TECHNISCHE UNIVERSITÄT DRESDEN

Adequacy of urban water infrastructure under climate change: A Bayesian Perspective

submitted in partial fulfillment of the requirements for the degree of

Master of Science (M.Sc.) in

Hydro Science and Engineering

By

Dawar M. Qureshi

(B.Eng. in Mechanical Engineering)

Matriculation Number: 4805643

Under the supervision of

Dr. Lauren Cook (First Supervisor, second examiner)

(Group Leader, Department of Urban Water Management
Swiss Federal Institute of Aquatic Science and Technology (EAWAG), Switzerland)

Dr. Omar Wani (Second Supervisor, second examiner)

(Postdoctoral Researcher, Environmental Systems Dynamics Lab, University of California Berkeley, U.S.A)

Prof. Dr. sc. techn. Peter Krebs (Responsible Professor, first examiner)

(Chair of Urban Water Management, TU Dresden, Germany)

April 2022, Dresden, Germany

Conceptual Overview

There is substantial evidence that climate change poses a risk to urban drainage infrastructure, which, due to previous assumptions of a stationary climate, may be undersized to withstand expected increases in frequency and intensity of extreme rainstorms. As we are now moving into a period of non-stationarity because of climate change, design practices must be updated consequently. One of the design practices in need of revision are Intensity-Duration Frequency (IDF) curves, which use statistical distributions to represent extreme rainfall in a particular region. Several studies have begun to update these extreme rainfall estimates for future climate change; however, uncertain climate signals, models, and input data increase the variability of the findings. Moreover, parameter uncertainty, which also leads to variation in the results of IDF curves, is sometimes neglected or omitted altogether. These factors reduce the reliability of the stormwater design process, which is currently highly standardized. The engineering community is thus in need of a robust technique that can account for multiple uncertainties in a reproducible manner.

One solution is to use Bayesian inference to update IDF curves to account for a range of uncertainties, including future climate change. The literature in this area is growing; however, it remains unclear whether these complicated methods can produce future updated IDF curves in a reliable manner, thus bringing added value to the stormwater design process. The goal of this thesis is to evaluate the feasibility and reliability of using Bayesian inference to update IDF curves for stormwater design. Several procedures to create future updated IDF curves are compared, with regard to the choice of statistical method used, representation of uncertainty, and the use of future climate data. Finally, in order to evaluate the added value for stormwater design, the rainfall estimates from these various procedures are used to evaluate how the dimensions of a drainage pipe will vary. The analysis is completed for Phoenix, Arizona, US.

Abstract

There is growing evidence that global climate change is increasing the risk of stormwater infrastructure to flooding. These infrastructure systems are traditionally designed using historical rainfall data in the form of intensity-duration-frequency (IDF) curves, assuming a stationary climate. IDF curves use statistical distributions to represent extreme rainfall in a particular region. The parameters of the distribution are estimated using classical methods such as the Maximum Likelihood Estimation (MLE), which calculates best guess estimates of the parameter set. However, several studies have shown that this approach is not able to account for uncertainty in these estimates in a reliable manner. As a result, a more sophisticated method known as Bayesian inference is used in this study to estimate the parameters of a Generalized Extreme Value (GEV) distribution and compare the results with the classical method for both historical and future rainfall data at a station in Phoenix, Arizona, the United States of America. Bayesian inference quantifies uncertainty by assuming a prior distribution on the distribution parameters. Three Bayesian scenarios are developed by varying the choice of prior distribution on the shape parameter of the GEV. Furthermore, the resulting IDF values are used to design a sewer pipe large enough to carry a 50-year storm.

It is observed that MLE underestimates the tail events of the probability distribution compared to the Bayesian method for both historical and future data. This results in lower values of return level or rainfall depth estimates for higher return periods (>25 years). The difference is especially notable for the longer 24-hour duration. Consequently, the pipe is undersized, necessitating the revision of conventional design practices. Furthermore, findings suggest that the choice of prior plays an important role in Bayesian data analysis.

Keywords: Bayesian Inference, Climate Change, GEV parameters, Intensity-Duration-Frequency (IDF) curves, Stormwater Infrastructure, Uncertainty

Declaration of authorship

I attest that this Master Thesis is the result of my own independent efforts. Material that has been taken directly or indirectly from other sources has been labeled as such. This work has not been partially or fully submitted as a graded academic work to an examination committee.

Dawar M. Qureshi

Date: 1.04.2022

Dresden, Germany

Acknowledgements

First, I would like to convey my heartfelt appreciation to Lauren Cook, my first supervisor. Her supervision and research approach throughout the thesis has been very inspiring to me. This work would not have been possible without her continuous support and belief in my abilities. Many thanks for all of your recommendations and brainstorming sessions, which greatly helped me in understanding the research problem.

My second supervisor and mentor Omar has been an invaluable source of support and guidance during this project. His notable work on uncertainty quantification and Bayesian inference has been of enormous benefit. Thank you very much for making Bayesian Inference somewhat less complicated.

I would also like to show my appreciation to Professor Peter Krebs, the responsible professor and chair of urban water department at TU Dresden for showing interest in this thesis. His classes on urban water in the degree program were one of the influencing factors in taking up this topic.

My sincere gratitude to all friends in Germany and back home in Kashmir. Thank you, Ibrahim, Alfonso, Ali, Rashid, Mirza, and Saher for all your kindness, support, and love especially in pandemic times.

I am whole-heartedly grateful for the selfless love and prayers from my amazing parents, Mukhtar and Sabia. All of this would have been impossible without them. My two loving and caring siblings, who have immensely supported me in my decisions. I would also like to mention my teacher Sheikh Nazim, who passed away in 2014, but is remembered every day.

Thank you everyone!

ACRONYMS

AMS	Annual Maximum Series
CDF	Cumulative Distribution Function
CI	Confidence/Credible Interval
ESS	Effective Sample Size
ESM	Earth System Model
GCM	General Circulation Model
GEV	Generalized Extreme Value
HMC	Hamiltonian Monte Carlo
IDF	Intensity-Duration-Frequency
IPCC	Intergovernmental Panel on Climate Change
KDDM	Kernel Density Distribution Mapping
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimation
NA-CORDEX	North American Coordinated Regional Downscaling Experiment
NOAA	National Oceanic and Atmospheric Administration
NUTS	No-U turn Sampler
PDS	Partial Duration Series
RCM	Regional Climate Model
RCP	Representative Concentration Pathway

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and Literature Review

The rapid development of urban areas combined with the complexity of infrastructure systems and increasing precipitation intensity from global climate change is progressively making cities vulnerable to flooding (Willems et al., 2012). Annual maximum daily rainfall extremes have risen significantly on a global scale (Westra et al., 2013), and climate models generally suggest that extreme precipitation events will continue to intensify over the 21st century (Pfahl et al., 2017). These extreme events can cause severe inundation, significant asset damages, and mortality, particularly in high-population urban areas (Mishra & Singh, 2010). In the year 2020 alone, twenty-two extreme events occurred in the US worth $95.8 billion in damages (NOAA, 2021). Improving the resilience of infrastructure systems to future conditions is therefore critical for informing disaster risk management(IPCC, 2021).

Stormwater infrastructure systems are particularly susceptible to changes in extreme rainfall because they are traditionally designed using historical rainfall information in the form of intensity-duration-frequency (IDF) curves. IDF curves estimate the frequency of occurrence of rare rainfall storms over different temporal durations (e.g., 1-hr, 6-hr, 24-hr) (Courty et al., 2019; Vu & Mishra, 2019). Extreme rainfall storms for a particular duration are quantified by fitting a probability distribution to the rainfall intensities that are typically extracted from the observed record(Cheng & Aghakouchak, 2014). The parameters of the distribution are typically calculated assuming they will not change over a time defined as stationarity, which indicates that the intensity of precipitation for a particular frequency and duration will also not change with time (Cook et al., 2020; Hosseinzadehtalaei et al., 2020). However, numerous studies (Mishra & Singh, 2010; Westra et al., 2013) suggest an increase in non-stationary behavior of extremes over time, meaning that the distribution parameters have and will vary over as climate change continues to accelerate. As a result, the assumption of a stationary climate is unsustainable. On this account, the IDF curves must be updated adequately to make urban water infrastructure adaptable to changing precipitation regimes.

Several studies have begun to update IDF curves around the world using various techniques. (Cook et al., 2020) used the annual maximum series (AMS) from observed rainfall and regional climate data from the North American Coordinated Regional Downscaling Experiment (NA-CORDEX) to create future intensity duration frequency (IDF) curves for 34 cities in the United States. They discovered that the modeling choices such as the spatial resolution and spatial adjustment techniques significantly alter rainfall predictions and uncertainty in updated IDF curves and the subsequent stormwater infrastructure decisions based on these curves. (Vu & Mishra, 2019) performed a non-stationary frequency analysis for six most recent extreme precipitation events in the United States, occurred over different durations by incorporating time-varying covariates (e.g., time, maximum temperature, etc.) into the nonstationary Generalized Extreme Value distribution to emphasize the importance of revising the IDF curves in infrastructure design. Over the selected time period, they found out that the majority of the extreme precipitation events followed a nonstationary pattern which could be due to an increase in the magnitude of recent extreme precipitation events, particularly hurricanes. Furthermore, the study showed that the return periods associated with extreme precipitation events were significantly reduced by the non-stationary extreme value analysis when compared to the stationary approach. Sharma et al (2021) quantified the upfront costs

required to achieve greater hydraulic reliability and robustness in the face of deep uncertainties in rainfall, surface runoff characteristics, and infrastructure lifetime projections (Sharma et al., 2021). They observed an increase in mean estimates of extreme precipitation intensity, which in turn increase with longer return periods. A continental level study over Europe for updating the IDF curves was carried out by (Hosseinzadehtalaei et al., 2020). An upward shift in IDF curves under climate change, with a larger degree under Representative Concentration Pathway (RCP) 8.5 compared to RCP4 was reported by them.

The above cited literature addresses either uncertainties in the parameters of the appropriate distribution for the historical time period or uncertainties in general circulation models for the future time period using classical statistical methods such as maximum likelihood estimation (MLE). With efficient and readily available computational resources, Bayesian method has also gained popularity recently. Given some data, the MLE calculates point estimate known as the best guess of the parameter values of the assumed distribution. This method expresses uncertainty by calculating error bars or confidence intervals on the data set. Under this setting the likelihood function (the likelihood describes the probability that each potential parameter value generated the data set) is a fixed parameter with some true value and considers the distribution of possible data sets to obtain confidence intervals or error bars. The true value, in this case, is not a random variable and the confidence interval is thus a statement that the true value may be between the upper and lower limits of that interval. Therefore, the probabilistic statement (e.g., 95% confidence interval) is made on the interval rather than the position of the actual value itself. This disparity is addressed by Bayesian scheme, where only observed data set is considered, and the uncertainty of a parameter is expressed by assuming a prior distribution for that parameter. This is particularly beneficial when only short-time series are available. The point estimate (e.g., the mean) of the parameter of interest is then estimated from the posterior distribution. The true value is therefore a random variable in such a setting. Bayesian 95% Credible interval would mean that given the evidence provided by the observed data, there is 95% probability that the true point estimate would lie in that interval (Coles, 2001). Bayesian Inference can therefore capture uncertainty in parameter estimations more effectively. As a result, the current point estimate procedure with classical methods for designing stormwater infrastructure may no longer be viable.

Additionally, the choice of prior distribution on distribution parameters is an essential subject in the field of Bayesian data analysis. There is no agreement in the literature (Hosseinzadehtalaei et al., 2020; Rupa et al., 2015; Sharma et al., 2021) on this issue with some studies adopting subjective while others restricting themselves to objective choices. As a result, it might be of interest to investigate all the commonly used prior distributions in a single research and compare the inference results with each other as well as with traditional statistical approaches.

These inconsistencies and uncertainties in the IDF values may lead to discrepancies in stormwater designs that may alter the level of protection provided by the infrastructure. There could be changes in design as a result of using traditional approaches. It is possible that the traditional techniques lead to an underestimation of IDF values, resulting in a significant risk associated with the infrastructure,

or that the design values are overestimated, thus increasing the cost-burden. As such, quantifying these uncertainties is critical for constructing safe and cost-effective stormwater infrastructure systems.

1.2 Research Aims

The research objectives of this scientific study are:

- To present a comparative analysis of historical IDF curves developed using classical Maximum Likelihood estimation (MLE) and Bayesian Inference for a station in Phoenix, Arizona, USA.
- Create future IDF curves from climate models using both fitting techniques in order to analyze the range of precipitation forecasts at the same site.
- Finally, evaluating how each fitting technique influences design of stormwater infrastructure by taking a simple example of an urban sewer pipe.

1.3 Structure of the Thesis

The thesis is made up of five chapters. Chapter 1, the current introductory chapter provides the motivation, literature review and aim of this work. Chapter 2 gives all the necessary background information while chapter 3 presents the data and methodology adopted in this study. In chapter 4, the results are analyzed and discussed. Finally, the chapter 5 presents the concluding remarks and recommendations for future studies. Furthermore, all the supplementary information supporting the results is provided in the form of appendices at the end.

Chapter 2

Background

2.1 Introduction to Stormwater Design and IDF Curves

Stormwater infrastructure systems are most commonly designed by considering the flowrate that would result from a particular rainfall "event" called the "design storm". This rainfall event is characterized by three factors: the intensity, i (depth per time), the frequency or return period, t and the duration, d. Intensity-duration curves (IDF) graphically depict these rainfall intensities of various durations and return periods. They are usually constructed through statistical analysis of long records of observed rainfall data (Canadian Standards Association, 2012).
This statistical analysis consists of three main steps. First, the extreme rainfall values are extracted from the observed record. Second, an appropriate probability distribution is fit to these extreme values to estimate the distribution parameters. Finally, the return level (or rainfall depth) can be extracted from the distribution for a particular probability.

Extreme rainfall values can be extracted in two main ways: (1) by calculating the Annual Maximum Series (AMS) which include only the highest values for each year of the recording period., and (2) by estimating partial duration series (PDS) through peak-over threshold (POT) analysis, which extracts a sequence of separate peak events that exceed a high threshold value. The contrast between the two methods lies in the treatment of extreme events. A significant event that is not the maximum for the year in which it occurs is not included in the AMS analysis, but it may represent a maximum for another year. The PDS, on the other hand, considers all events that exceed a certain threshold. In general, the AMS and PDS are analyzed using different distributions and the results are interpreted differently. To represent the probability of occurrence of independent extreme values of rainfall in any given year, an event with a return period T is likely to be equaled or exceeded by 1/T.

Once extracted, the AMS are fit to extreme value distribution, which are statistical models that describe the likelihood of extreme events. PDS, on the other hand, are typically fit to generalized Pareto distributions (Hans Van de Vyver & Demarée, 2010).

2.2 Extreme Value Analysis

Calculation of AMS is the most widely used approach to estimate extreme values. Therefore the Generalized Extreme Value Distribution (GEV) is the most often utilized distribution to estimate meteorological extremes (Canadian Standards Association, 2012). The GEV distribution (equation 1) is characterized by three parameters: location, μ, scale, σ and shape, ξ. It combines three Extreme Value (EV) distributions into a single distribution, which means there is no need to choose which extreme value distribution to adopt. The three EV distributions, Gumbel, Fréchet, and Weibull, vary according to the shape parameter, which depicts that tail of the distribution. Gumbel (EV-I, $\xi = 0$) has an exponential tail, Fréchet (EV-II, $\xi > 0$) has a more slowly decaying tail (fat tail) and this distribution has an upper limit. Weibull (EV-III, where $\xi < 0$) has an exponentially decaying tail (Coles, 2001).

The Cumulative Distribution Function (CDF) of the Generalized Extreme Value (GEV) family of distributions is given by:

$$G(x) = exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \quad (1)$$

The parameters satisfy: $-\infty < \mu < \infty, \sigma > 0$ and $-\infty < \xi < \infty$.

Inverting the equation 1 yields return level quantiles:

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1 - y_p{}^{-\xi}\right], & \text{for } \xi \neq 0 \\ \mu - \sigma \log y_p, & \text{for } \xi = 0 \end{cases} \quad (2)$$

$$y_p = -\log(1-p) \quad (3)$$

where z_p is the return level associated with the return period 1/p. The return level z_p equates to rainfall depths while analyzing Annual Maximum Series. The intensity is computed by dividing z_p with the duration. As an example, the density function of all three kinds of GEV distributions with μ= 0 and σ=1 is illustrated in Figure 1.

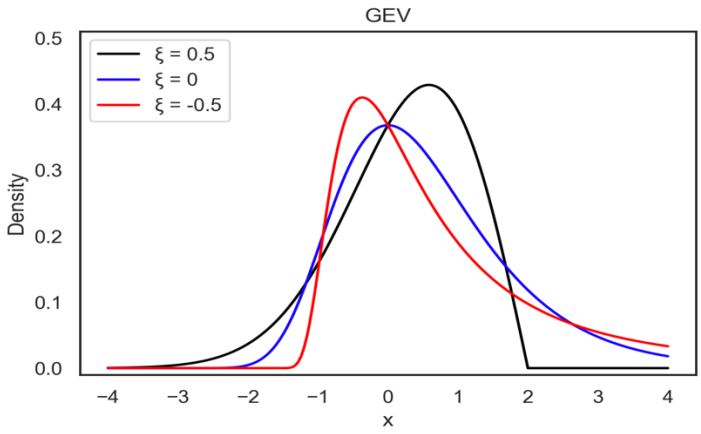

Figure 1: Density functions of a Generalized Extreme Value distribution

2.3 Parameter Estimation Methods

2.3.1 Classical Maximum Likelihood Estimation (MLE)

Maximum Likelihood Estimation (MLE), a traditional method used to estimate the density of a probability distribution. MLE entails developing a likelihood function to the conditional probability of a data sample, given a desired probability distribution and distribution parameters. In this analysis,

MLE is used to fit a GEV distribution to the annual maximum rainfall values, represented by x1, x2...xn, which are considered independent variables. Three parameters must be determined: shape (ξ), location (μ) and scale (σ). Under this assumption, the log likelihood for the GEV-parameters when the shape parameter is not zero is given by the equation:

$$L(\xi,\mu,\sigma) = -nlog\sigma - (1+\frac{1}{\xi})\sum_{i=1}^{n} Log[1+\xi(\frac{x_i-\mu}{\sigma})] - \sum_{i=1}^{n} Log[1+\xi(\frac{x_i-\mu}{\sigma})]^{-\frac{1}{\xi}} \quad (4)$$

Where,

x1, x2, …, xn represents the annual maximum rainfall values

μ, σ and ξ represents the location, scale, and shape parameters respectively

Given that

$$1+\xi\left(\frac{x_i-\mu}{\sigma}\right) > 0 \;\; for\; i = 1,2,\dots,n \quad (5)$$

The parametric relation $\frac{\mu-\sigma}{\xi}$ determines the end points of a GEV distribution accordingly when $\xi < 0$ (upper-end point) and $\xi > 0$ (lower-end point). If at least one observed data point falls beyond the endpoint, equation 5 is violated and therefore the likelihood is zero and the log-likelihood is -∞ .The Gumbel case of the GEV distribution is followed if $\xi = 0$. The Log-Likelihood in this case is:

$$L(\mu,\sigma) = -nlog\sigma - \sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right) - \sum_{i=1}^{n} exp-\left(\frac{x_i-\mu}{\sigma}\right) \quad (6)$$

The maximum likelihood estimate for the entire GEV family can be therefore obtained by maximizing the equations 4 and 6 with respect to the parameter values (ξ,μ,σ). This is obtained by simple numerical optimization algorithms.

2.3.2 Bayesian Estimation

Bayesian statistics employs probabilistic assertions to draw conclusions about a parameter vector θ or an unobserved data set $\hat{z}$. These probabilistic statements are conditional on the observed value of data vector z = $(z_1, z_2, \dots, z_n)$ and are written as π (θ|z) or π ($\hat{z}$|z). Prior to examining the data, the assumptions about θ are recorded, in the form of prior probability distribution π (θ). A joint probability distribution function for θ and $\hat{z}$ is written as a product of prior distribution π (θ) and likelihood function π (z|θ) to make probabilistic statements. For distinctive settings of the parameter vector θ, the likelihood function expresses how probable the observed data set is. To yield the posterior probability, Bayes' theorem combines the prior probability and the likelihood function as:

$$\pi(\theta|z) = \frac{\pi(\theta)\pi(z|\theta)}{\pi(z)} \quad (7)$$

The denominator in this equation is the normalization constant which makes sure that the posterior distribution on the left side of the equation 7 is a valid probability density and integrates to one. The distribution of the unknown but observable z is

$$\pi(z) = \int \pi(z,\theta)d\theta = \int \pi(\theta)\pi(z|\theta)d\theta \quad (8)$$

This is called as the marginal or prior predictive distribution of z. The marginal distribution does not depend on the parameters for which one want to make inference as these parameters have been integrated out, which makes it difficult to calculate the integral in closed form. Therefore, the equation 7 is generally written in the form

$$\pi(\theta|z) \propto \pi(\theta)\pi(z|\theta) \quad (9)$$

Or in words, Posterior $\propto$ Likelihood $\times$ Prior

The problem to solve the integral in the normalization constant is realized by simulation-based techniques such as Markov Chain Monte Carlo (MCMC) which allow us to sample from the posterior and draw distributions over our sample without having to worry about the evidence or the normalization constant. In other words, for most probabilistic models, precise inference is impossible, thus we must resort to some form of approximation instead (Bishop, 2006; Stuart G . Coles and Jonathan A, 2010).

2.4 Bayesian Inference Techniques

In this study, the Hamiltonian Monte Carlo (HMC), discussed in the next section is employed. HMC is a MCMC technique that uses Hamiltonian dynamics for obtaining a sequence of random samples. This technique reduces the problem of producing heavily correlated samples observed in other commonly employed MCMC methods such as the Metropolis-Hastings algorithm and Gibb's sampling. For a detailed theoretical background on MCMC methods and HMC, the reader is referred to Michael Betancourt's work on the topic (Betancourt, 2017)and Chris Bishops Pattern Recognition and Machine Learning (Bishop, 2006).

2.4.1 Hamiltonian Monte Carlo (HMC)

Originally developed as Hybrid Monte Carlo to address the problems in molecular dynamics, it was Radford Neal who realized the potential of this method for applications in applied statistics in his ground-breaking research on Bayesian neural networks (Neal, 1995). HMC is a sophisticated class of algorithms that samples on an enhanced target density and amplifies the target posterior by adopting synthetic momentum variables. This enhanced target is proportional to the exponential of the total energy also called as the Hamiltonian function, where the total energy is the sum of potential and kinetic energies analogous to physical systems (Betancourt, 2017).

The purpose of sampling is to draw from a Bayesian posterior $\pi(\theta|y)$ of parameters θ given the data y. HMC introduces auxiliary momentum variables denoted by **p** and draws from the joint density

$$\pi(p,\theta) = \pi(p|\theta)\pi(\theta) \qquad (10)$$

The joint density defines a Hamiltonian

$$H(p,\theta) = -log\ \pi(p,\theta)$$
$$= -log\pi(p|\theta) - log\pi(\theta) \qquad (11)$$

Where the term $\mathrm{K(p)} = -log\pi\ (p|\theta)$ is called the kinetic energy and the term $E(\theta) = -log\pi(\theta)$ is called the potential energy. The Hamiltonian dynamics describes the evolution of total energy over time. One of the most intriguing aspects of the Hamiltonian is that it conserves energy over time. Two stages are essential for the transition of the parameters θ from the current state to the new state. First, A value for the momentum variable is drawn

$$\mathrm{P} \sim \text{Multinormal}\ (0, \mathrm{M}).$$

The covariance matrix M is independent of the current parameter θ values. Next, the system (θ, p) comprising of the current parameter values and new momentum is evolved by Hamilton Equations

$$\frac{d\theta}{dt} = +\frac{\delta H}{\delta p} = \frac{\delta K}{\delta p} \qquad (12)$$

$$\frac{dp}{dt} = -\frac{\delta H}{\delta \theta} = -\frac{\delta K}{\delta \theta} - \frac{\delta E}{\delta \theta} \qquad (13)$$

Where $\frac{\delta K}{\delta p}$ is zero

The Kinetic energy is chosen to be Gaussian

Therefore, $K(\theta,p) = \frac{1}{2}p^T M^{-1}\ p\ +\ \log|\mathrm{M}| +\ \mathrm{const}$

With M=I,

$K(\theta,p) = \frac{1}{2}p^T\ p\ +\ const$

So $\frac{\delta K}{\delta p} = p$

And $\frac{\delta K}{\delta \theta} = 0$

The Hamiltonian equations are simplified to

$$\frac{\delta \theta}{\delta t} = p \qquad (14)$$

$$\frac{\delta p}{\delta t} = -\frac{\delta E}{\delta \theta} \qquad (15)$$

2.4.2 Leapfrog Integration

The two differential equations 14 and 15 need to be simulated or solved numerically, practically by a time stepping algorithm. Generally, leapfrog integrator is employed for this purpose. The leapfrog algorithm works in discrete steps separated by brief time intervals. The leapfrog algorithm starts by generating a new momentum term that is independent of the parameter values θ or prior momentum value. It entails updating the momentum p by half a step, then the position by a full step, and finally updating p by the other half of a step. If τ is the small-time step,

$$p \longleftarrow p - \frac{\tau}{2}\frac{\delta E}{\delta \theta}$$

$$\theta \leftarrow \theta + \tau \sum p$$

$$p \leftarrow p - \frac{\tau}{2}\frac{\delta E}{\delta \theta}$$

With the applications of S leapfrog steps, $S\tau$ time is simulated. The resulting state after the end of the above three steps is (p^*, θ^*)However, the Leapfrog integrator is not numerically perfect. Therefore, to account for the error bounds during leapfrog integration, a common practice is to apply the Metropolis acceptance step. The probability of keeping the proposal (p^*, θ^*) is

$$Min(1, exp(H(p, \theta) - H(p^*, \theta^*))) \qquad (16)$$

If the proposal is rejected, the previous parameter value for the next draw is returned and used to start the following iteration (Stan Development Team, 2020).

2.4.3 Choice of Priors for Bayesian Estimation

The choice of priors is non-trivial in Bayesian data analysis, and expert knowledge is typically required to augment information with which to supplement the data. The class of priors based on expert knowledge are known as *informative priors*. In the field of extreme value analysis, this would translate to having a solid understanding of the physical generation of rainfall, as well as detailed knowledge of the rainfall properties at the specific study location under consideration. However, if no such knowledge is available, *non-informative priors* or priors with a wide variance are commonly adopted in practice. There is currently no consensus in the literature on the choice of priors for Extreme Value Analysis. (Huard et al., 2010) and (Rupa et al., 2015) used a non-informative, uniformly distributed prior for the location parameters and a prior following Jeffreys distribution for the scale parameter. For the shape parameter, they chose a weakly informative Beta prior. (H Van De Vyver, 2015) used non-informative normally distributed priors for the location and scale as suggested by (Coles, 2001) and weakly informative beta prior for the shape parameter. (Hosseinzadehtalaei et al., 2020) chose

non- informative normal distributions for the location and scale parameters, and a normal distribution with a standard deviation of 0.3 for the shape parameter. (Sharma et al., 2021) used a non-informative Gaussian prior distribution with a mean zero and a wide variance for all the three parameters. In their detailed examination of the annual maximum daily rainfall of over 15,000 records from throughout the world, (Papalexiou & Koutsoyiannis, 2013) discovered that the location parameter for yearly maxima of daily rainfall spans between 6 and 700mm/day and the range or scale parameter varies from 2-400mm/day. They also concluded that the shape parameter is likely to fall in the range of 0 to 0.23.

Chapter 3

Data and Methods

This chapter describes the data and methods used in the analysis. Section 3.1 introduces the study site and input data used for both historical and future analysis. Section 3.2 presents the analysis for IDF curves and extreme value analysis while section 3.3 provides the description of parametric inference methods including the computational tools used for inference in the study. Finally, the post-processing of the IDF values is presented in section 3.4.

3.1 Study Site and Input data

Phoenix, Arizona, is a hot-arid metropolis in Southwest America with a population of roughly 5.05 million people (in 2020). The city was chosen as a case study because it has experienced record-breaking floods in recent times, especially the floods of 2014 (Battaglia, 2019). The Phoenix weather station (PHOENIX AIRPORT AZ US) is located at an elevation of 337m and geographical coordinates 33.4-degree latitude and - 112-degree longitude. The annual total rainfall is, on average, about 187 mm, which is substantially lower than the average U.S. contiguous precipitation of about 760mm (Cook et al., 2020; NOAA, 2022b).

A total of 65 years of continuous, hourly precipitation observations from 1950-2014 are used for historical analysis. These data were obtained from the NOAA National Center for Environmental Information (NOAA, 2022a) data repository.

For the future analysis, output from regional climate models (RCMs) from the North American Coordinated Regional Downscaling Experiment (NA-CORDEX) (Mearns, L.O., 2017) is used. RCMs are used for estimating climatic variables such as temperature, precipitation, wind, etc., at regional scales and were initially intended to bridge the gap between the coarse estimates from General Circulation Models (GCM), used to model the response of global climate system to rising greenhouse gas emissions (Tapiador et al., 2020). These RCMs, which are driven by the Coupled Model Intercomparison Project Phase 5 (CMPI5)-GCM boundary conditions, were selected due to their availability at an hourly step over a continuous 150-year timespan from 1950-2100. Presented in Table 1, three RCM-GCM combinations are used with all model simulations using only RCP-8.5 greenhouse gas pathway, which represents the business as usual or worst-case scenario (Riahi et al., 2011).

Table 1: Climate data from RCM-GCM combinations for future analysis

RCM	GCM	Time Period	Years	RCP
RegCM4	Max Planck Institute Earth System Model (MPI-ESM-LR)	2014-2099	86	8.5
WRF	(MPI-ESM-LR)	2014-2099	86	8.5
WRF	Geophysical Fluid Dynamics Laboratory Earth Systems Model, Modular Ocean Version (GFDL-ESM2M)	2014-2099	86	8.5

In this study, output from RCMs are bias corrected by the technique presented in (Cook et al., 2020). This technique of non-parametric bias correction called Kernel Density Distribution Mapping (KDDM) (McGinnis et al., 2015) includes altering the data from the climate model output based on the discrepancies with the observed dataset by constructing a transfer function that maps the model data's Cumulative Distribution Function (CDF) to the observed data's inverse CDF.

3.2 IDF Curves and Extreme Value Analysis

The Annual maximum Series (AMS) from the observed rainfall series and Regional Climate Model (RCM) simulations of the study site is extracted for two durations (1h and 24h). Two fitting techniques are used to estimate the parameters of extreme value distribution (shape $\boldsymbol{\xi}$, location $\boldsymbol{\mu}$, and scale $\boldsymbol{\sigma}$) from the annual maximum series (AMS). The first is conventional Maximum Likelihood Estimation (MLE). The maximum likelihood estimates of the three parameters of GEV distribution are computed with SciPy stats *genextreme* function in Python programming language (van Rossum, 2017).

The second fitting technique used is the Bayesian method to fit the GEV that employs the Hamiltonian Monte Carlo (HMC) method, described in section 2.4.1 (details regarding the Bayesian model are provided in the next section).

The Bayesian predictive distribution of return levels is then obtained by considering 5000 samples from the posterior distributions of the three GEV parameters using *SciPy's genextreme.rvs* function in the Python programming language. The equivalent return level distribution simulated using MLE best guess estimates is also derived to present a comparative overview.

Furthermore, the ensemble of rainfall depths or return levels produced with three RCM-GCM forcing's by both MLE and Bayesian Methods is averaged to account for intermodal uncertainty in the point estimates.

3.3 Bayesian Inference Model

To solve the mathematical terms described in section 2.4.1 and 2.4.2 , an open-source probabilistic programming language written in python called PYMC3 (Salvatier et al., 2016) is used. PYMC3 is built on Theano, a robust optimization toolkit that enables the efficient computing of complex tensor operations by performing computation either on a Central Processing Unit (CPU) or a Graphics Processing Unit (GPU). Multi-parametric model estimation given very big data sets is feasible using GPU-based optimization. This makes Bayesian Inference less cumbersome, which is otherwise known to be computationally very complex. The syntax used for model building in PYMC3 is very instinctive and easy to read. It includes a self-tuning version of HMC called NUTS (No-U-turn Sampler)(Hoffman & Gelman, 2014) which doesn't require the user to have any specialized knowledge about how the algorithm works.

3.3.1 Choice of Prior

Based on the reviewed literature, a non-informative prior for the location and scale parameters is chosen. For both parameters, a normal distribution (μ=1) with a large variance (σ=100) is selected. For the shape parameter, three different priors as specified in Table 2 are tested. The three priors for the shape parameter include: (1) a non-informative normal distribution with a with mean of zero and a large variance to reflect prior ignorance, (2) a truncated normal distribution with mean of zero, a standard deviation of 0.4, and bounds [-6,6], and (3) a weakly informative Beta prior with hyper parameters a = 6 and b=9. The latter choice constraints the shape parameter to a reasonable range of [-5,5] and was originally recommended by (Martins & Stedinger, 2000).

Table 2.Prior Choice for GEV parameters

Scenario	Shape	Location	Scale
One	$\mathcal{N}(1,100)$	$\mathcal{N}(1,100)$	$\mathcal{N}(1,100)$
Two	$\Psi(0, 0.4, a = -6, b = 6)$	$\mathcal{N}(1,100)$	$\mathcal{N}(1,100)$
Three	$Beta(\alpha = 6, \beta = 9)$	$\mathcal{N}(1,100)$	$\mathcal{N}(1,100)$

3.3.2 Posterior Estimation

To estimate the posterior parameters using Bayesian inference, the Hamiltonian Monte Carlo (HMC) Algorithm is applied in PYMC. To do so, several parameters associated with the HMC Algorithm need to be selected, which includes the tuning, the draw, and the number of chains. The tuning lets the chains converge to the distribution of interest and sample unbiased samples from the model after which the model draws the specified number of samples with the draw parameter in PYMC. The number of Markov Chains can be specified with the chain parameter. It is always recommended to run more than one chain in case that one of the chains would fail to converge, it is then possible to use the alternate sampled chains for the inference model. In this analysis, sampling is done with 2000 tuning and 5000 draws in four chains. Furthermore, two numerical indicators, the Gelman-Rubin convergence diagnostic, R_hat, $\hat{r}$, and Effective Sample Size (ESS), are used to examine the convergence of chains. R_hat, $\hat{r}$, employs the variance technique to assess convergence, calculating both between- and within-chain variance. R_hat should ideally approach 1, otherwise it shows that the chains have not yet converged. Bulk Effective Sample Size (ESS_bulk) measures the sampling efficiency in the bulk of the distribution while Tail Effective Sample Size (ESS_tail) measures the sampling efficiency in the tails of the distribution. For a Markov Chain to be considered reliable, both ESS bulk and ESS tail should be at least 100. Further discussion on convergence diagnostics is presented in Appendix A.

3.4 Post-Processing of IDF Values

To demonstrate the impact of parametric uncertainties in return levels on stormwater infrastructure, a stormwater sewage pipe large enough to contain a 50-year storm is designed. The peak storm water run-off is calculated with the Rational method using the formula

$$Q = 0.00278\, CIA \qquad (17)$$

Where,

Q is the peak discharge,

C is the dimensionless runoff coefficient

I is the rainfall intensity in mm/h (Intensity is obtained by dividing rainfall depths with the duration under consideration)

A is the drainage area in acres

The pipe is designed for intensities corresponding to 50-year, 1- and 24-hour storm events. The character of drainage area, soil types, vegetation amount and types, land cover, and infiltration determines the run-off coefficient C. For a moderately urbanized watershed C= 0.65 and the area A is taken as 10 acres for the 1-hour event and 40 acres for the 24-hour event. The pipes are designed for open channel flow using the Manning's formula

$$Q = \frac{0.31}{n} D^{\frac{8}{3}} S^{\frac{1}{2}} \qquad (18)$$

n is the Manning roughness coefficient

D is the pipe diameter in mm

S is the pipe bottom slope

A pipe with Manning's roughness coefficient n = 0.013 for ordinary concrete lining and s = 0.005 is considered. The resulting diameters are rounded off to the nearest US pipe sizes: 450, 525, 600, 675, 750, 900, and 1050 mm

Chapter 4

Results and Discussion

This chapter presents the general trends and results of the study. Historical parametric inference and IDF values are discussed in Section 4.1, while future IDF values are described in Section 4.2. Finally, the results of sewer pipe sizing are described in section 4.3

4.1 Historical GEV distribution and IDF values

The first research question discussed in this study is to present a comparative analysis of historical GEV parameter estimates and rainfall depths for the 1-and 24-hour duration storms. Parameters and return levels calculated using traditional Maximum Likelihood Estimation are compared to the posterior parameters for three Bayesian prior scenarios (see Section 3.3.1 in methods).

4.1.1 Historical GEV parameter estimates

Table 3 presents the comparison of GEV parameters calculated by Bayesian Inference and MLE for the 1- and 24- hour duration. Shown in brackets are the 95% Confidence Intervals for the MLE and the 95% Credible Intervals for the Bayesian method.

Table 3: Point estimates with 95% confidence/credible intervals shown in brackets of GEV parameters for 1- and 24-hour durations of the historical data set.

Method	Duration	ξ	μ	σ
MLE	1-h	0.06[-0.21, 0.32]	12.87[11.09,14.66]	6.05[4.67, 7.43]
	24-h	0.08[-0.15, 0.32]	23.43[20.99, 25.87]	6.05[4.67, 7.43]
Bayesian Scenario One	1-h	0.082[-0.16,0.36]	13[11,15]	6.3[4.9,7.8]
	24-h	0.11[-0.13, 0.34]	23[21,26]	8.9[7,11]
Bayesian Scenario Two	1-h	0.07[-0.17, 0.33]	12[11,14]	6[4.7,7.5]
	24-h	0.096[-0.11,0.32]	23[21,25]	8.7[6.7,11]
Bayesian Scenario Three	1-h	0.28[0.11, 0.45]	13[11,15]	6.3[4.9, 7.9]
	24-h	0.27[0.11, 0.44]	23[21,26]	8.9[7,11]

The Bayesian parameter estimates are consistent with the MLE estimates for Bayesian prior scenarios one and two, which assume Gaussian and Truncated Gaussian distributions for the shape parameter, respectively. However, the MLE parameters are considerably different from the commonly employed weakly-informative Beta prior for both the durations. Additionally, for all four methods, the shape parameter is positive, which translates to heavy-tailed Fréchet distribution. The distribution of each shape parameter for the three Bayesian prior scenarios is shown in Figure 2.

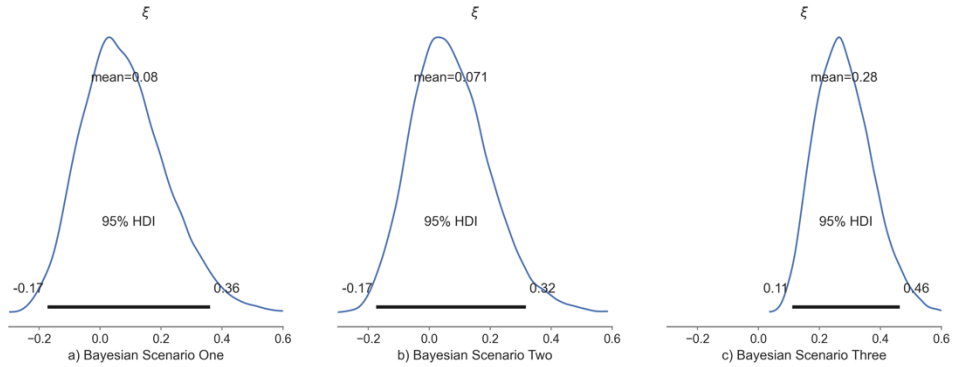

Figure 2: Bayesian posterior distribution of shape parameter for 1-hour duration across different scenarios of the historical data set

4.1.2 Historical rainfall return level estimates

The risk estimation of an event is expressed by calculating the return levels or rainfall depths for different return periods. This estimation has a direct dependence on the parametric uncertainty. Figure 3 presents the GEV probability density function for the three Bayesian prior scenarios compared to the MLE for the two durations. The Bayesian approach leads to more probability mass in the tails of the distribution of rainfall intensities compared to the MLE. This difference is more prominent in the longer 24-hour duration (right) than the shorter 1-hour duration (left). Results for the return level estimates or rainfall depths corresponding to the 2-, 5-, 10-, 25-, 50-, and 100-year return periods are listed in Table 4.

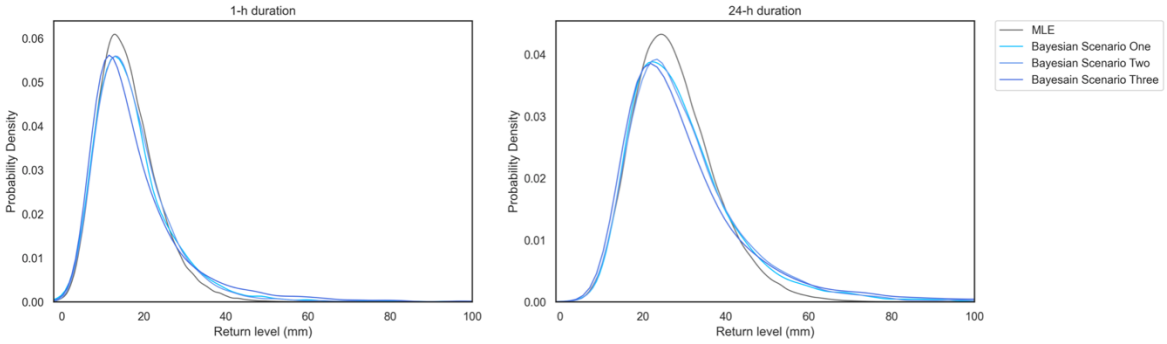

Figure 3: Comparison of predictive distributions between MLE and Bayesian inference for 1- and 24-hour durations of the historical data set.

Table 4: Historical Return Level estimates of 1-h and 24-h durations for different probabilities of exceedance in Phoenix, Arizona, U.S.

Return Period 'T' years	Exceedance Probability (p) p = 1/T	Duration (hours)	Return Level (mm) MLE	Return Level (mm) Bayesian Scenario One	Return Level (mm) Bayesian Scenario Two	Return Level (mm) Bayesian Scenario Three
2	0.50 (50%)	1-h	15.07	15.15	15.20	14.70
		24-h	26.47	26.73	26.68	26.15
5	0.20 (20%)	1-h	21.05	22.30	22.93	23.36
		24-h	35.41	38.01	37.56	38.90
10	0.10 (10%)	1-h	25.64	28.60	28.40	30.90
		24-h	40.73	46.20	45.61	49.96
25	0.04 (4%)	1-h	30.59	36.21	36.0	44.0
		24-h	47.08	57.27	57.16	67.89
50	0.02 (2%)	1-h	34.15	43.71	42.41	55.50
		24-h	51.63	66.77	66.90	85.41
100	0.01 (1%)	1-h	37.47	51.50	50.18	73.20
		24-h	55.88	78.82	77.48	106.06

Fig 4 presents a comparison of percentage change (%) of the return level estimates of the three Bayesian Inference scenarios with respect to the maximum likelihood estimates for different probabilities of exceedance. As expected from the results of parametric inference, the Bayesian estimates of rainfall depths are larger than the MLE estimates for the same probability of occurrence. This difference is quite visible for longer return periods greater than 25-years (p = 4%). The increase in rainfall depths of the Bayesian predictions is to be expected because the Bayesian estimates incorporate additional source of information in the form of prior distributions compared to MLE estimates. The difference in estimates ranges from about 18% for a 25-year 1h storm to 37% for a 100-year 1-h storm between the MLE and Bayesian scenario one. The longer 24-hour duration follows the same pattern. Previous research (Huard et al., 2010; Sharma et al., 2021) have found similar trends, confirming these findings. Therefore, using the point estimates calculated from the MLE technique for infrastructure design, leads to a considerable underestimation of the return levels for return periods larger than 25-years, undermining estimations of pluvial flood risk that depend on rainfall estimates for longer return periods. This undervaluation could result in increased risk of infrastructure failure.

The difference between MLE and Bayesian estimation is especially apparent in the third Bayesian scenario, which assumes a weakly informative Beta prior. For a 1-hour storm, the difference between MLE and the Bayesian estimate for the 25-year and 100-year return periods is 13.4 mm (+43.8%) and 35.73 mm (+95.35%), respectively. This shows that choice of Beta prior might be overly informative for this location, and dependent on local conditions. Furthermore, constraining the Gaussian distribution to bounds, as in the case of Bayesian scenario two, has little effect on the inference results, with scenario one and two showing minimal difference between each other which means that this choice can be considered weakly-informative. These assertions could, however, be only validated by analyzing multiple locations.

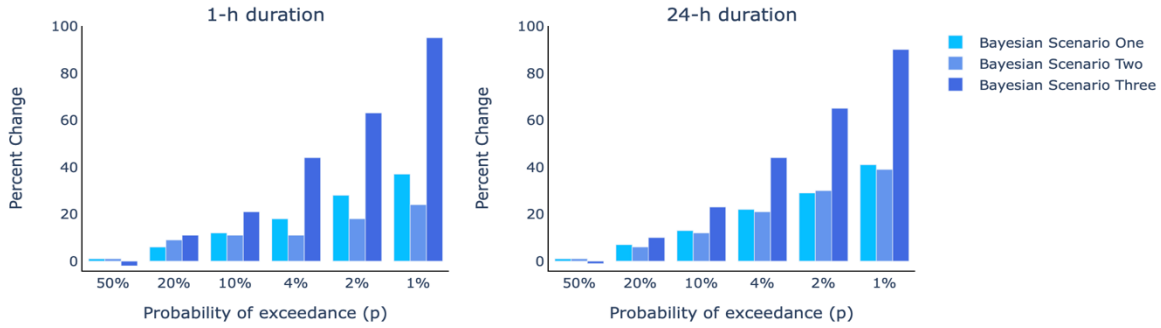

Figure 4: Percent change of Bayesian rainfall depth estimates with respect to the MLE in Phoenix, Arizona, U.S of the historical period.

4.2 Future GEV distribution and Rainfall depth values

Not only does MLE underestimates the rainfall depths when compared to the Bayesian method as observed in the previous section, but the assumption of a stationary climate also plays an important role in the analysis of precipitation extremes. Therefore, future rainfall depths are obtained by fitting the future rainfall data from the three climate models described in section 3.1 to the stationary GEV distribution using both the classical MLE and Bayesian scenario one viz normally distributed prior choice with a large variance. The results of return level estimates are obtained for individual model simulations as well as the ensemble average of three-model forcing's to account for model uncertainty. The return level results for the 1- and 24-hour durations are listed in Table 5 for the model ensemble. Shown in brackets are the upper and the lower uncertainty bounds in mean point estimates resulting from different model combinations used in this study (see section 3.1). As expected, Bayesian method results in broader uncertainty ranges compared to MLE for higher return periods across both the durations. (Rupa et al., 2015) found similar trends for RCP 8.5 scenario which suggests that the uncertainty corresponding to return levels for RCP 8.5 is high. This also underlines the significance of including climate model uncertainty in developing future IDF curves.

The return level estimates of the 3 RCM-GCM combinations are presented in Appendix B.

Table 5: Future Return Level estimates of 1- and 24-h durations for different probabilities of exceedance in Phoenix, Arizona, U.S.

Return Period 'T' years	Exceedance Probability (p) p = 1/T	Duration (hours)	Return Level (mm) Bayesian Scenario One (Future)	Return Level (mm) MLE (future)
2	0.50 (50%)	1-h	14.21 [13.26, 15.77]	13.87[13.1, 15.24]
		24-h	42.75[40.21, 44.10]	41.10[39.08, 42.50]
5	0.20 (20%)	1-h	23.22[22.40, 25.59]	19.80[19.83, 24.26]
		24-h	76.26[69.16, 84.24]	74.2 [68.58,81.79]
10	0.10 (10%)	1-h	30.54[29.04, 34.03]	28.61 [27.0, 31.0]
		24-h	106.56[95.13, 119.79]	102.06 [93.11, 115.71]
25	0.04 (4%)	1-h	42.06[39.43, 48.62]	39.03[36, 44.0]
		24-h	158.32[141.12, 178.94]	147.68[133.91, 171.95]
50	0.02 (2%)	1-h	53.30[49.01, 65.04]	48.72[44, 56.1]
		24-h	212.55[187.33, 242.10]	190.27[173.34,223]
100	0.01 (1%)	1-h	67.31[60, 85.06]	58.28 [52, 69]
		24-h	278.30[248.40, 330.13]	243.6 [219.81, 290]

Another important research question addressed in this study is to assess the effect of incorporating parametric uncertainty on the historical and future rainfall data for future engineering applications. As a result, the historical and future (model ensemble) results for the reference 1- and 24-hour durations obtained from MLE and Bayesian analysis and are compared. Fig 5 compares the percentage change of MLE historical results with the future results obtained from both the classical MLE and Bayesian analysis. For the 1-hour duration in Phoenix, the MLE future rainfall depths estimates show a percentage gain of up to 56% compared to MLE historical values. However, the Bayesian future estimates show an increase in return levels by up to 80% which is anticipated because the Bayesian inference quantifies uncertainty more effectively as described in the previous section. For the 25-year 1-hour storm, a difference of 11.47 mm (+37%) is observed while for the 100-year 1 hour storm, the difference between the historical MLE and Bayesian future periods increases to 29.84 mm, a percentage gain of 80%. Therefore, the flood risk associated with infrastructure designed for longer return periods is expected to be higher in the future in this region.

For the 24-hour duration, a very large difference across all return periods is observed when comparing MLE historical with Bayesian future results. For the 50% exceedance probability or return period of 2 years, a difference of 16.28 mm or a percentage increase of 62% is observed. Similarly, for the longer return period of 100 years, this difference is as high as 222.42 mm (+398%). Furthermore,

for the 24-h duration, MLE future estimates also show an increase in rainfall depths by about 55% for the return period of 2-years to 336% for the longer 100-year period with respect to MLE historical results. These findings also suggest that longer-duration precipitation events are likely to be more intense in the future at this location. Moreover, the ensemble averaged mean rainfall depth estimates are nearly equal for shorter return periods across both the Bayesian and MLE methods, however, the uncertainty increases as the return period increases. It indicates that the MLE appropriately accounts for model uncertainty, but the Bayesian technique additionally accounts for parametric uncertainty in a more rigorous manner, which is represented in the uncertainty range. The infrastructure's health and safety may be jeopardized if this uncertainty is underestimated. The observed variations in extreme precipitation events also raises concerns about the infrastructure's ability to resist such occurrences in the future. All these findings highlight to the need to include parametric as well as other kinds of uncertainties such as the uncertainty from climate model outputs in extreme rainfall predictions, implying the use of a Bayesian technique that incorporates this uncertainty by assuming prior distribution on model parameters.

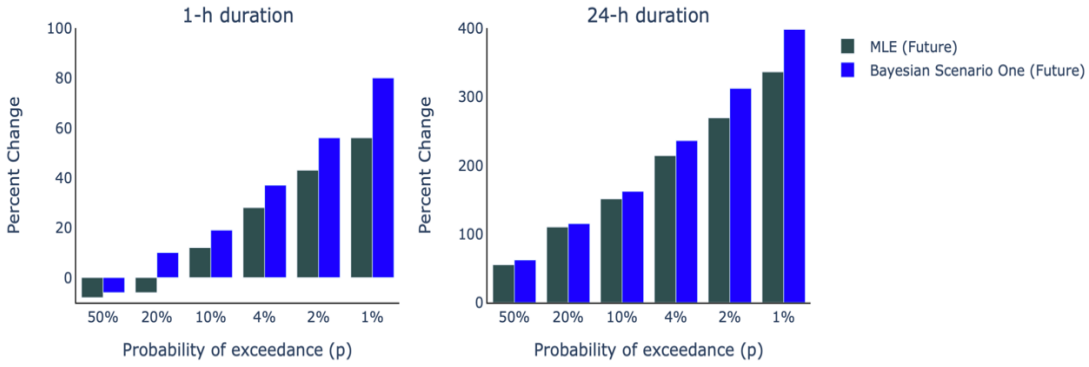

Figure 5: Percent change of Bayesian and MLE future rainfall depth estimates from model ensemble average results with respect to MLE historical values at Phoenix, Arizona, U.S.

4.3 Drainage Pipe Sizing Analysis

This section discusses the effects of IDF values calculated in the preceding sections on stormwater pipe diameters, calculated with the rational method as described in section 3.4. Figure 6 presents the comparison of resulting pipe diameters from both Bayesian and MLE techniques in the historical and future periods.

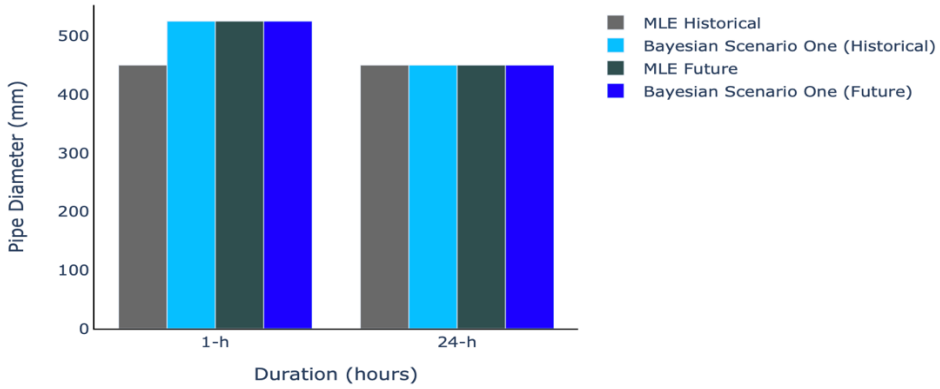

Figure 6: Pipe diameters corresponding to intensities of mean point estimates calculated from MLE and Bayesian inference for 50 year, 1- and 24-h storm events in the historical and future periods.

As a result of explicitly incorporating parametric uncertainty in rainfall intensity estimates with Bayesian inference, a substantial effect is observed on stormwater design choices at the study site. For the 50-year 1-hour storm event, an increase of one pipe size with both Bayesian Historical and future intensity estimates compared to MLE historical estimates is noted. The MLE future also shows an increment of one pipe size compared to MLE historical for the 1-hour duration while for the 24-hour 50-year event the pipe size remains the same across all scenarios. However, for the 1 and 24-hour future storms, Bayesian method also shows an increment of one pipe size in the upper uncertainty interval (see Table 5) to account for climate model uncertainty while with MLE, the diameter remains same across both the lower and upper intervals. This increment in pipe size may be critical for informing disaster risk management and raises concerns about using design values based on traditional statistical methods like the MLE.

The existing design practices put the infrastructure at high risk of not withstanding extreme events, which may ultimately result in severe material damage. In wake of the changing precipitation regimes, there is a need for updating the current design practices of using point estimates calculated from methods such as MLE. As observed in the current study, Bayesian inference is more effective at quantifying parametric as well as climate model uncertainty, implying that engineers and planners must integrate this approach of probabilistic IDF values in their analyses to provide maximal flood protection and to increase the resilience of infrastructure systems to future changes.

Chapter 5

Conclusions and Recommendations

The goal of this thesis is to check the adequacy of urban water infrastructure systems which are typically designed by incorporating precipitation characteristics such as intensity, duration, and frequency in the form of Intensity-duration frequency (IDF) curves. There is, however, a substantial uncertainty surrounding the extreme value distribution parameters usually calculated using classical statistical methods like the Maximum Likelihood estimation (MLE). In this study, the Bayesian method is used to quantify parametric uncertainty in GEV parameters, and the results are compared to classical MLE by fitting the annual maximum series of 1- and 24-hour durations from both the historical data as well as the bias-corrected future rainfall data from RCM forcing's under the RCP 8.5 pathway. The effect of formally incorporating this uncertainty on design decisions is also analyzed by taking a simple example of a stormwater pipe. Hamiltonian Monte Carlo (HMC) method is used to obtain the Bayesian posterior distribution of the parameter set by assuming prior distributions on the parameters and multiplying it with the likelihood function. Bayesian inference involves an additional step of identifying a prior distribution, which can be informative priors based on past research or non-informative, also called as flat priors if no such information is available. Three Bayesian scenarios are developed by varying the prior distribution on the shape parameter of the GEV distribution by assuming Normal, Truncated normal and Beta distributions.

The results show that the prior assumption of a Beta distribution results in higher values of the shape parameter compared to MLE and the other Bayesian scenarios for both the 1- and 24-h durations. This suggests that the choice of commonly used Beta distribution might be overly informative for the chosen study site and should generally be used with caution. However, this can only be confirmed by analyzing data from multiple locations. As a result, flat priors with a large variance devoid of any expert opinion might be the appropriate choice for Bayesian analysis of extreme rainfall series.

The historical return level estimates or rainfall depths are consistent across all the methods for shorter return periods. But, for longer return periods the uncertainty increases especially for the 24-hour duration with the MLE underestimating the tail events compared to Bayesian inference resulting in lower values of the return level estimates. Comparing the future Bayesian scenario one return level estimates with MLE historical values, a large percentage gain is observed after the return period of 25 years. For longer 24-hour duration, this difference is more prominent with a percentage increase of as much as 398% for return period of 100 years. The future MLE results also show an increase in the return level estimates for the 1-h duration by up to 56% compared to the historical MLE values while for the 24-hour duration percentage gain ranges from 55% for the shorter return period of 2-years to as high as 336% for the 100-year return period, which means that the risk associated with infrastructure designed for shorter return periods is comparatively lower than those designed for longer return periods at this location. Furthermore, as a result of quantifying parametric uncertainty, the Bayesian historical analysis alter the size of the stormwater pipe by one increment compared to MLE historical for the 1-hour duration. The future results from both the MLE and Bayesian method also show an increment of one pipe size diameter compared to MLE historical.

Moreover, the pipe size remains same for the 24-hour duration across both the techniques for the historical and future periods. However, Bayesian Inference also showed an increment of one pipe size in the upper interval of future 1- and 24-hour 50-year events, accounting for intermodal uncertainty from climate models, which could be used in the design decisions if additional protection is desired. These findings suggest that infrastructures designed using traditional statistical methods may no longer be able to withstand the extreme events in wake of a changing climate and should therefore be updated to increase the resilience of such infrastructure systems.

As observed in this study, Bayesian method is able to account for uncertainty more effectively by incorporating additional source of information in the form of prior distributions. Therefore, it should be incorporated into current design practices by engineers and planners to produce safe and futuristic designs. Furthermore, the selection of priors is important in Bayesian Inference and should be approached with prudence. It is recommended to use objective choices with large variance if the modeler lacks thorough knowledge about the characteristics of the study site.

While this study presented a comparative analysis of MLE and Bayesian Inference for rainfall extremes, there are several limitations and areas for future work. This study was restricted to a single location. Thus, additional sites can be studied to help researchers better understand the Bayesian Inference method and the resulting probabilistic IDF values and their impact on hydraulic infrastructure. Moreover, only three climate models and one emission scenario (RCP8.5) were considered which may not accurately reflect the actual uncertainty range. As a result, to properly represent model uncertainty, more models and emission scenarios should be incorporated. Additionally, it is recommended to use Reliability Ensemble Average (REA) method (Rupa et al., 2015) to address uncertainty resulting from multiple model simulations and scenarios.

Unfortunately, the quantification of uncertainties is still unpopular in flood risk assessment methods despite having obvious advantages (Cook et al., 2020). This is particularly important in wake of the changing climate patterns. Bayesian inference is no longer computationally complex as a result of advances in computational resources and development of efficient algorithms, and it offers apparent advantages over traditional statistical techniques due to its rigorous nature. Therefore, this method should be made a part of hydrologic infrastructure design and planning.

References

Battaglia, S. M. (2019). No Fairy Tale Ending: The Future of Water and the American Southwest. *Weatherwise*, *72*(6), 36–43. https://doi.org/10.1080/00431672.2019.1659034

Betancourt, M. (2017). *A Conceptual Introduction to Hamiltonian Monte Carlo*. http://arxiv.org/abs/1701.02434

Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York : Springer, [2006] ©2006. https://search.library.wisc.edu/catalog/9910032530902121

Canadian Standards Association. (2012). *Development, interpretation, and use of rainfall intensity-duration-frequency (IDF) information: Guideline for Canadian water resources practitioners*. 214.

Cheng, L., & Aghakouchak, A. (2014). Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific Reports*, *4*, 1–6. https://doi.org/10.1038/srep07093

Coles, S. (2001). An Introduction to Statistical Modeling of Extreme. In *Springer*.

Cook, L. M., McGinnis, S., & Samaras, C. (2020). The effect of modeling choices on updating intensity-duration-frequency curves and stormwater infrastructure designs for climate change. *Climatic Change*, *159*(2), 289–308. https://doi.org/10.1007/s10584-019-02649-6

Courty, L. G., Wilby, R. L., Hillier, J. K., & Slater, L. J. (2019). Intensity-duration-frequency curves at the global scale. *Environmental Research Letters*, *14*(8). https://doi.org/10.1088/1748-9326/ab370a

Hoffman, M. D., & Gelman, A. (2014). The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, *15*, 1593–1623.

Hosseinzadehtalaei, P., Tabari, H., & Willems, P. (2020). Climate change impact on short-duration extreme precipitation and intensity–duration–frequency curves over Europe. *Journal of Hydrology*, *590*(March), 125249. https://doi.org/10.1016/j.jhydrol.2020.125249

Huard, D., Mailhot, A., & Duchesne, S. (2010). Bayesian estimation of intensity-duration-frequency curves and of the return period associated to a given rainfall event. *Stochastic Environmental Research and Risk Assessment*, *24*(3), 337–347. https://doi.org/10.1007/s00477-009-0323-1

IPCC. (2021). Technical Summary. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In *Climate Change 2021: The Physical Science Basis*.

Martins, S., & Stedinger, J. (2000). *Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data*. *36*(3), 737–744. https://doi.org/https://doi.org/10.1029/1999WR900330

McGinnis, S., Nychka, D., & Mearns, L. O. (2015). A New Distribution Mapping Technique for Climate Model Bias Correction. In V. Lakshmanan, E. Gilleland, A. McGovern, & M. Tingley (Eds.), *Machine Learning and Data Mining Approaches to Climate Science* (pp. 91–99). Springer International Publishing.

Mearns, L.O., et al. (2017). *The NA-CORDEX dataset, version 1.0. NCAR Climate Data Gateway, Boulder, CO*. https://doi.org/https://doi.org/10.5065/D6SJ1JCH

Mishra, A. K., & Singh, V. P. (2010). Changes in extreme precipitation in Texas. *Journal of Geophysical Research Atmospheres*, *115*(14). https://doi.org/10.1029/2009JD013398

Neal, R. M. (1995). *Bayesian Learning for Neural Networks*.

NOAA. (2021). *NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters (2021). https://www.ncdc.noaa.gov/billions/*. https://doi.org/10.25921/stkw-7w73

NOAA. (2022a). *NOAA (2022) National Centers for Environmental Information*. http://www.ncdc.noaa.gov/

NOAA. (2022b). *NOAA National Centers for Environmental information, Climate at a Glance: National Time Series, published March 2022*. https://www.ncdc.noaa.gov/cag/

Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions : A global survey on extreme daily rainfall. *Water Resources Research*, *49*(1), 187–201. https://doi.org/10.1029/2012WR012557

Pfahl, S., O'Gorman, P. A., & Fischer, E. M. (2017). Understanding the regional pattern of projected future changes in extreme precipitation. *Nature Climate Change*, *7*(6), 423–427. https://doi.org/10.1038/nclimate3287

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., & Rafaj, P. (2011). RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change*, *109*(1), 33. https://doi.org/10.1007/s10584-011-0149-y

Rupa, C., Saha, U., & Mujumdar, P. P. (2015). Advances in Water Resources Model and parameter uncertainty in IDF relationships under climate change. *ADVANCES IN WATER RESOURCES*, *79*, 127–139. https://doi.org/10.1016/j.advwatres.2015.02.011

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, *2016*(4), 1–24. https://doi.org/10.7717/peerj-cs.55

Sharma, S., Lee, B. S., Nicholas, R. E., & Keller, K. (2021). *A safety factor approach to designing urban infrastructure for dynamic conditions*. https://arxiv.org/abs/2102.04496

Stan Development Team. (2020). Stan Modeling Language User's Guide and Reference Manual, Version 2.19.2. *Interaction Flow Modeling Language*, 1–534. http://mc-stan.org.

Stuart G . Coles and Jonathan A. (2010). *A Bayesian Analysis of Extreme Rainfall Data Author (s): Stuart G . Coles and Jonathan A . Tawn Source : Journal of the Royal Statistical Society . Series C (Applied Statistics), Vol . 45 , No . 4 Published by : Blackwell Publishing for the Royal Stat.* *45*(4), 463–478. doi: 10.2307/2986068

Tapiador, F. J., Navarro, A., Moreno, R., Sánchez, J. L., & García-Ortega, E. (2020). Regional climate models: 30 years of dynamical downscaling. *Atmospheric Research*, *235*, 104785. https://doi.org/https://doi.org/10.1016/j.atmosres.2019.104785

Van de Vyver, Hans, & Demarée, G. R. (2010). Construction of Intensity–Duration–Frequency (IDF) curves for precipitation at Lubumbashi, Congo, under the hypothesis of inadequate data Elaboration des courbes Intensité-Durée-Fréquence (IDF) des précipitations à Lubumbashi, Congo, sous l'hypothèse des . *Hydrological Sciences Journal*, *55*(4), 555–564. https://doi.org/10.1080/02626661003747390

van Rossum, P. development team. (2017). *The Python Language Reference Release 3.6.0* . https://www.goodreads.com/work/best_book/8418412-python-3-reference-manual-python-documentation-manual-part-2

Vu, T. M., & Mishra, A. K. (2019). Nonstationary frequency analysis of the recent extreme precipitation events in the United States. *Journal of Hydrology*, *575*(March), 999–1010. https://doi.org/10.1016/j.jhydrol.2019.05.090

Vyver, H Van De. (2015). Bayesian estimation of rainfall intensity – duration – frequency relationships. *Journal of Hydrology*, *529*, 1451–1463. https://doi.org/10.1016/j.jhydrol.2015.08.036

Westra, S., Alexander, L. V., & Zwiers, F. W. (2013). Global increasing trends in annual maximum daily precipitation. *Journal of Climate*, *26*(11), 3904–3918. https://doi.org/10.1175/JCLI-D-12-00502.1

Willems, P., Arnbjerg-Nielsen, K., Olsson, J., & Nguyen, V. T. V. (2012). Climate change impact assessment on urban rainfall extremes and urban drainage: Methods and shortcomings. *Atmospheric Research*, *103*, 106–118. https://doi.org/10.1016/j.atmosres.2011.04.003

Appendix A

Trace Diagnostics in Bayesian Inference

One of the most crucial tasks in Bayesian data analysis is attaining convergence of Markov Chain Monte Carlo (MCMC) process. It is recommended to use more than one chain so that if one chain fails to converge, others could be used for data analysis. A numerical indicator known as Gelman Rubin diagnostics is commonly employed as a model convergence indicator, and it should ideally approach one. It denotes the variance between chains and a lower number indicates good mixing of chains around the stationary distribution. According to German and Rubin (Stan Development Team, 2020), values larger than 1.2 for any of the model parameters should imply nonconvergence and will require more tuning. The trace diagram for the three GEV parameters is shown in Figure 7. As an example, the figure shows the trace plot of the 1-h duration Annual Maximum Series (AMS) of the study site under the assumption of Bayesian scenario one. The sampling efficiency in the bulk of the distribution is measured by Bulk Effective Sample Size (ESS bulk), whereas the sampling efficiency in the tails of the distribution is measured by Tail Effective Sample Size (ESS tail). The model summary statistics is presented in Table 7. It can be observed that the r_hat, $\hat{r}$ is one and ESS well over the acceptable range of at least 100, implying that the chains have converged successfully which can also be visually confirmed in Figure 7. The lower and the higher limits of the highest density interval or the credible interval is represented by hdi_3% and hdi_97% respectively. The Markov Chain Monte Carlo standard error is represented by mcse_mean and mcse_sd for the mean and standard deviation respectively. It represents the quality of the calculated mean and standard deviation and gives an idea of how long the simulation must be run and to prevent useful samples from being discarded. In other words, it indicates how much of the posterior variation in the model parameters is because of the sampling error.

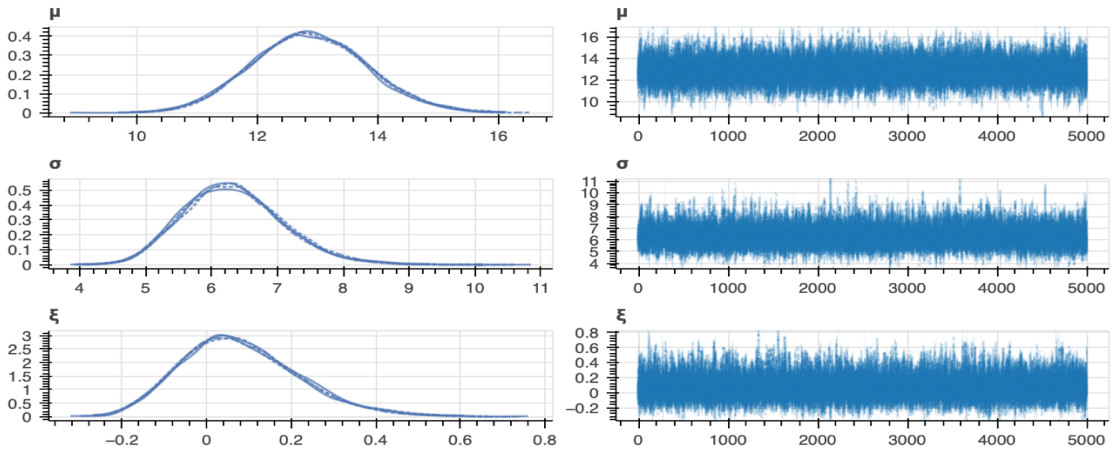

Figure 7: Trace plot of GEV parameter set for 1-hour duration at Phoenix, Arizona, U.S. (generated with PYMC) of the historical data set.

Table 6: Model Summary Statistics of GEV parameter set for 1-hour duration at Phoenix, Arizona, U.S. (generated with PYMC) of the historical period.

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
μ	12.87	0.94	11.13	14.68	0.01	0.01	6686.18	8388.49	1.0
σ	6.33	0.76	4.99	7.79	0.01	0.01	6598.64	8378.71	1.0
ξ	0.08	0.14	-0.17	0.35	0.00	0.00	5965.53	6756.88	1.0

Appendix B

The model simulation results from the future rainfall data of three RCM-GCM combinations used in the analysis are presented in tables below:

Table 7: Return Level estimates of 1- and *24-h duration for different probabilities of exceedance at Phoenix, Arizona, U.S* of RCM-GCM combination *WRF (GFDL- ESM2M)*

Return Period 'T' years	Exceedance Probability (p) p = 1/T	Duration	Return Level (mm/h) MLE	Return Level (mm/h) Bayesian Scenario One
2	0.50 (50%)	1-h	13.10	13.26
		24-h	42.50	44.10
5	0.20 (20%)	1-h	19.83	22.40
		24-h	81.79	84.24
10	0.10 (10%)	1-h	27.0	29.04
		24-h	115.71	119.79
25	0.04 (4%)	1-h	36.0	39.43
		24-h	171.95	178.94
50	0.02 (2%)	1-h	44.0	49.01
		24-h	223	242.10
100	0.01 (1%)	1-h	52.0	60.0
		24-h	290	330.13

Table 8: Return Level estimates of 1- and *24-h duration for different probabilities of exceedance at Phoenix, Arizona, U.S* of RCM-GCM combination *RegCM4(MPI-ESM-LR)*

Return Period 'T' years	Exceedance Probability (p) p = 1/T	Duration	Return Level (mm/h) MLE	Return Level (mm/h) Bayesian Scenario One
2	0.50 (50%)	1-h	15.24	15.77
		24-h	42.38	43.52
5	0.20 (20%)	1-h	24.269	25.59
		24-h	72.23	73.90
10	0.10 (10%)	1-h	31.0	34.03
		24-h	97.37	100.76
25	0.04 (4%)	1-h	44.0	48.62
		24-h	137.12	145.70
50	0.02 (2%)	1-h	56.1	65.04
		24-h	173.83	187.33
100	0.01 (1%)	1-h	69.0	85.06
		24-h	221.0	248.40

Table 9: Return Level estimates of 1- and *24-h duration for different probabilities of exceedance at Phoenix, Arizona, U.S* of RCM-GCM combination *WRF(MPI-ESM-LR)*

Return Period 'T' years	**Exceedance Probability (p) p = 1/T**	**Duration**	**Return Level (mm/h) MLE**	**Return Level (mm/h) Bayesian Scenario One**
2	0.50 (50%)	1-h	13.13	15.68
		24-h	39.08	40.21
5	0.20 (20%)	1-h	21.0	25.13
		24-h	68.58	69.16
10	0.10 (10%)	1-h	27.84	33.45
		24-h	93.11	95.13
25	0.04 (4%)	1-h	37.11	47.30
		24-h	133.91	141.12
50	0.02 (2%)	1-h	45.01	63.63
		24-h	173.34	188.17
100	0.01 (1%)	1-h	53.84	82.33
		24-h	219.81	253.35