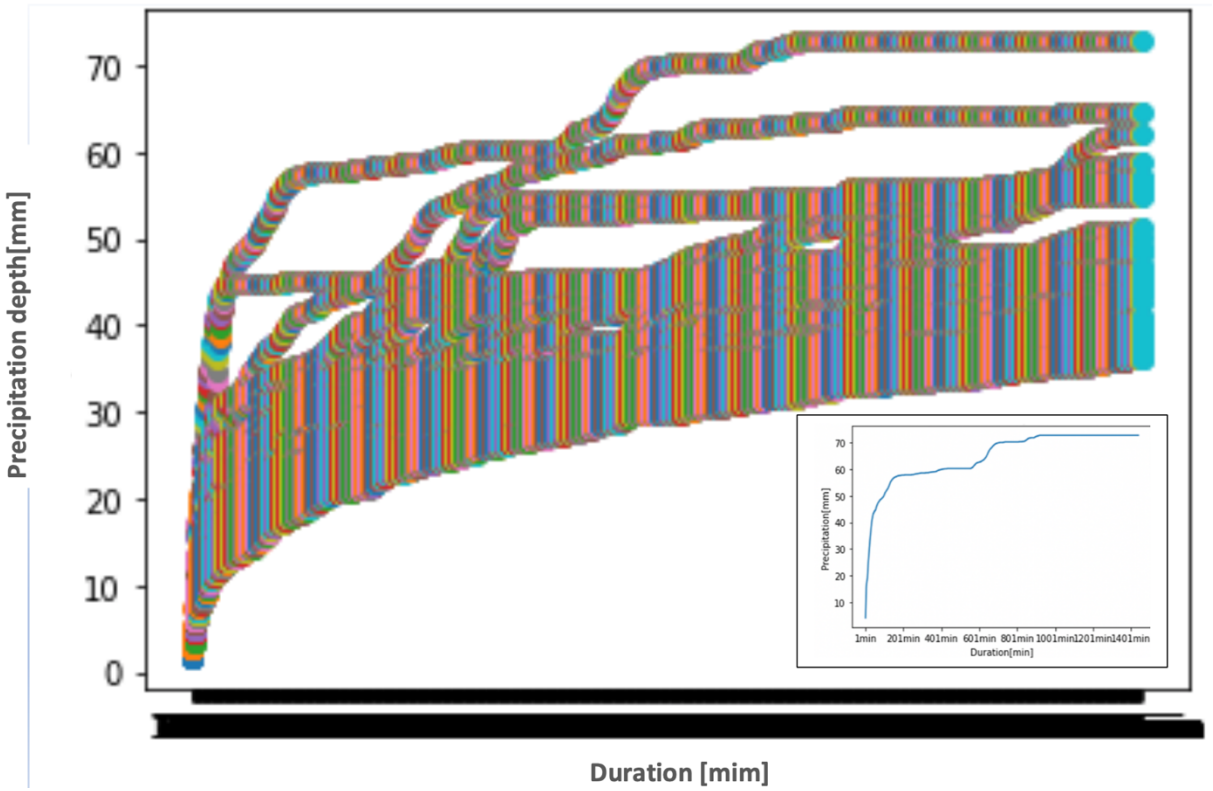


Figure 9: Scattered plot of first 54 extreme for different durations represented in the subplot. Both graphs were obtained from the method described section 2.2 and with zero spacing value.

The diagram above shows the scattered plot of the durations, and their corresponding extreme precipitation depths. The figure shows that most of the events are concentrated below the precipitation depth of 45 mm, and above that there are lack of records in certain areas of the plot. One can conclude that cumulative precipitation above 45 mm in a specific duration rarely happens, this phenomenon is independent of duration. There is also extreme values variation during the first 200 minutes duration shown in the subplot which present the highest extreme values on each duration. Most extreme values variation with duration happens in events with low duration, and the overall gradient value of the diagram (and subplot) tries to stabilize toward a zero value as the duration increases, and this shows that at high duration events, the extreme values tend to stabilize toward a constant value. This event characteristic was as well observed in graphs of independent merged events as function of both spacing function and maximum duration as show in figure 13 and 14 respectively.

The diagram of the second analysis is shown in figure 10 below, with the main diagram representing high duration's overlapping probabilities, and the subplot representing low durations and mean duration's overlapping probabilities. From the subplot one can observe that lowest duration events appear to have low ability to overlap with events which has lower duration their own durations, in contrast they happen to overlap with events with higher durations than their own. In low events one can observe that lines which stand for overlapping probability appears to be non-continuous on the lower duration side of their own durations. When event's durations were increased to find their ability to overlap with other events, the lines became more and more continuous, signifying that, events overlap frequently with other events, with e.g., events of duration above an hour having far better continuous overlapping probability lines than for instance five-minute duration. An overall observation is that events with low duration has higher probability of overlapping with higher duration than their

own durations, but their ability to overlap with higher duration event decreases over a large interval as the duration of the events they supposed to overlap with increases. These phenomena lead to low chances of finding low duration rain event in a very large duration precipitation event, and this trend try to be reciprocated in a situation of high duration precipitations events, for which they tend always to overlap with low duration precipitations. Their ability to overlap with low duration events keep increasing over a large interval until it reaches their own durations, that is where there is a slightly drop in their overlapping ability showing the overlapping probability start to drop (see diagram 9 below).

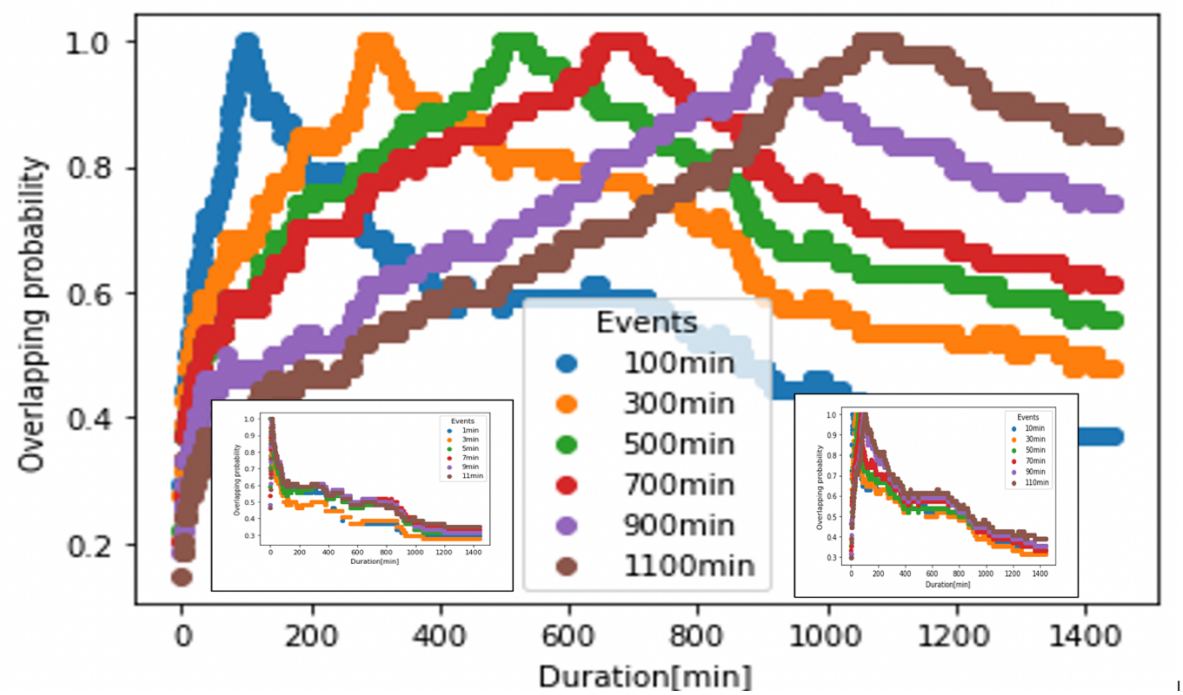


Figure 10: Graphical representation of overlapping probabilities of extreme values intervals of specific durations against the extreme values intervals of all other durations(1-1440min). The diagrams were obtained from the method described in section 2.2 and 2.4.

Event's overlapping probabilities which have been visualized in figure 10 can be used to compute a summative EAD of different precipitation durations using equation (1.2). It was mentioned that equation 1.2 consists of several terms of EAD of different durations, and the alpha symbols in the equation represents the contribution weight of each duration, in such a way that, if events were totally independent (not overlapping), therefore all alpha values will be one, and if they were totally dependent (complete overlap), all alphas would be zero except for one duration, thus the equation would have only a single term, and finally if they partially overlap, then alpha values would be represented by the degree of overlapping. The last statement can be completed by the analysis done in this thesis, thus the alpha values in this regard are event's abilities to overlap with other events. For instance, consider a 30 min event's probability values to overlap with other events as shown in figure 11. According to the diagram the event has 100% chances of overlapping with 30 min event, and 85% with 200 min events. Therefore, a summative EAD of these two durations, thus 30 min duration and 200 min can be computed by considering two terms in equation (1.2), for which 100% stands for the

alpha value in 30 min term, and 15% (100%-85%) stands for the alpha value in 200 minutes term. The reason of considering 15% instead of 85% as alpha value in 200 min term is because of securing independent relationship between the 30 min and 200 min term, and as result of this, the result obtained in this thesis couldn't be used to compute the EAD of three durations (terms). This is because the independence information of the third term in relation to the two other terms couldn't be retrieved in the constructed probability diagrams and this problem is beyond the scope of this thesis.

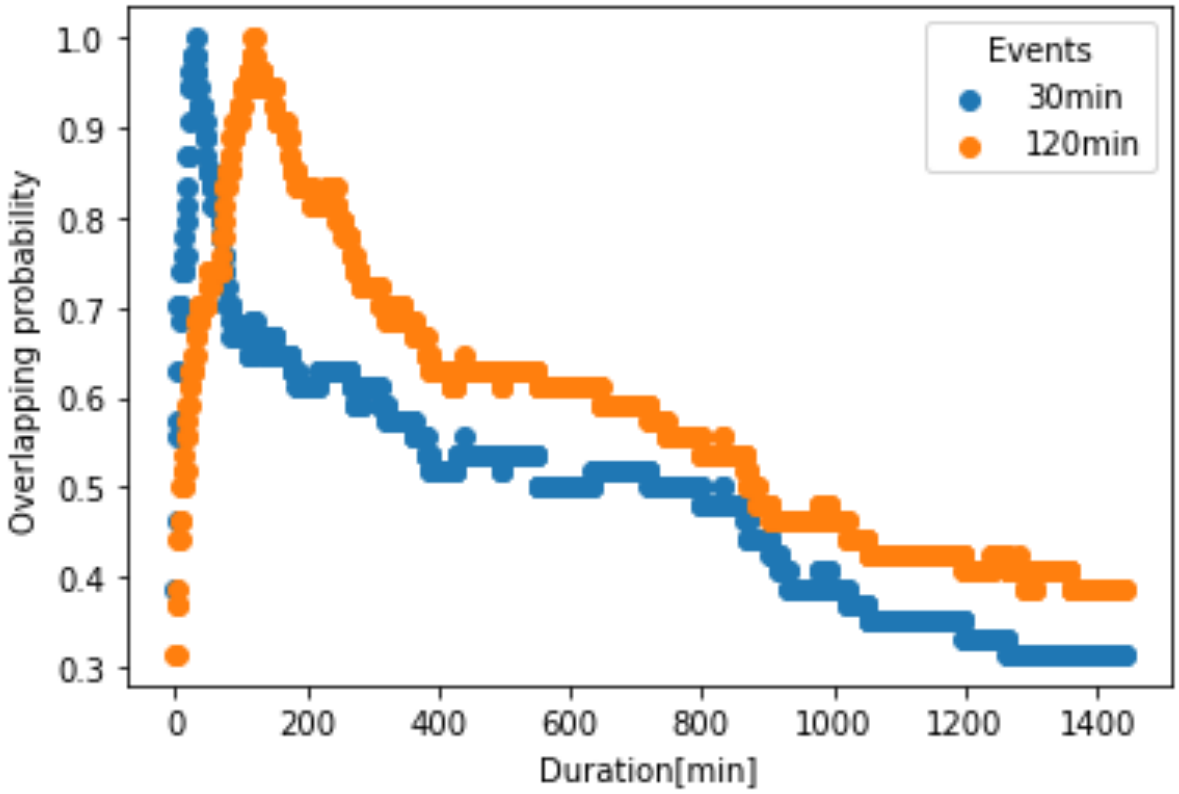


Figure 11: Graphical representation of overlapping probabilities of extreme values intervals of 30 & 120min duration against the extreme values intervals of all other durations(1-1440min). The graphs were obtained from the method described in section 2.2 and 2.4.

Considering the equation (1.2) and the explanation of computing the summative EAD for different durations above, one can conclude that non-consideration of event's overlapping probabilities will lead to either overestimation of EAD computation in a situation where events are regarded to be independent of each other, or underestimation of EAD computation in situation where events are regarded as fully dependent of each other. Therefore, the procedure of considering event's abilities to overlap is very essential in determining the accurate EAD calculations.

Spacing function was implemented to secure independent selection of extreme values described in section 2.2. The independent selection was achieved by assigning a zero value to the events which overlapped with selected extreme event, in such way that the assigned zero value events wouldn't be considered during the selection of the next extreme value on the same event duration. To understand the effect of this function better, diagram 9 which had a spacing value of zero minute, had to be plotted with higher spacing values, therefore diagram 12a,1b and 12c denotes the same diagram of extreme values with spacing values of 100, 1000 and 10000 minutes respectively.

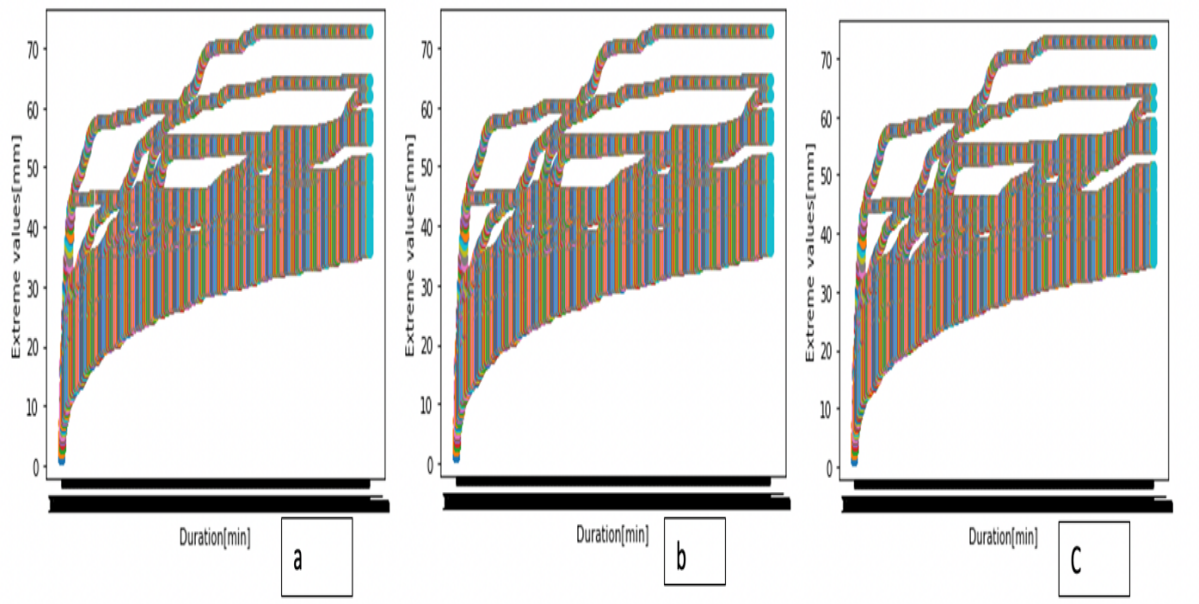


Figure 12: Scattered plot for first 54 extreme for different durations and spacings. With 12a, 12b and 12c representing spacing value 100, 1000 and 10000 minutes respectively.

The results in 12a and 12b with spacing of 100 and 1000 minutes respectively, shows little difference between the extreme value's scattered plot for spacing of 100 and 1000. There gap's sizes between extreme values increased in diagram 12b compared to diagram 12a, but the increment in gap sizes was not that significant, to achieve significant gap sizes, the spacing size had to be increased to 10000 minutes as shown in c. One can observe that gap became more significant compared to what was achievable in b. However, the three diagrams displayed almost the same data distribution despite the variation of spacing values. Particularly on low duration's extreme values, and the results obtains here substantiate on the statement that low duration events are more independent of each other than large durations. High variation of the spacing values with little variation on extreme values distribution obtained in these diagrams indicates that, the extreme values are very independent of each and happens far form each other. This result was again displayed when plotting the cumulative number of independent merged events obtained by the analysis described 2.3 as a function spacing. The result displayed very little variation (at low spacing value) on the cumulative number of independent merged events, which indicates that more overlapping of precipitation events with extreme values lead to a slightly increment on number of cumulative independent events as shown in figure 13. In overall, figure 12 shows little effect that the spacing function had on the distribution of independent extreme values, and this effect propagated itself to the few numbers of independent merged event obtained when varying the spacing.

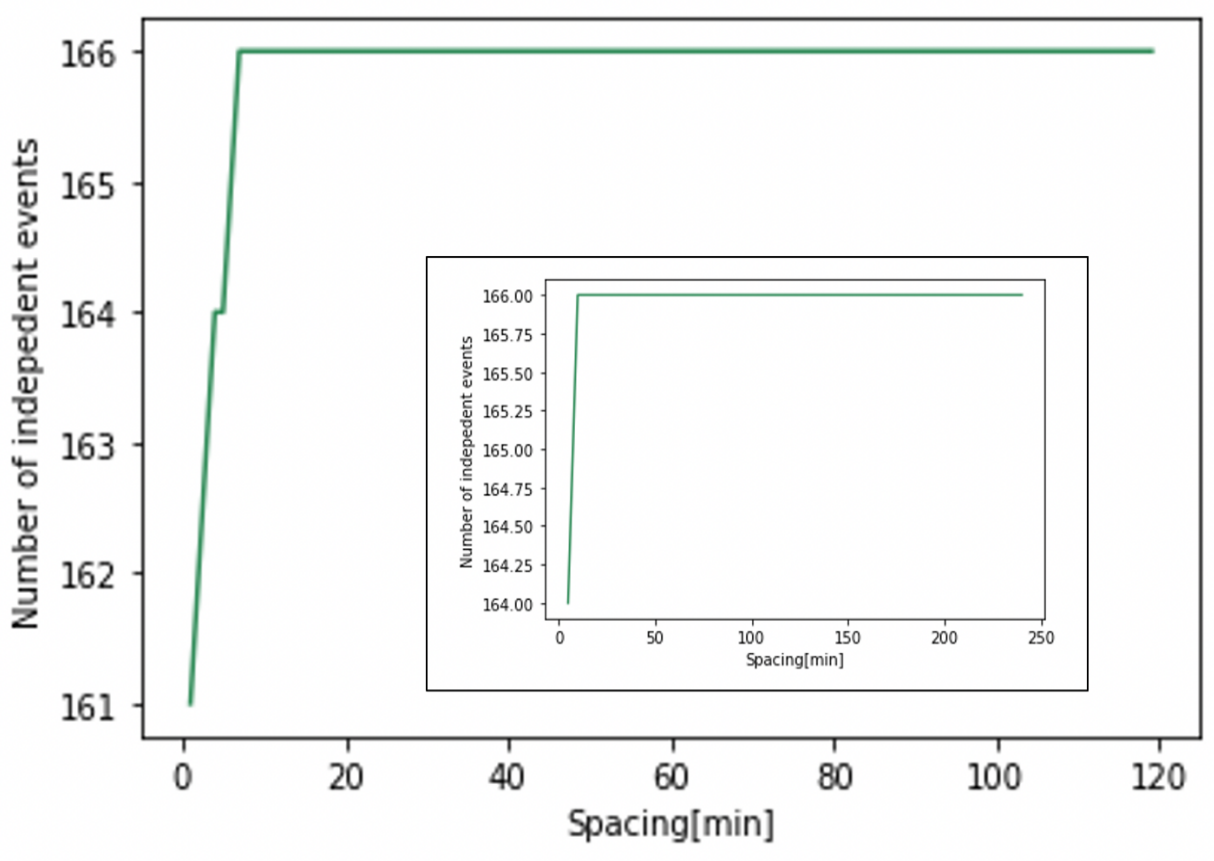


Figure 13: Graph for number of independent merged events as a function of spacing obtained from the method described in section 2.2. The subplot represents the graph extracted from extending the duration from 1440 to 2000 minutes.

Number of independent merged events obtained was again visualized as a function of maximum duration as shown in figure 14. The result showed that most independent merged events were obtained on low maximum duration, and as the maximum duration was increased the gradient approached zero, showing that the number of independent merged events stabilized itself at 160 events. For further investigation, the length of maximum duration was increased beyond what was sat for this analysis, thus from 1440 to 2000 minutes (see subplot figure 14), however the stabilization for the number of independent events at 160 became more and more obvious. The variation of the number of independent merged events obtained on low duration in this regard represent the frequency at which the number of independent events occurred with event's duration. In this regard one could observe that most of independent events happened in low duration events, and the results obtained here correlated with the results obtained from previous analyses.

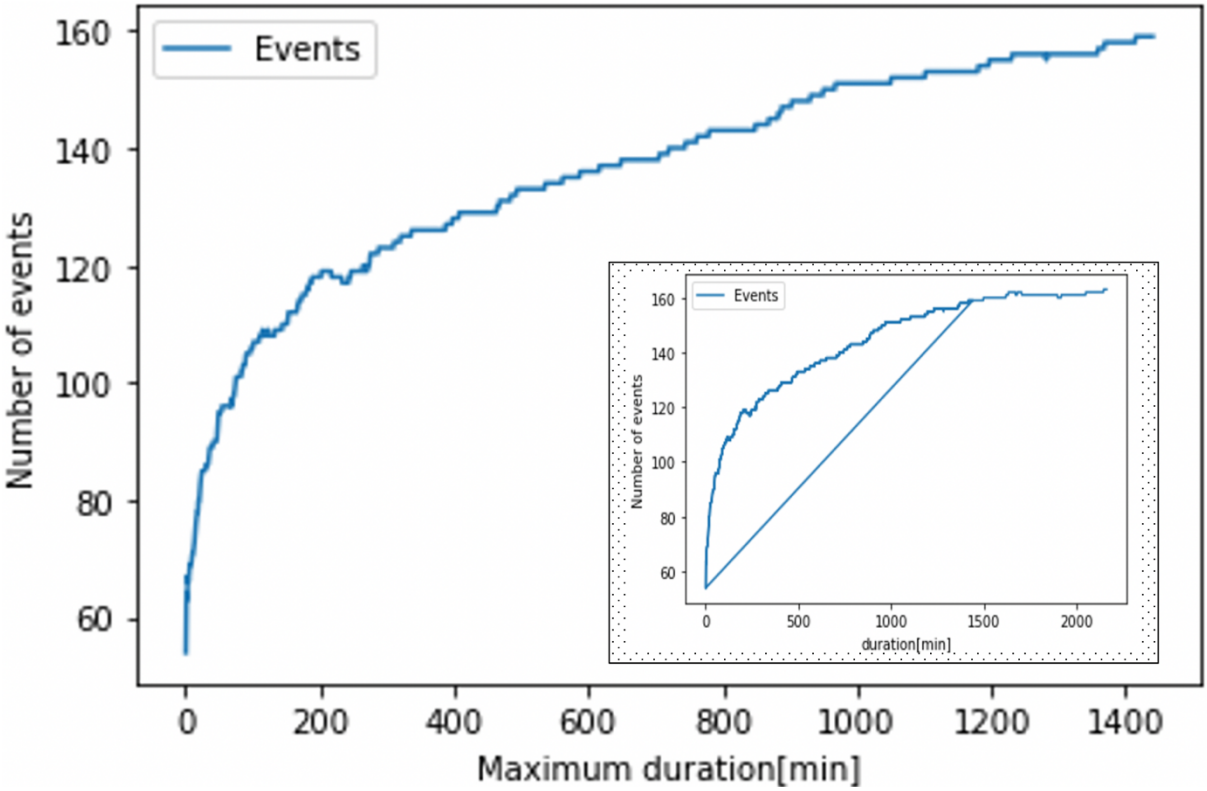


Figure 14: Cumulative independent events as a function of maximum event duration obtained from the method described in section 2.4 of the thesis. The subplot represents of the graph extracted from extending the duration from 1440 to 2000 minutes.

To investigate the result obtained in figure 14 futhere, that most of independent events happened in low duration events, there was a need of visualizing the 159 merged independent events into the original Blindern's IDF curves as shown in figure 15, by the method described in section 2.5. This was to observe how these events propagated on the IDF diagram. One could observe that most those events had precipitation depth below 45 mm, which correlated with the results obtained from extreme values analysis in figure 9. The IDF diagram with all 159 merged independent events in it was an inappropriate way of visualizing data, thus the best way was to find events with highest return periods in each of merged independent event and present them in the IDF diagram instead (see figure 16). The method to obtain the data for this diagram was described in section 2.5.

.

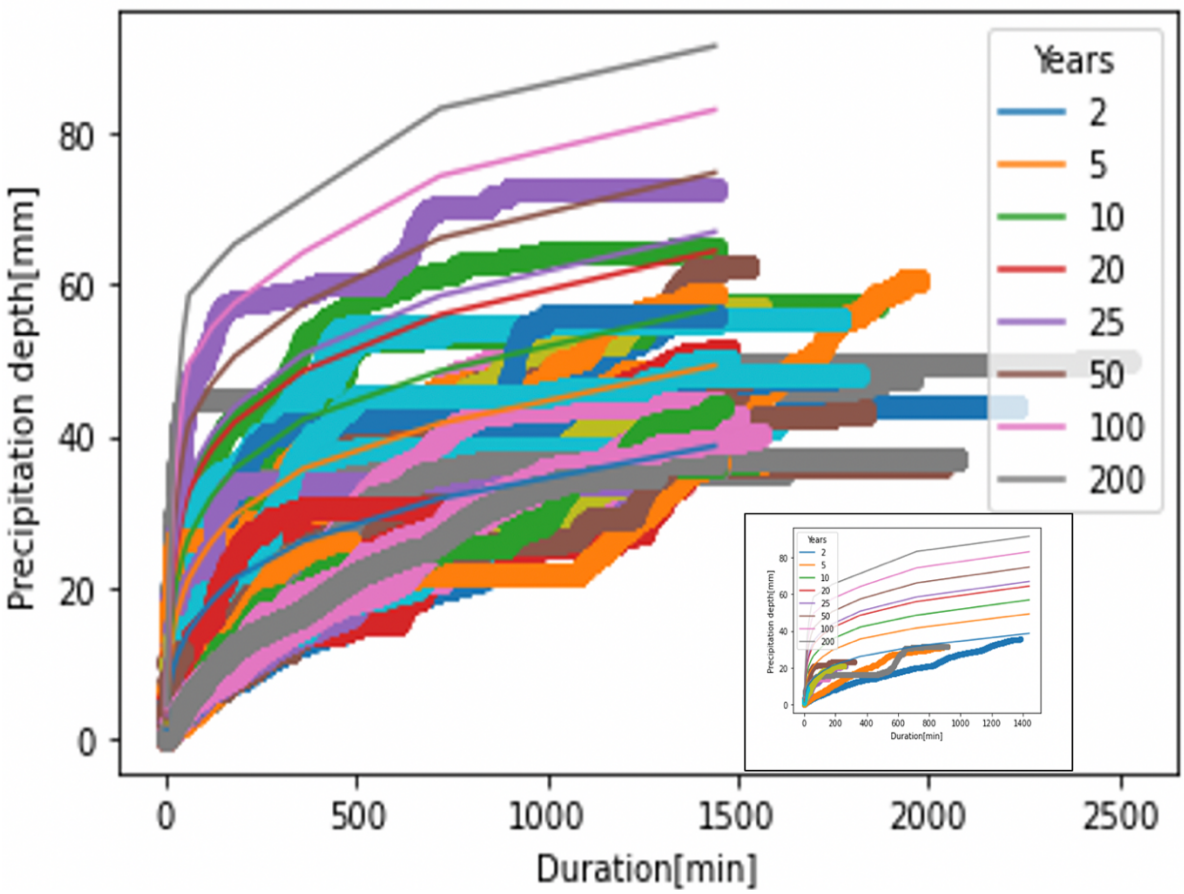


Figure 15: The diagram of showing the propagation of independent merged event in the original Blindern's IDF curves with zero spacing size. The subplot represents the propagation of the first ten independent merged events. The diagram was extracted from the method described in 2.5.

From the diagram one could observe that most of the merged independent events were concentrated on low duration, which correlated with the result obtained from previous analysis. The other observation was that the independent merged events were concentrated around two-year return period, and when the return period increased, then these events happened to occur rarely and are found individually. The result showed that these independent events propagated mostly below 20 years return period (45 mm precipitation depth), which correlated with the result obtained from previous analyses presented in figure 9 and 15.

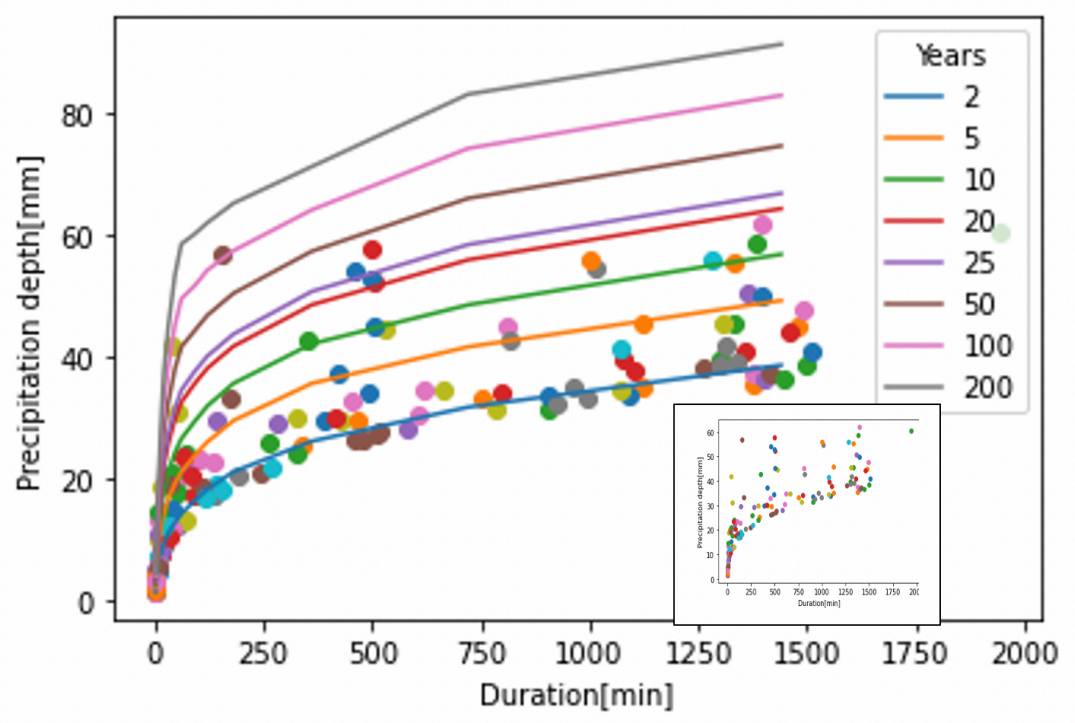


Figure 16: The diagram of showing the distribution of highest return period of the independent merged event in the original Blindern's IDF curves with zero spacing size. The subplot represents the distribution of highest return period of the independent merged event alone. The diagram was extracted from the method described in 2.5.

5 CONCLUSIONS AND RECOMMENDATIONS

- The result obtained from analyzing the statistical dependence (overlapping probability) of events with different durations shows that, events with low duration has higher probability of overlapping with events of higher duration than their own durations, but their ability to overlap with higher duration events decreases over a large interval as the duration of higher duration events increases. But the reciprocal happened with events of higher durations, that they have higher probability of overlapping with events which have lower duration than their own durations, and their ability to overlap with low duration events increases over a large interval as the duration of low durations events increases. I addition their ability to overlapping with higher duration events than their own starts at high value and drop as the duration increases.
- It was observed that the process of mapping event's overlapping probabilities could enhance the computation of EAD, thus one can conclude that non-consideration of event's overlapping probabilities such calculations can result to either overestimation of EAD computation in a situation where events are regarded to be independent of each other or underestimation of EAD computation in situation where events are regarded as fully dependent of each other. Therefore, the procedure of considering event's abilities to overlap is very essential in determining the accurate EAD calculations.

- It was also observed that extreme events happened far from each other, as result the spacing function which secured the independent selection of extreme value had little effect on the distribution of these values. This tendency propagated itself to the result obtained on the number of independent merged events when the spacing values were varied.
- It was also observed that precipitation events have abilities to overlap across durations, and independent events are mostly found in short duration events, and most of these independent events were distributed around two years return period.

With the results achieved in this thesis, there is still a need of describing or sum up event's overlapping properties statistically. This will open the door for mapping statical dependence of three or more precipitation events with different durations, therefore facilitating the use of more than two terms of the summative EAD equation and improving the computation of EAD beyond the scope of this thesis. In addition, there is a need of putting the event's overlapping properties into application in computing the EAD, and apply the principle of future mapping, including the consideration of climate effect. They can be also a need of applying the spacing function in construction of IDF curves to secure independent selection of extreme values instead of the traditional methods which are based on both previous and current method mentioned in section 1.4.5. The inconsistencies between durations obtain by using the spacing function could be compared to the one obtains by the traditional methods and see if there a possibility of reducing them to the minimum value.

6 REFERENCES

Anita Verpe Dyrrdal, J. L., Lars Grinde (2022). "METreport IVF-verdier for norske nedbørstasjoner." **2/2022**.

Arnbjerg-Nielsen, K. (2011). "Past, present, and future design of urban drainage systems with focus on Danish experiences." Water Science and Technology **63**(3): 527-535.

Arnbjerg-Nielsen, K. and H. S. Fleischer (2009). "Feasible adaptation strategies for increased risk of flooding in cities due to climate change." Water Sci Technol **60**(2): 273-281.

Bernad, M. M. (1932). "Formulas for rainfall intensities of long duration " ASCE **96**: 592–624.

Berndtsson, R. and J. Niemczynowicz (1988). "Spatial and temporal scales in rainfall analysis—Some aspects and future perspectives." Journal of Hydrology **100**(1-3): 293-313.

Chin, D. A., et al. (2000). Water-resources engineering, Prentice Hall Englewood Cliffs.

Coles, S., et al. (2001). An introduction to statistical modeling of extreme values, Springer.

Donna Wilson, A. K. F., Deborah Lawrence, Hege Hisdal, Lars-Evan Pettersson, Erik Holmqvist (2011). A Review of NVE's Flood Frequency Estimation Procedures. Norwegian Water Resources and Energy Directorate, Middelthunsgata 29. P.O.Box 5091 Majorstua, N 0301 Oslo, Norway, Norwegian Water Resources and Energy Directorate. **:** 50.

Gumbel, E. J. (1954). Statistical theory of extreme values and some practical applications: a series of lectures, US Government Printing Office.

Gumbel, E. J. (1958). Statistics of extremes. Statistics of Extremes, Columbia university press.

Gyasi-Agyei, Y. (2005). "Stochastic disaggregation of daily rainfall into one-hour time scale." Journal of Hydrology **309**(1-4): 178-190.

Habib, E. H., et al. (2008). "Effect of local errors of tipping-bucket rain gauges on rainfall-runoff simulations." Journal of hydrologic engineering **13**(6): 488-496.

He, L. and G. V. Wilkerson (2011). "Improved Bankfull Channel Geometry Prediction Using Two-Year Return-Period Discharge 1." JAWRA Journal of the American Water Resources Association **47**(6): 1298-1316.

Hendriks, M. R. (2010). Introduction to physical hydrology. Oxford, Oxford University Press.

Keefer, T. O., et al. (2008). "An event based comparison of two types of automateÑ recording, weighing bucket rain gauges." Water Resources Research **44**.

Koutsoyiannis, D. (2004). "Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations." Hydrological Sciences Journal **49**(4).

Lutz, J., et al. (2020). "Estimating Rainfall Design Values for the City of Oslo, Norway—Comparison of Methods and Quantification of Uncertainty." Water **12**(6): 1735.

Madsen, H., et al. (2009). "Update of regional intensity–duration–frequency curves in Denmark: tendency towards increased storm intensities." Atmospheric Research **92**(3): 343-349.

Marsalek, J. (1981). "Calibration of the tipping-bucket raingage." Journal of Hydrology **53**(3-4): 343-354.

McCuen, R. and E. Beighley (2003). "Seasonal flow frequency analysis." Journal of Hydrology **279**: 43-56.

Olsen, A. S., et al. (2015). "Comparing Methods of Calculating Expected Annual Damage in Urban Pluvial Flood Risk Assessments." Water **7**(1): 255-270.

Papalexiou, S. M. and D. Koutsoyiannis (2013). "Battle of extreme value distributions: A global survey on extreme daily rainfall." Water Resources Research **49**(1): 187-201.

Rahimi, A., et al. (2003). "Use of dual-frequency microwave links for measuring path-averaged rainfall." Journal of Geophysical Research: Atmospheres **108**(D15).

Roksvåg, T., et al. (2021). "Consistent intensity-duration-frequency curves by post-processing of estimated Bayesian posterior quantiles." Journal of Hydrology **603**: 127000.

Rosbjerg, D. (2017). "Optimal adaptation to extreme rainfalls in current and future climate." Water Resources Research **53**(1): 535-543.

Vegvesen, S. (2018). "Vegbygging Håndbok N200

https://fileserver.motocross.io/trafikksiden/HB_N200_Vegbygging_2018.pdf."

7 APPENDIXES

Appendix 1: Data used to extract Blindern IDF curves for 24 hr by the METmethod described in section 1.4.5.2

	1	2	3	5	10	15	20	30	45	60	90	120	180	360	720	1440
2	1.5	2.6	3.5	4.8	7	8.3	9.5	11.1	12.8	14.3	16.4	18.3	21.2	26.2	31.8	38.7
5	2.1	3.7	5	6.9	10.3	12.6	14.4	16.7	19.2	21.2	23.8	26.1	29.6	35.7	41.7	49.3
10	2.6	4.4	6	8.4	12.7	15.7	18.1	21.1	24.3	26.6	29.4	32	35.7	42.2	48.6	56.9
20	3	5.2	7	10	15.1	18.9	22.1	25.8	29.7	32.6	35.6	38.2	41.8	48.6	56	64.4
25	3.1	5.5	7.4	10.5	15.9	19.9	23.4	27.4	31.5	34.6	37.6	40.2	43.8	50.8	58.5	66.9
50	3.6	6.2	8.5	12.2	18.4	23.5	27.8	32.6	38	41.7	44.3	46.9	50.5	57.4	66.1	74.7
100	4	7.1	9.6	13.9	21	27.4	32.6	38.4	45.2	49.6	51.7	54.3	57.5	64.2	74.3	83
200	4.5	7.9	10.7	15.7	23.7	31.6	37.8	45	53.4	58.6	60.2	62.1	65.3	71.3	83.2	91.4

Appendix 2. Codes for computing merged independent events described in section (section 2.1-2.3)

```
>>> import pandas as pd
import numpy as np
import datetime
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
from scipy.interpolate import interp1d

df = pd.read_csv("blindern_130122.csv", header=0, names=("time", "precip"))
df = df[df.precip > 0]
df["time"] = pd.to_datetime(df["time"])
df = df.set_index("time")
df.sort_index()

def find_events(duration, n_events, df, spacing= 0):
    sums = df.rolling(f"{duration}min").sum()

    dt = datetime.timedelta(minutes=duration)
    dts = datetime.timedelta(minutes=duration+spacing)
    res = []
    for i in range(n_events):
        t = sums.precip.idxmax()
        res.append(pd.Interval(t-dt, t, closed='both'))

        sums.precip[t-dts:t+dts] = 0.0

    return res

def merge_overlapping_intervals(a, b):
    return pd.Interval(min(a.left, b.left), max(a.right, b.right), closed='both')

def merge_list_of_intervals(a):
    c = [a[0]]
    for iv in a:
        if pd.Interval.overlaps(iv, c[-1]):
            c[-1] = merge_overlapping_intervals(iv, c[-1])
        else:
            c.append(iv)
    return c

eventdict = {i: find_events(i, 54, df)  for i in range(1,1441)}
all_intervals = []

for duration in range(1,1441):
    all_intervals.extend(eventdict[duration]) ?

a = merge_list_of_intervals(all_intervals)
```

Appendix 3: Codes for extraction of events overlapping probabilities.

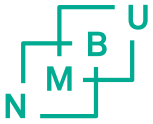
```
new_dict = {1:[1.5,2.1,2.6,3,3.1,3.6,4,4.5],2:[2.6,3.7,4.4,5.2,5.5,6.2,7.1,7.9],3:[3.5,5,6,7,7.4,8.5,9.6,10.7],5:[4.8,6.9,8.4,10,10.5,12.2,13.9,15.7],10:[7,10.3,12.7,15.1,15.9,18.4,21,23.7],
15:[8.3,12.6,15.7,18.9,19.9,23.5,27.4,31.6],20:[9.5,14.4,18.1,22.1,23.4,27.8,32.6,37.8],30:[11.1,16.7,21.1,25.8,27.4,32.6,38.4,45],45:[12.8,19.2,24.3,29.7,31.5,38,45.2,53.4],60:[14.3,21.2,
26.6,32.6,34.6,41.7,49.6,58.6],90:[16.4,23.8,29.4,35.6,37.6,44.3,51.7,60.2],120:[18.3,26.1,32,38.2,40.2,46.9,54.3,62.1],180:[21.2,29.6,35.7,41.8,43.8,50.5,57.5,65.3],360:[26.2,35.7,42.2,48.6,50.8,57.4,6
720:[31.8,41.7,48.6,56,58.5,66.1,74.3,83.2],1440:[38.7,49.3,56.9,64.4,66.9,74.7,83,91.4]}


y_group = []
x_list =[t for t in range(1,1441)]
for dur in range(100,1300,200):
  event_duration = dur



  timeoneday = find_events(event_duration, 54, df)
  timeoneday.sort()
  y_list = []
  for f in range(1,1441):

        timeoneday2 = find_events(f, 54, df)

        number = 0
        for s in timeoneday:

                count_list = 0
                for g in timeoneday2:
                  if pd.Interval.overlaps(s, g):
                      count_list += 1

                if count_list > 0:
                    number += 1
        y_list.append(number/54)
  y_group.append(y_list)

pdy = pd.DataFrame(y_group,columns =[u for u in range(1,1441)])

new_df = pdy.transpose()

plt.scatter(new_df[0].index,new_df[0],label="100min")
#plt.plot(new_df[1],label="20")
plt.scatter(new_df[1].index,new_df[1],label="300min")
#plt.plot(new_df[3],label="40")
plt.scatter(new_df[2].index,new_df[2],label="500min")
#plt.plot(new_df[5],label="60")
plt.scatter(new_df[3].index,new_df[3],label="700min")
#plt.plot(new_df[7],label="80")
plt.scatter(new_df[4].index,new_df[4],label="900min")
#plt.plot(new_df[9],label="100")
plt.scatter(new_df[5].index,new_df[5],label="1100min")
#plt.plot(new_df[6],label="120")


plt.legend(title='Events')
plt.ylabel('Overlapping probability')
plt.xlabel("Duration[min]")
plt.show()
```

Appendix 4: Codes for extraction independent merged intervals from the original database.

```
def get_interval_df(interval, df1):
    start_idx = np.searchsorted(df1.time, interval.left)
    end_idx = np.searchsorted(df1.time, interval.right)
    df = df1.iloc[start_idx : end_idx+1]
    return df.set_index("time")

def get_timeseries_with_precip(a, df):
    df1 = df.resample("1min").sum()
    df1.reset_index(inplace=True)
    return [get_interval_df(interval, df1) for interval in a]

dflist = get_timeseries_with_precip(a, df)
```

Appendix 5: Codes for plotting independent merged events into the original Blindern's IDF curves by extrapolating the return periods from the original IDF curve.

```
returnperiods = [2,5,10,20,25,50,100,200]
durations = [1,2,3,5,10,15,20,30,45,60,90,120,180,360,720,1440]
precip_values = np.load('blindern_ivf_prec.npy')

interp1d_precip = interp1d(durations, precip_values, fill_value='extrapolate')

fin_list = []

for v in comb_list:
  new_l = []

  def interp_returnperiod(dur,prec):

    precipvaluesforduration = interp1d_precip(dur)

    interp1d_retper = interp1d(precipvaluesforduration, returnperiods, fill_value='extrapolate')

    returnperiod = interp1d_retper(prec)

    new_l.append([dur,prec,returnperiod])


  T = [interp_returnperiod(i[0], i[1]) for i in v]

  ls = max([i[2] for i in new_l])
  x = [(i, r.index(ls))
  for i, r in enumerate(new_l)
  if ls in r]
  x = list(x[0])
  fin_list.append(new_l[x[0]])


pds = pd.DataFrame.from_dict(new_dict, orient='index')

plt.plot(pds[0],label="2")
plt.plot(pds[1], label="5")
plt.plot(pds[2],label="10")
plt.plot(pds[3],label="20")
plt.plot(pds[4],label="25")
plt.plot(pds[5],label="50")
plt.plot(pds[6],label="100")
plt.plot(pds[7],label="200")

for o in fin_list:
  y_a = [o[1]]
  x_a = [o[0]]

  plt.scatter(x_a,y_a)

plt.legend(title='Years')

plt.ylabel('Precipitation depth[mm]')
plt.xlabel('Duration[min]')
plt.show()
```



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway