FLOOD FREQUENCY ANALYSIS IN CONTEXT OF CLIMATE CHANGE OR WITH MIXED POPULATIONS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Xin Yu

May 2017

FLOOD FREQUENCY ANALYSIS IN CONTEXT OF CLIMATE CHANGE OR WITH MIXED POPULATIONS

Xin Yu, Ph. D.

Cornell University 2017

The thesis addresses two challenges in flood frequency analysis (FFA). The first is how to analyze annual maximum series (AMS) with maxima from two or more distinct processes (e.g. rainfall and snowmelt). The second is how one might incorporate climate change trends into flood risk models.

The mixed-population flood-risk estimators considered include a joint model that includes correlation between rainfall and snowmelt events, a mixture model that treats the two as independent, and an AMS model. The mixture estimator is simple and the most efficient when the complete series of both events are available and the log-cross-correlation is 0.5 or less. When the rainfall distribution dominants the large flood risk, using just the rainfall flood distribution works well. We explore a Kirby-estimator and an Expected Moments Algorithm (EMA) for situations when only the AMS is available. Kirby used the conditional distributions for snowmelt and for rainfall given they are the annual maximum for their year. EMA employs a censored sampling paradigm to represent each data series. EMA generally performs better than the Kirby estimator.

A fundamental assumption of FFA is that flood series are stationary. This thesis evaluates FFA methods that might be used when flood records have trends due to climate change. We consider six estimators. The “Stationary” estimator retains the time-invariance assumption and employs the AMS. Possible methods with time-

varying parameters are represented by 3 estimators: Trend_0 uses the true trends in the AMS mean and variance; Trend_1 estimates the trend in the mean of the log-AMS; Trend_2 estimates trends in both the mean and the variance of the log-AMS. “30-year record” is the “Stationary” estimator using only the most recent 30 years of data. “Safety factor” increases or decreases the 100-year flood estimator by a prescribed percentage. With modest trends ($\leq \pm 0.25\%$ per year), the stationary estimator works well for short records (n=40), but is inferior to Trend_1 with larger trends. With longer records (n=100), Trend_1 performs well for most cases except when the trend in both the mean and variance was ±1%, when Trend_2 is a good alternative. FFA in a dynamic world is a challenge.

BIOGRAPHICAL SKETCH

Xin Yu was born in Zhengzhou, in China's Henan Province, on November 11, 1986. After obtaining her high school diploma from Zhengzhou Foreign Language School in 2005, she attended Zhejiang University in Hangzhou. In 2009, she received a B.S. in Environmental Science from Zhejiang University. In 2008, she entered graduate study via a joint Zhejiang University/Michigan State program, and graduated from Michigan State University with a Master of Public Policy in 2010. After receiving her master's degree, she had internships with the Michigan Environmental Council and then the Michigan Department of Environmental Quality. In spring 2012, she joined the Civil and Environmental Engineering School at Cornell to pursue a Ph.D. in Environmental Water Resources Systems Engineering. During the summer of 2014, she worked as an intern with the U.S. Geological Survey at their National Center in Reston, Virginia. Upon completion of her Ph.D. program, she will assume a position as a consulting data analyst with Axtria in Berkeley Heights, New Jersey. She looks forward with enthusiasm to applying statistics and system engineering to much broader fields.

I dedicate this thesis to my family and my mentors, who made my success possible.

ACKNOWLEDGMENTS

My success of obtaining the doctorate degree won't be possible without my family's support. First, I would like to acknowledge their dedication and generosity toward my education.

At the same time, I would like to acknowledge my academic advisor and special committee chair Professor Jery R. Stedinger, who always encourages me with creative suggestions and constant patience, and is very dedicated to helping me whenever I meet difficulties.

Specially, I would like to acknowledge my internship supervisor Timothy A. Cohn, who initiated one of the research topics in my thesis and inspired me with precious insights in my study field.

Finally, I would like to acknowledge my friends and colleagues, Dr. Jonathon Lamontagne, Dr. Calvin Whealton, and Ph.D. candidate Julianne Quinn, who have always stood by my side in the past five years, and Dr. Jiefei Mao, who has been encouraging me since college.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 2

Appendices for Chapter 2

Chapter 3

LIST OF TABLE

Chapter 2

Appendices for Chapter 2

Chapter 3

Appendices for Chapter 3

CHAPTER 1

INTRODUCTION

Hydrologists use flood frequency analysis to evaluate flood risk for specific locations near a channel. This information can provide parameters for designing infrastructure (dams, levees, etc.), reservoir system operational plans, floodplain zoning, and determination of flood insurance premium. Usually the analysis involves fitting a probability distribution to the peak-flow data (e.g. an annual maximum series), then extrapolating the fitted distribution to estimate the quantile with a target exceedance probability. For instance, the quantile corresponding to 1% annual exceedance probability is defined as 100-year flood. Some widely-used probability distributions for flood frequency analysis include the two- or three-parameter lognormal distribution [Stedinger, 1980], the log-Pearson type 3 (LP3) distribution [Bobée, 1975; Bobée and Robitaille, 1977; Bobée and Ashkar, 1991; Singh, 1998, pp.252-268; Griffis and Stedinger, 2007ab, 2009], and the generalized extreme value (GEV) distribution [Stedinger et al., 1993; Hosking and Wallis 1997]. In United States, the detailed procedures for determining the flood flow frequency are documented in a national guideline Bulletin #17B [IACWD, 1982] and its recent update Bulletin #17C [Draft, 2017].

Fundamental Challenge – Important Questions

The fundamental challenge in flood frequency analysis is to estimate the frequency of large floods, such as the 100-year flood flow value, exceeded with a probability of 1% in any year using short records. In U.S., the common lengths of the annual maximum series are around 30-60 years, and many gauged stations have the records less than 30 years. The short records may not provide sufficient information to accurately estimate the distribution parameters, which results in an even less accurate

quantile estimation with small exceedance probabilities (e.g. 10%, 1%, and 0.2%) that people are interested in. Thus, the estimation for large quantiles such as the 100-year flood is often associated with inevitable errors.

This thesis addresses two issues in flood frequency analysis. The first is the impact of two or more physical processes that may give rise to the largest floods in a year. In Chapter 2 this situation is described as the mixed distribution issue where floods in different seasons perhaps, come from distinctly different processes. Chapters 3 and 4 address another issue, which is on many people's mind: how should frequency analyses reflect possible climate change.

Mixed Distributions

What should hydrologist do if the annual maximum series are consist of events that are generated from two or more distinct physical processes (e.g. rainfall and snowmelt)? This is the so-called mixed population problem.

As described in Bulletin 17B, flooding in some watersheds is created by distinct types of events, which can result in flood frequency curves with abnormally large skew coefficients [IACWD, 1982, pp. 16 and 28] and even kinks [USACE, 1958, 1982; Canfield, 1980; Jarrett, 1982]. Various methods that reflect the distributions of multiple physical processes have been explored to estimate flood risks in these situations by developing different models for the different phenomena [Kite 1977; Waylen and Woo, 1982; Cudworth, 1989; Murphy 2001; Singh et al. 2005].

Chapter 2 discusses such methods and the resultant estimators that can be employed when one has a mixed population of annual maxima. A conceptually attractive approach is to employ a bivariate distribution (joint model) that incorporates interdependence between the two-component series. A simpler mixture model ignores any correlations among the component series assuming they are independent. A

special case of those two models is when the distribution of snowmelt floods has little effect on the distribution of the annual maximum series; thus, the annual maximum rainfall series defines the risk of major flooding and the rainfall-peak-only data set can be used to evaluate flood risk.

However, many records include only the maximum annual flood for each year. For such cases, a method introduced by Kirby [Parrot, personal com., 2011] uses meteorological information with the annual exceedance probability to develop conditional flood risk estimators for both processes; the two combined provide a joint model of the flood risk. Alternatively, the expected moments algorithm (EMA) [Cohn et al., 1997] can estimate the distribution parameters for the mixture model's annual maximum rainfall series and the maximum snowmelt series. Another estimator is developed from a single 3-parameter lognormal distribution to fit the annual maximum series.

Climate Change

It is well understood in scientific circles that we are in a changing world. Scientist and the public are very concerned about the climate change. Scientists talk about an accelerated hydrology cycle that produces more extremes: both floods and droughts. A hotter atmosphere can evaporate more water, and the moisture content of warmer air can be larger. Presentations of flood frequency analysis are met with the question: what about climate change? So, what would be a good strategy to incorporate possible climate change impacts into flood frequency analysis? If such analyses include a climate change mechanism, what will happen to the precision of flood risk estimates?

The fundamental assumption in classical flood frequency analysis is that geomorphologic characteristics, and hydroclimatic and hydrologic statistics for a

watershed, are stable over time. This allows hydrologists to assume available data represent a single time-invariant population of extreme events. However, when the flood risk associated with long-term climate change becomes a concern, the time-invariance assumption of the probability distribution may not be valid anymore. The claim in *Science* of the death of stationarity by Milly et al. [2008] caused much debate [Lins and Cohn, 2011; Matalas, 2012; Koutsoyiannis and Montanari, 2014; Montanari and Koutsoyiannis, 2014; Milly et al. 2015; Serinaldi and Kilsby, 2015]. Despite the complex and controversial answers to this question, Bulletin 17C does encourages hydrologists to incorporate the climate change knowledge when there is sufficient scientific evidence to facilitate quantification of its impacts.

Chapter 3 evaluates different flood frequency estimation methods that can incorporate climate change. Several estimators employ time-varying parameters that can capture trends; other estimators ignore trends, use a limited flood window, or adopt a safety factor (e.g. 25% increase). Basic models are proposed and appropriate parameter estimation algorithm is developed. Data from over 400 sites are analyzed to determine what might be reasonable trends in the log mean and variance. A Monte Carlo re-sampling study considers LP3 estimators of the 100-year-flood some 25 years beyond the end of a 100-year flood record.

Chapter 4 expands upon the study in Chapter 3. Rather than resampling from a single sample that represented a behaved LP3 sample, samples were randomly drawn from LP3 distributions with randomly generated log-space skewness coefficients, given a postulated regional skew value. The Monte Carlo study explores the impact of different skewness coefficients and sample sizes on quantile and exceedance probability estimation in the context of climate change. An improved regression model was developed for estimating the time-varying scale parameter of the LP3 distribution.

The at-site and regional average skewness coefficients were weighted as recommended by Bulletin 17B and 17C.

Result and Conclusion

Finally, Chapter 5 summarizes the results in Chapters 2-4, the conclusions, and provides recommendations for the questions addressed in Chapters 2-4. The chapter also discusses the potential for future research that would help answer the questions addressed by this thesis.

References

Bobée, B. (1975), The Log Pearson Type 3 Distribution and Its Application in Hydrology, *Water Resources Research,* 11(5), 681-689.

Bobée, B., and R. Robitaille (1977), The Use of the Pearson Type 3 and Log Pearson Type 3 Distributions Revisited, *Water Resources Research,* 13(2), 427-443.

Bobée, B., and Ashkar, F. (1991). *The gamma family and derived distributions applied in hydrology.* Water Resources Pubns. ISBN 0918334683.

Canfield, R. V., D. R. Olsen, R. H. Hawkins, and T. L. Chen (1980), *Use of Extreme Value Theory in Estimating Flood Peaks from Mixed Populations,* Reports, Paper 577. Utah Water Research Laboratory, Utah State University.

Cohn, T. A., W. L. Lane, and W.G. Baier (1997), An algorithm for computing moments-based flood quantile estimates when historical flood information is available, *Water Resources Research*, 33(9), 2089-2096.

Cudworth, A.G., Jr. (1989), *Flood Hydrology Manual,* Water Resources Technical Publication, U.S. Bureau of Reclamation, U.S. Dept. of the Interior, Denver, CO, pp. 205-207 and 219-220.

Griffis, V. W., and J. R. Stedinger (2007a), The LP3 distribution and its application in flood frequency analysis, 1. Distribution Characteristics, *J. Hydrol. Eng.,* 12(5), 482-491.

Griffis, V. W., and J. R. Stedinger (2007b), The LP3 distribution and its application in flood frequency analysis, 2. Parameter estimation methods, *J. Hydrol. Eng.,* 12(5), 492-500.

Griffis, V. W., and J. R. Stedinger (2009), The log-Pearson type 3 distribution and its application in flood frequency analysis, 3. Sample skew and weighted skew estimators, *J. Hydrol. Eng.*, 14(2), 209-212.

Hosking, J.R.M. & J.R. Wallis (1997). Regional Frequency analysis. An approach based on L-Moments. Cambridge University Press. ISBN 0-521-43045-3

Interagency Advisory Committee on Water Data [IACWD] (1982), *Guidelines for determining flood flow frequency, Bulletin #17B of the Hydrology Subcommittee,* U.S. Geological Survey, Reston, Virginia.

Jarrett, R. D., and J. E. Costa (1982), Multidisciplinary approach to the flood hydrology of foothill streams in Colorado, Johnson, A. J., and R.A. Clark eds., *International Symposium on Hydrology*: Bethesda, Md., American Water Resources Association, 565-569.

Kite, G. W. (1977, 1988), *Frequency and Risk Analysis in Hydrology,* Water Resources Publications, Littleton, Colorado, USA.

Koutsoyiannis, D., and A. Montanari (2014), Negligent killing of scientific concepts: The stationarity case, *Hydrol. Sci. J.*, 60, 1174–1183, doi:10.1080/02626667.2014.959959.

Lins, H. F., and T. A. Cohn (2011), Stationarity: Wanted dead or alive? *J. Am. Water Resour. Assoc.*, 47, 475–480.

Matalas, N. C. (2012), Comment on the announced death of stationarity, *Water Resour. Plann. Manage.*, 138, 311–312.

Milly, P.C.D., et al. (2008). Stationarity is dead: wither water management? *Science.* 319, 573-574

Milly, P.C.D., et al. (2015). On critiques of "Stationarity is dead: wither water management?" *Water Resour. Res.*, 51, 7785-7789, doi:10.1002/2015WR017408.

Montanari, A., and D. Koutsoyiannis (2014), Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resour. Res.*, 50, 9748–9756, doi:10.1002/2014WR016092.

Murphy, P. J. (2001), Evaluation of mixed-population flood-frequency analysis, *J. Hydrol. Eng.*, *6*(1), 62-70

Singh, V. P., S. X. Wang, and L. Zhang (2005), Frequency analysis of nonidentically distributed hydrologic flood data, *Journal of Hydrology 307,* 175–195.

Serinaldi, F., and C. G. Kilsby (2015). Stationarity is undead: Uncertainty dominates the distribution of extremes. *Advances in Water Resources,* 77:17-36

Singh, V. P. (1988), *Entropy-Based Parameter Estimation in Hydrology*, Springer Science + Business Media, B.V., U.S.

Stedinger, J. R. (1980), Fitting log normal distributions to hydrologic data, *Water Resources Research, 16*(3), 481-490.

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou, Frequency Analysis of Extreme Events, Chapter 18, *Handbook of Hydrology*, D. Maidment (ed.), McGraw-Hill, Inc., New York, 1993.

U.S. Army Engineer District, Corps of Engineers (1958), Research Note No. 1, *Frequency of New England Floods.*

U.S. Army Corps of Engineers (1982), *Mixed-population frequency analysis*, Hydrologic Engineering Center, Institute for Water Resources, Davis, CA.

Waylen, P., and M-K. Woo (1982), Prediction of Annual Floods Generated by Mixed Processes, *Water Resources Research, 18*, 1283-1286.

CHAPTER 2

FREQUENCY ANALYSIS OF MIXED POPULATION FLOOD SERIES

Abstract

A challenge in flood frequency analysis is how best to derive the distribution of the annual maximum peak-flows when they arise from distinct processes, such as snowmelt and rainfall events. A conceptually attractive approach is to employ a bivariate distribution (joint model) that incorporates interdependence between the two-component series. A simpler mixture model ignores correlations among the component series and often performs well. A special case of those two models is when the distribution of snowmelt floods has little effect on the distribution of the annual maximum series (AMS); thus, the annual maximum rainfall series defines the risk of major flooding (rainfall-only model). However, many records include only the maximum annual flood for each year. For such cases, the Kirby method uses meteorological information with the AMS to develop conditional flood risk estimators for both processes; the two combined provide a joint model of the AMS risk. Alternatively, the expected moments algorithm (EMA) can estimate the distribution parameters for the mixture model's annual maximum rainfall and maximum snowmelt series. Our sixth estimator is developed from a single 3-parameter lognormal (LN3) distribution that fits the AMS. Monte Carlo simulation results show that the simple estimator of mixture model provides a reasonable estimation for large flood quantile when the log-space cross correlations (ρ) between rainfall and snowmelt floods are less than 0.5. Correlation has modest effects on the accuracy of quantile estimators. The rainfall-only estimator is a reasonable alternative when rainfall events dominate the annual flood risks. If only the annual maximum series is available, for 100-year flood estimation EMA performs better than the Kirby and single LN3 estimators when

ρ is modest (less than 0.5). With climate change increasing in importance, mixed models allow climate change to have different impacts on different flood mechanisms, as might be expected physically.

Table of Contents

1. Introduction

Accurate estimation of peak flood quantiles is important for flood risk assessment. Hydrologists employ appropriate models to describe the annual maximum series, such as the log-Pearson III (LP3) or the lognormal distributions, both of which have been proven effective in hydrology frequency analysis [Stedinger, 1980; IACWD, 1982]. However, an annual maximum series composed of events that arise from different processes may not be well described by a single LP3 or lognormal distribution. As described in the National Guidelines for Determining Flood Flow Frequency (Bulletin 17B), flooding in some watersheds is created by different types of events, which can result in flood frequency curves with abnormally large skew coefficients [IACWD, 1982, pp. 16 and 28] and even kinks [USACE, 1958, 1982; Canfield, 1980; Jarrett, 1982]. In some situations, the frequency curve of annual events can best be described by computing separate curves for each type of event. In the ASCE *Hydrology Handbook* [Task Committee on Hydrology Handbook, 1996, p.490], two questions are asked: When is it advisable to model several different component flood series separately? And when is it just as reasonable to model the composite annual maximum series directly?

To define a mixed population, assume an annual maximum series Q can be viewed as the maximum of the maximum rainfall event R and the maximum snowmelt event S:

$Q_t = max \{R_t, S_t\}$.

Here R and S may be classified by a rigidly specified calendar period, a loosely defined season (e.g. the rainy winter and relatively dry summer in California) [Watt, 1989; Lecce, 2000; Veilleux, 2011; Lamontagne, 2012], climate anomalies (e.g. El Nino or La Nina oscillations) [Alia and Mtiraoui, 2002; Escalante–Sandoval, 2007],

the weather records [Murphy, 2001ab], or physical characteristics reflected on hydrographs [Elliot et al. 1982; Hirschboeck, 1987].

To illustrate the potential value of a mixed distribution model, consider the 27-year flood record of South Platte River in Colorado (1926-1953, the record in 1951 is missing) [Elliott, et al. 1982]. The gage site is located at North Saint Vrain Creek at Longmont Dam (near Lyons). The drainage area of the watershed is 274.5 km^2 and the gage datum is 1853 m high above mean sea level. The rainfalls event and snowmelt events were classified by hydrograph and season. Snowmelt-runoff peaks have a hydrograph with a diurnal pattern and daily peaks rise and fall over a period of several days or weeks. Rainfall affects streamflow much more quickly and such events are less regular in occurrence than meltwater from the snowpack [Elliott, et al., 1982]. Table 1 lists the parameters of lognormal distributions fit to the annual maximum rainfall and maximum snowmelt series. The real-space mean and standard deviation are denoted as μ and σ; the log-space mean and standard deviation are denoted as μ_L and σ_L.

Figure 1 shows the probability density functions (PDFs) for the rainfall, snowmelt, and annual maximum distribution given the lognormal parameters in Table 1. The PDF of rainfall flood series has a larger variance and heavier upper tail than the snowmelt flood series; this means almost all the largest floods (e.g. 50-year or 100-year flood) were caused by rainfall events, despite the fact that snowmelt flood series has the larger mean and median. On the other hand, the PDF of snowmelt flood series has a tighter distribution and snowmelt floods appear to always exceed 5 m^3/s. Overall, the snowmelt flood series has a smaller standard deviation, mainly because the snow-pack melting process depends on a local energy balance (i.e. the sunlight and the temperature), while the rainfall events can be extremely large when extreme rainfall events are experienced. The annual maximum distribution shows a low

threshold that is consistent with the lower bound of the snowmelt distribution, and an upper tail that almost overlaps with the PDF curve of rainfall floods. The example shown in Figure 1 is not a special case. Flood records with the similar pattern of PDF curves (as in Figure 1) can be found at other gauge stations in Colorado [Elliott, et al. 1982; Jarrett and Costa, 1982] and British Columbia [Waylen and Woo, 1982].

Table 1 Real Space (μ, σ) and log-space (μ_L, σ_L) parameters of LN2 Distribution
for annual peak flows from 1926 to 1953 at North Saint Vrain Creek, CO (m^3/sec)

Moments	Rainfall	Snowmelt
μ	15.10	16.84
σ	11.53	4.51
μ_L	2.45	2.78
σ_L	0.73	0.29

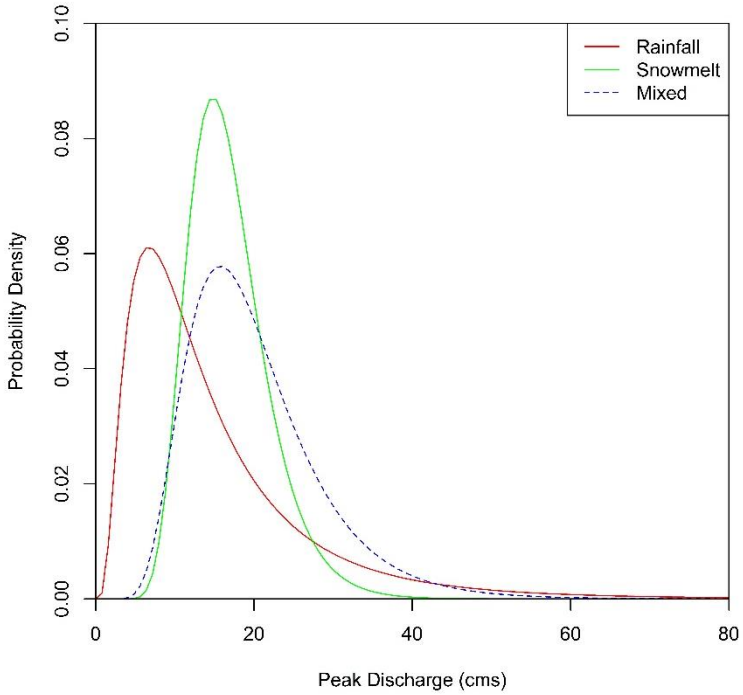


Figure 1 Probability Density Functions of Rainfall, Snowmelt, and Annual Maximum Flood Series for annual peak flows from 1926 to 1953 at North Saint Vrain Creek, CO

Climate change is an increasing concern because of its potential impact on flood-flow frequency relationships. Mixed floods models would appropriately allow climate change to have different impacts on floods caused by different mechanisms, such as summer convective storms, hurricanes, major winter storms, or snowmelt events. It is reasonable to expect that global warming would have different impacts on the distribution of floods from different sources [Kidson and Richards 2005; Smith at el. 2011].

This paper focuses on different methods for frequency analysis with mixed population flood series. Section 2 introduces various methods people have used to model maximum flood series that result from multiple physical processes. Using the framework in Lu [2013], this paper considers six different estimators that make use of the annual maximum series in different ways, or build a model of the annual maximum series using complete records of the individual component series. Those estimators are described in Section 3. Section 4 reports a Monte Carlo study that evaluates the accuracy and precision of all the 6 estimators for 10-year and 100-year estimation. Based on the Monte Carlo results, Section 5 provides recommendations for which estimator is likely to be best for practical hydrologic investigations under specific situations.

Stedinger [2000, Figure 12.1] notes that flood frequency estimates at a site can often be improved by regional or historical information. None of our experiments make use of regional information. However, there is no reason that regional information on snowmelt or on rainfall floods could not be employed using any number of regional estimation procedures [Stedinger and Lu, 1995; Hosking and Wallis, 1997]. Similarly parameter estimation could employ historical information that is available at a site [Stedinger and Cohn, 1986]. If one knows that over the last T years, some threshold Q_T was not exceeded by an annual maximum, then it must be

the case that the annual maximum rainfall and the annual maximum snowmelt events in those years were less than Q_T. However, if historical information shows that in year t, the exceptionally large historical maximum was a rainfall event of magnitude $Q(r_t)$, then we also know the annual snowmelt event in that year was less than $Q(r_t)$.

2. Studies Using Models of Mixed Flood Series

Various methods that reflect the distributions of multiple physical processes have been developed to estimate flood risks. As early as 1958, the U.S. Army Corps of Engineers separated hurricane and non-hurricane floods from a single series in order to understand the physical explanation of unusual flood-frequency curves. Assuming the hurricane and non-hurricane flood events are independent [USACE, 1958, 1982], they developed a mixture representation of the flood risks due to both hurricane and non-hurricane events. Let P_H be the exceedance frequency of annual maximum hurricane floods, and P_N be the exceedance frequency of annual maximum non-hurricane floods. Then P_Q the exceedance frequency of annual maximum floods of either type can be computed using:

$$P_Q = P_H + P_N - P_H P_N \qquad (1.1)$$

Jarrett and Costa [1982] employed that same mixture model to generate flood-frequency curves for 69 stations in Colorado. They compared curves for low- and high-elevation stations. Results showed that the flood frequency distributions have different coefficient of variations and skew depending upon elevation in Colorado foothill streams. Snowmelt floods are more frequent annual maximums above about 7,500 feet and rainfall-produced floods are more frequent below about 7,500 feet. The Flood Hydrology Manual [Cudworth, 1989] discussed both the meteorological (rainfall and snowmelt) and hydrological (infiltration, cover, channel roughness etc.) factors that can cause mixed populations. The manual also expanded the mixture

presentation to describe mixed populations when more than two mutually independent processes compete to be the annual maximum.

Kite employed implicitly the same mixture representation to carry out frequency analysis for floods caused by precipitation and snowmelt [1977, pp. 6-13]. He computed the annual maximum distribution as a product of non-exceedance probabilities of two independent flood series, which is just the complement of Equation (1.1). Let F_P be the non-exceedance probability due to rainfall, and F_S be the non-exceedance probability due to snowmelt. F_Q the non-exceedance probability in any year is calculated as:

$$F_Q(q) = F_P(q)\, F_S(q) \qquad (1.2)$$

Waylen and Woo [1982] employed the precipitation information to identify the flood-generating processes in British Columbia. The snowmelt and rain generated annual floods are modeled by Gumbel distribution individually, and these cumulative distribution functions (CDFs) are combined using Equation (1.2) which provided a good fit to the annual flood series.

The pursuit of representations of the distribution of the annual maximum series also employed models that were a weighted sum the respective non-exceedance or exceedance probabilities for the component series. U.K. Natural Environment Research Council (NERC) in their Flood Studies Report [NERC, 1975, pp. 57-58, 543] recognized the possible advantage of sorting data according to season or climatic origins and studying the different series separately; they represented the distribution of the annual maximum using the equation:

$$F_Q(q) = p_1 F_1(q) + p_2 F_2(q) + \ldots + p_n F_n(q) \qquad (1.3)$$

Subject to $p_1 + p_2 + \ldots p_n = 1$

Here n is the number of component distributions in the mixture; they defined p_i to be the probability that a single event drawn at random from the mixed population is from distribution F_i. In the special case of two causes, Equation (1.3) reduces to

$$F_Q(q) = p_1 F_1(q) + p_2 F_2(q) \qquad (1.4)$$

which is very different than Equation (1.2).

The logic given by NERC to support Equations (1.3) and (1.4) seems correct, but that is not necessarily so – it depends upon how the F_i are defined. Equations (1.3) and (1.4) yield the correct results if F_i are the *conditional* distributions of q_i given that q_i was the year's annual maximum. Thus, to be clear for Equations (1.3) – (1.4) to be correct, F_i should be written as:

$F_i (q_i \mid q_i > q_j \text{ for all } j \neq i)$.

From the example in NERC, it seems clear that they envisioned that F_i in Equation (1.3) to be the unconditional distribution of the annual maximum of events in category i. This would result in an incorrect result as a simple example shows. If F_i is the unconditional distribution for the annual maximum of component series *i*, the correct expression is given by Equation (1.2).

To illustrate the problem with Equations (1.3) and (1.4), consider two components representing snowmelt and rainfall floods, wherein snowmelt always has a value of 100, and rainfall flood is drawn from $N(100, 10^2)$. The distribution of the annual maximum will be 100 with probability 0.5 (because the probability that Snowmelt < Rainfall is 50%), and a half normal distribution with lower bound of 100 with probability 0.5. Just averaging the two CDFs using Equations (1.4) clearly gives the wrong result because it specifies that 25% of the time the annual maximum will be less than 100 (which never occurs). This example also illustrates how different the distribution of the annual maximum for a component can be from the distribution of those component values are that are annual maximums for the watershed: in the

example above the first distribution is $N(100, 10^2)$, whereas the second is only those values greater than 100.

Others have used Equation (1.3) incorrectly. Using lognormal distributions for the F_i, Alila and Mtiraoui [2002] showed that frequency models that explicitly account for floods generated by a mixture of two or more populations are both hydrologically more appropriate and statistically consistent with the LN distribution. Using Equation (1.4), Escalante-Sandoval [2007] derived a mixed Gumbel (F_1) – GEV (F_2) distribution, one of which describes the "regular" floods, and the other describes the floods affected by extreme climatic factors such as El Nino/La Nina oscillations.

Other research has employed a different and correct approach that utilizes the conditional distributions. Murphy [2001a, 2001b] expanded the single lognormal distribution to mixed populations by weighing the conditional exceedance probabilities for floods due to different causes (tropical cyclone, ice-jam-release, and "ordinary"):

$$P_Q(q) = p(o)\, P_O(q|o) + p(t)\, P_T(q|t) + p(i)\, P_I(q|i) \qquad (1.5)$$

where P_Q is the total annual exceedance probability; $p(o)$, $p(t)$, and $p(i)$ are the probabilities that the annual flood will be of the ordinary, tropical cyclone, and ice-jam types; particularly important, $P_O(q|o)$, $P_T(q|t)$, and $P_I(q|i)$ are respective conditional probabilities that the specific type of flood will have a peak-flow rate exceeding q when that type is the annual maximum. Equivalently, if one uses conditional non-exceedance probabilities, Equation 1.5 can be written

$$F_Q(q) = p(o)\, F_O(q|o) + p(t)\, F_T(q|t) + p(i)\, F_I(q|i) \qquad (1.6)$$

Equivalence is easily demonstrated by the substitution $F_J(q|j) = 1 - P_J(q|j)$

Murphy's method doesn't assume flood series from different physical processes are independent, which is convenient. The results show that the mixed-population analyses with local regional skewness gave statistics (from probability-plot

correlation-coefficient tests) that marginally improved the quality of fit of the single-population, log-Pearson III distributions. Singh et al. [2005] expresses the frequency distribution of the annual maximum flood by a similar approach, except that the subpopulations of the flood record were categorized by seasons. He also observes that this method requires having the whole series of annual maxima for each component.

Kidson and Richards [2005] discussed the effect of different flood generating mechanisms on the power law model for flood frequency analysis. They mentioned that modeling the component flood series separately based on their generating mechanisms reduces the sample size for the fitting of individual conditional PDFs if one splits up the annual maximum by component as required when using Equations 1.5-1.6, and thereby makes the use of multi-parameter extreme event distributions more difficult due to limited sample sizes. They also suggested that there may be better theoretical grounds for assuming that the annual flood record is more likely to be generated by mixed, simple distributions (lognormal or power law) than by single, complex ones.

3. Development of Mixed Flood Series Estimators

For the estimation of flood risk at sites with mixed populations, we consider here three models denoted Mixture, Joint, and Annual Maximum. For these models, we consider three categories of parameter estimators, corresponding to use of the annual maximums for each component series R and S, use of only the annual maximum series Q with/without information as to which component yield the annual maxima in each year, and use of the only the rainfall maximum series (Table 2).

Simple estimators of the parameters of the first two models (Mixture, Joint) require that the annual maximums for each series in each year are available. Rainfall-only estimator uses only the rainfall series, which is a special case of the mixture

model. The third category of estimators – Kirby, EMA and Single LN3, employ only the annual maximum series. Kirby and EMA require that one knows for each year in the annual maximum series from which population the peak came, in our case rainfall and snowmelt. Kirby and EMA use that extra information to build models of the two flood populations, even though the entire annual maximum series for each population is not available. Finally, the parameter estimator of the Annual maximum model ignores the mixed population issue and just fits a single three-parameter lognormal distribution to the series of annual maximum floods. Fitting a single distribution to the annual maximum series is of course a legitimate approach. The issue is if one can get more accurate and reliable flood risk estimators by recognizing that the annual maximum series arises from two or more physically distinct meteorological/hydrological processes that can be modeled by a well-behaved 2-parameter distribution (lognormal in this study).

Sections 3.1 – 3.2 present the simple estimators of the parameters of mixture and joint models for describing the joint and marginal distributions of rainfall and snowmelt floods in a watershed. The mixture model is a special case of the joint model when the two series are assumed to be independent. Section 3.3 introduces the 2-parameter lognormal estimator for rainfall-only series which ignores the impact of snowmelt; this is appropriate at sites where the annual maximum flood risk is dominated by rainfall floods and one has a record of the annual maximum rainfall series.

Table 2. Models and parameter estimators for describing mixed population flood risk

	Model 1 Mixture	**Model 2 Joint**	**Model 3 Annual maximum**
Rainfall Series	Rainfall-only		
Tow Complete Series	Mixture	Joint	
Annual Maximum Series	EMA	Kirby	Single LN3

While the mixture and joint models start by describing the marginal distributions of the rainfall and snowmelt maximum series, the Kirby method starts by describing the conditional distribution of rainfall maxima that are also annual maximum, and snowmelt maxima that are also annual maximum. Section 3.1 and 3.2 derive the annual maximum distribution and the Kirby conditional distributions from the mixture and joint models, respectively. Section 3.4 derives of the annual maximum distribution from the Kirby framework.

The Expected Moments Algorithm (EMA), presented in Section 3.5, also employs the annual maximum series and the information of the physical flood process. However, it uses that information with the idea of censored sampling to estimate the parameters of the unconditional distribution of the annual maximum series of the individual rainfall and snowmelt series. Finally, Section 3.6 describes the single 3-parameter lognormal estimator for the Annual maximum model. Section 3.7 provides conclusions.

In this paper, we made use of the lognormal distribution. The lognormal distribution is widely used to describe annual maximal discharges. This idea appears to have been introduced into hydrologic practice by Horton [1914]. Hazen [1914] was probably the first to state explicitly that if the logarithms of the numbers representing the several floods are used, the agreement with the normal probability curve is closer than without such a transformation. This paper adopts the two- and three-parameter lognormal distributions (LN2 and LN3) to fit the mixed population Q. Parameter estimation for the lognormal distribution is well studied, with results provided by Wilson and Worcester, 1945; Cohen, 1951, 1980; Aitchison and Brown, 1957; Stedinger, 1980; Johnson et al. 1994]. The further discussion on the differences between LN2 and LN3 is in section 3.3.

The relationships between the real-space and log-space moments are relatively simple. Let μ and σ^2 be the real-space mean and variance of the two-parameter lognormal distribution. Similarly, let μ_L and σ^2_L denote the log-space mean and variance. The relationships between the two is as follows:

$$\sigma_L^{\,2} = \ln\left(1+\frac{\sigma^2}{\mu^2}\right) \qquad (3.1)$$

$$\mu_L = \ln(\mu) - \frac{1}{2}\sigma_L^2 \qquad (3.2)$$

3.1 Mixture Model

The mixture model corresponds to the simple situation where the magnitudes of rainfall and snowmelt maxima are statistically independent [Kite, 1977; Waylen and Woo, 1982]. Then the CDF for Q is:

$$F_Q(q) = P\{Q < q\} = P\{ R < q \cap S < q\} = F_R(q)\, F_S(q) \qquad (3.3)$$

For rainfall and snowmelt events whose maximum discharge series has a lognormal distribution, one obtains that

$$F_Q(q) = F_R(q)F_S(q) = \Phi\left[\frac{\ln(q)-\mu_{RL}}{\sigma_{RL}}\right]\Phi\left[\frac{\ln(q)-\mu_{SL}}{\sigma_{SL}}\right] \qquad (3.4)$$

where $\Phi(\bullet)$ is the CDF of the standard normal distribution.

The mixture model assumes the cross-correlation between rainfall and snowmelt events is zero. In reality, the same-year rainfall and snowmelt floods for a watershed are often correlated due to interdependence seasonal weather, and soil moisture and perhaps groundwater levels. However, the mixture model should still a reasonable estimator for the mixed population, provided the cross-correlation is not too large. We are also concern about the situation when the cross-correlation is very large (perhaps greater than 0.9). Section 3.2 introduces the "joint model" that includes the correlation between the two series. This will allow the evaluation of the impact of cross-correlation on computed values of $F_Q(q)$.

3.2 Joint Model

When the annual maximum rainfall and snowmelt floods are correlated, the CDF of the mixed population can be computed from the joint distribution of rainfall and snowmelt floods using:

$$F_Q(q) = P\{R < q, S < q\} = \int_{-\infty}^{q}\int_{-\infty}^{q} f_{R,S}(r,s)drds \qquad (3.5)$$

To simplify the calculation, consider using the conditional distribution for R given S. The double integration becomes:

$$F_Q(q) = \int_{-\infty}^{q}\int_{-\infty}^{q} f_{R|S}(r \mid s)\, f_S(s)drds = \int_{-\infty}^{q}\left[\int_{-\infty}^{q} f_{R|S}(r \mid s)dr\right] f_S(s)ds = \int_{-\infty}^{q} F_{R|S}(q \mid s)\, f_S(s)ds \qquad (3.6)$$

This requires the conditional CDF of R given S. We employ the analytic conditional CDF of R given S and then integrate over S. Because the PDF of R has a heavy tail, numerical integration over R will struggle to get accurate answers.

If R and S have a joint lognormal distribution, then the conditional distribution of ln(*R*) is simply

$$\{\ln(R) \mid S = s\} \sim N\left[\mu_{RL} + \sigma_{RL}\rho\sigma_{SL}^{-1}\left(\ln(s) - \mu_{SL}\right), (1-\rho^2)\sigma_{RL}^2\right] \quad (s > 0) \qquad (3.7)$$

where ρ is the correlation between the log-space rainfall and log-space snowmelt variates. More explicitly, the conditional CDF of R given S is:

$$F_{R|S}(r \mid s) = \Phi\left\{\frac{\ln(r) - \left[\mu_{RL} + \sigma_{RL}\rho\sigma_{SL}^{-1}\left(\ln(s) - \mu_{SL}\right)\right]}{\sqrt{(1-\rho^2)\sigma_{RL}^2}}\right\} \quad (r > 0,\ s > 0) \qquad (3.8)$$

For the joint model, a numerical method (composite Simpson's rule), is employed to compute $F_Q(q)$.

Using the same annual peak flow statistics in Table 1 for the North Saint Vrain Creek in Colorado with Equation 3.6, yields Figure 2 showing the resulting CDF for the annual maximum series with different log-space correlations.

The correlation between the rainfall and snowmelt series reflects the meteorological and physical process in a watershed. Consider the Sierra Nevada

Mountains in California as an example; the floods caused by rainfall events usually occur during the winter, while the floods result from snowmelts usually occur in spring [Lamontagne, 2012]. Another example we use in this chapter involves the rainfall and snowmelt flood records in Colorado, which are categorized by U.S. Geological Survey. Snowmelt floods usually result from the seasonal ablation of the snow pack, which occur mostly between April and June; while the rainfall floods usually occur during the late spring or summer (due to a convective storm or frontal system) when thunderstorm activity is greatest [Elliott, et al., 1982]. The differences in physical mechanisms and seasons make it unreasonable to see a correlation as high as 0.9.

the common correlations between these two events range from 0 to 0.5. Thus $0 \leq \rho \leq 0.5$ is a physically reasonable range, while $\rho = 0.9$ is an extreme value included to show what can happen.

Overall, with the joint model and the parameters considered, the correlation between rainfall and snowmelt events does not have noticeable impacts on the CDF of annual maxima, unless the log-space cross-correlation is very close to 1. In the case that $\rho = 1$, the CDF of mixed population will be identical with the CDF of snowmelt flood series below the intersection in Figure 2 and the CDF of rainfall flood series above the intersection. Even more significant, for all correlations $\rho \leq 0.9$, the correlation had almost no impact on the distribution of flows larger than the 90 percentile – the rainfall distribution determines the risk of large floods. While the appropriate parameters of the joint model will vary regionally and from watershed to watershed, these results generally should be widely applicable.

In the case that the value of one source of flood risk determines the value of the other ($\rho = 1$), the CDF of the maximum is simply

$$F_Q(q) = min\{ F_R(q), F_S(q) \}$$

because 1-$F_S(q)$ is the risk of flooding below their intersection in Figure 2, and 1-$F_R(q)$ is larger and determines the risk of flood above their intersection; equivalently, in terms of quantiles, $Q_P = max\{ R_P, S_P \}$.

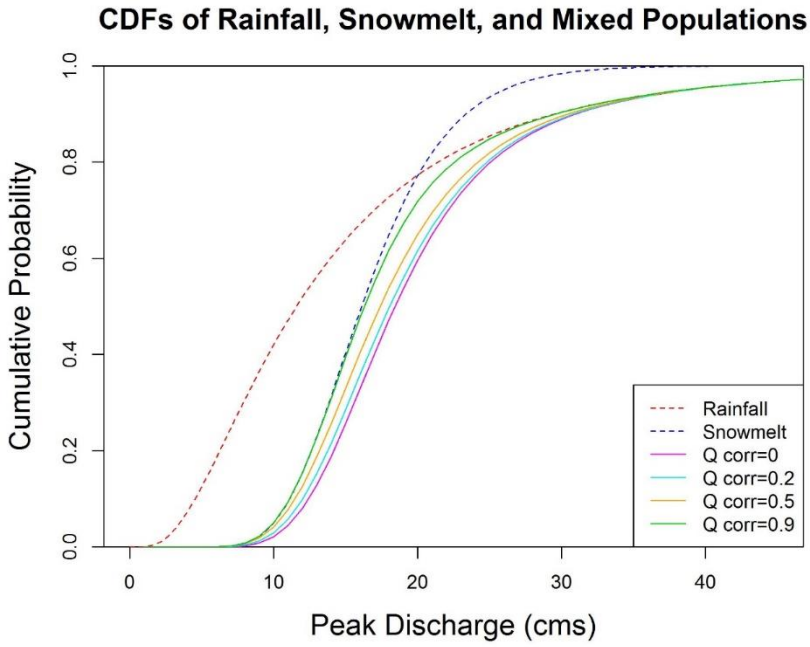


Figure 2 Cumulative Distribution Functions of rainfall, snowmelt, and mixed population based upon annual peak flows from 1926 to 1953 at North Saint Vrain Creek, CO.

3.3 Rainfall-only Model

A special case of both the mixture model and the joint model is when the distribution of snowmelt floods has no effect or little effect on the distribution of the annual maximum series. Figure 2 above provides an illustration of such an effect for quantiles with flood less than 30 cms corresponding to non-exceedance probabilities less than 90%. In such cases, the distribution describing the mixed population for events which have an annual probability of occurrence less than 20% is determined by just the CDF of the rainfall flood series. Thus, our third model, called rainfall-only,

uses only the annual maximum rainfall series to describe the risk of flooding, neglecting the occurrence of snowmelt events.

$$F_Q(q) = F_R(r) = \Phi\left[\frac{\ln(r) - \mu_{RL}}{\sigma_{RL}}\right] \qquad (3.9)$$

3.4 Kirby Method

Theoretically, a joint distribution is an ideal for computing the annual maximum flood risk. It would be easy to fit a joint model if the individual rainfall and snowmelt series (R and S) are both available. However, in practice, hydrologists often have only data for the mixed population (Q). A clever method developed by William Kirby (USGS, Reston VA) addresses this situation.

3.4.1 Theory and Parameter Estimate

Kirby method fits conditional distributions to rainfall maxima and snowmelt maxima that are also the annual maxima for their years; it then weights the two conditional distributions by their probabilities to compute the CDF of annual maximum series [Parrot, personal com., March 2011; Lu, 2013]. To describe the Kirby method, let:

R’ = rainfall maxima that are also annual maxima in their years

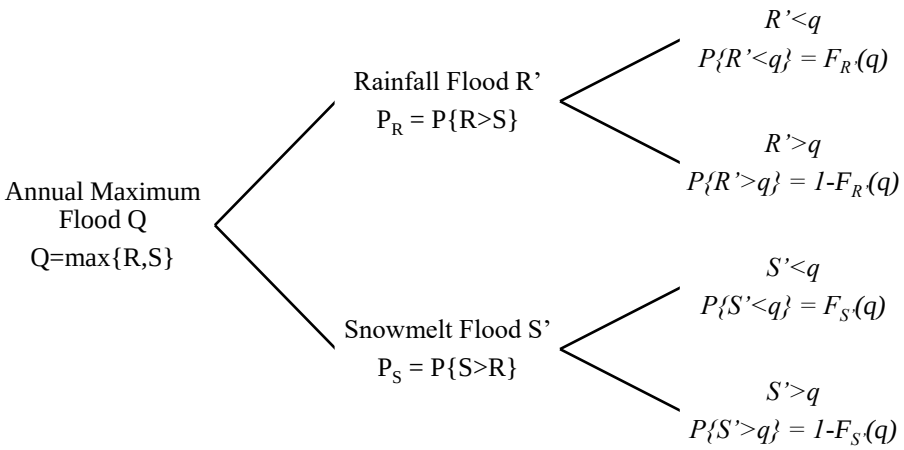
S’ = snowmelt maxima that are also annual maxima in their years

P_R = probability that rainfall maximum is also annual maximum = P{R > S}

P_S = probability that snowmelt maximum is also annual maximum = P{S > R}

$P_S = 1 - P_R$

The key conceptual relationships for Kirby method is described by the event tree below:



Analysis of the event tree yields

$$F_Q(q) = P\{Q<q\} = P\{R<q \cap R>S\} + P\{S<q \cap S>R\} = P_R F_{R'} + P_S F_{S'} \qquad (3.10)$$

where

$$F_{R'} = P\{R'<q\} = P\{R<q \mid R>S\} \qquad (3.11)$$

$$F_{S'} = P\{S'<q\} = P\{S<q \mid S>R\} \qquad (3.12)$$

Given a mixed population Q, it is easy to estimate P_R and P_S from the observations.

$$\hat{P}_R = \frac{n_r}{n} \qquad (3.13)$$

$$\hat{P}_S = \frac{n_s}{n} = 1 - \hat{P}_R \qquad (3.14)$$

where n_r is the number of rainfall floods that are also the annual maxima in {Q}, n_s is the number of snowmelt floods that are also the annual maxima in Q, and $n = n_r + n_s$.

3.4.2 Illustration of Conditional Distributions

One concern is that, R' and S' are the values extracted from complete rainfall and snowmelt floods population, which may have appreciably different distribution types from the one used for R and S.

In order to test if LN2 is still a valid distribution for R' and S', the PDF of R' and S' need to be calculated first. Let R, and S have a 2-parameter (μ and σ) lognormal distributions.

$$f_{R'}(q) = \frac{d\left[F_{R'}(q)\right]}{dr} \qquad (3.15)$$

As stated above, $F_{R'}$ is the conditional probability that rainfall maxima is also the annual maxima for their years. Utilizing the relationship between conditional probability and joint probability, $F_{R'}$ is obtained as follows:

$$\begin{aligned} F_{R'}(q) &= P\{R' < q\} = \frac{P\{R < q \cap R > S\}}{P\{R > S\}} \\ &= \frac{1}{P_R}\int_{-\infty}^{q}\int_{s}^{q} f_{R,S}(r,s)\,dr\,ds = \frac{1}{P_R}\int_{-\infty}^{q}\left[\int_{s}^{q} f_{R|S}(r \mid s)\,dr\right] f_S(s)\,ds \end{aligned} \qquad (3.16)$$

Taking derivative with respect to *r*, the PDF of R' is represented as:

$$f_{R'}(q) = \frac{1}{P_R}\int_{-\infty}^{q} f_{R|S}(\mathrm{q} \mid s)\, f_S(s)\,ds \qquad (3.17)$$

The same procedures are applied to the PDF of S':

$$f_{S'}(q) = \frac{1}{P_S}\int_{-\infty}^{q} f_{\mathrm{S|R}}(\mathrm{q} \mid \mathrm{r})\, f_R(\mathrm{r})\,dr \qquad (3.18)$$

Similar to what have been done in Section 3.2 (Equation 3.6), P_R can be derived from the integral of probability density functions (PDF) and conditional probability functions:

$$\begin{aligned} P_R &= P\{R > S\} = \int_{-\infty}^{\infty}\int_{s}^{\infty} f_{R,S}(r,s)\,dr\,ds \\ &= \int_{-\infty}^{\infty}\left[\int_{s}^{\infty} f_{R|S}(r \mid s)\,dr\right] f_S(s)\,ds = \int_{-\infty}^{\infty}\left[1 - F_{R|S}(s \mid s)\right] f_S(s)\,ds \end{aligned} \qquad (3.19)$$

All the integrations above have analytical solutions determined by μ_{RL}, σ_{RL}, μ_{SL}, σ_{SL}, and ρ. Since the equations are extremely complicated, numerical method (Simpson's rule) used in Section 3.2 is employed here again.

Figure 3 (A-D) display the probability plot for R, R', S, and S' with the log-space cross-correlation of 0 and 0.9. Comparing with the CDF of R when fitting a LN2 distribution, the CDF of R' has a larger mean and a tendency to converge to the CDF of R at the upper tail (Figure 3A and 3B) The log-space cross-correlations have modest effects on these differences. The R' also has a lower threshold, which is

caused by the lower threshold of snowmelt floods each year. There is not much difference between the CDFs of S and S' (Figure 3C and 3D), expect when ρ= 0.9, which is an extreme situation and unlikely to occur in reality.

Figure 3A and 3B (with log-space cross correlation 0 and 0.9) also compares the CDFs of R' by fitting them to a two-parameter lognormal distribution with the CDFs of R' calculated from the integral of Equation 3.17 (derived from the two-parameter lognormal R and S). It shows that the LN2 distribution is not a good fit for R', whose CDF has a smaller low threshold and diverges from the CDF of the derived LN2 distribution from R at the upper tail. Figure 3C and 3D (with log-space cross correlation 0 and 0.9) compares the CDF of S' by fitting a LN2 distribution with the CDF of S' calculated from the integral of Equation 3.18. Two CDF curves match almost perfectly, which indicates LN2 is a proper distribution for describing the sub-population of snowmelt floods that are also annual maxima.

3.4.3 3-parameter Lognormal Distribution

The results in Section 3.4.2 indicate that 2-parameter lognormal distribution may not be a good fit for R'. A 3-parameter in lognormal distribution is considered. The methods of fitting the LN3 distribution and estimating the third parameter (a lower bound τ) were discussed in Sangal and Biswas [1970], Burges at el. [1975], Charbeneau [1978], and Stedinger [1980]. Here we employ the estimator of τ from Stedinger [1980]:

$$\hat{\tau} = \begin{cases} \dfrac{\hat{q}_p \hat{q}_{1-p} - \hat{q}_{0.5}{}^2}{\hat{q}_p + \hat{q}_{1-p} - 2\hat{q}_{0.5}} & \text{if } \hat{q}_p + \hat{q}_{1-p} - 2\hat{q}_{0.5} > 0 \text{ and } \hat{q}_p \hat{q}_{1-p} - \hat{q}_{0.5}{}^2 > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (3.20)$$

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} q_i \qquad (3.21)$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \left(q_i - \hat{\mu}\right)^2 \qquad (3.22)$$

$$\hat{\mu}_L = \ln\left[\frac{\hat{\mu}-\hat{\tau}}{\sqrt{1+\hat{\sigma}^2/\left(\hat{\mu}-\hat{\tau}\right)^2}}\right] \quad (3.23)$$

$$\hat{\sigma}_L^2 = \ln\left[1+\hat{\sigma}^2/\left(\hat{\mu}-\hat{\tau}\right)^2\right] \quad (3.24)$$

Where q_p and q_{1-p} are the largest and smallest observations in the sample. When the sample size is greater than 100, p=0.05 is applied.

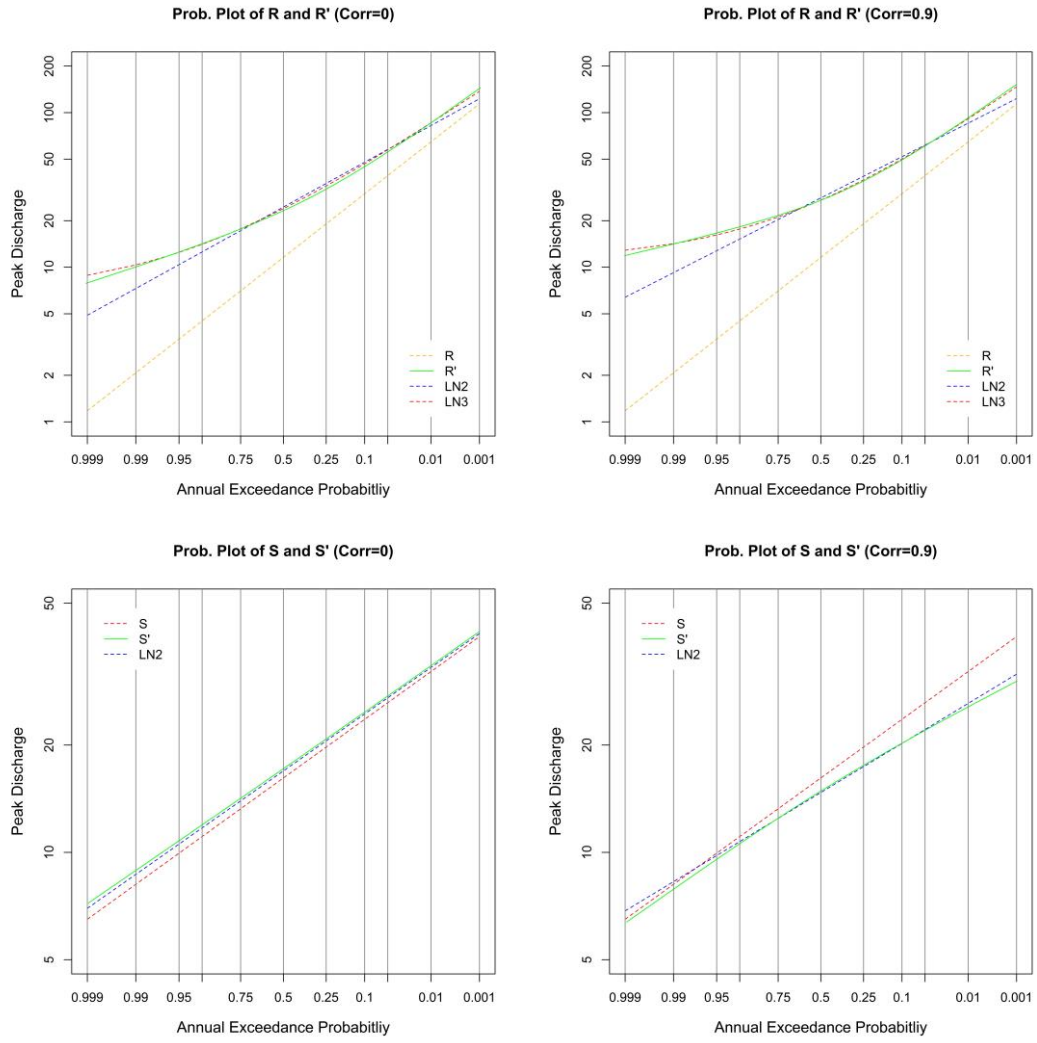


Figure 3 Probability plots of R and R' (3A-B), S and S'(3C-D), LN2 distributions with moments of R' and S', and LN3 distribution with moments of R'

The CDFs of R' when fitting a three-parameter lognormal distribution are also shown in Figure 3A and 3B. The parameters of LN3 are calculated from Equation 3.20-3.24. Compared with CDFs of R' when fitting a LN2 distribution, three-parameter lognormal is a more reasonable distribution for describing the sub-population of rainfall floods that are also annual maxima, though none of these two distributions describe the 1000-year flood risks perfectly.

Table 3 lists the Kolmogorov-Smirnov (K-S) statistics (the maximum difference between two CDFs) of the distributions shown in Figure 3. In spite of the different cross-correlations between R and S, R' always fits LN3 distribution better. The K-S statistics in the third row (Table 3) are even smaller, which indicates a good match between S' and the LN2 distribution with its moments.

Table 3 K-S Statistics between R'(S') and LN distributions with their moments

ρ	0	0.2	0.5	0.9
R' vs. LN2	0.0571	0.0495	0.0544	0.0876
R' vs. LN3	0.0317	0.0248	0.0177	0.0261
S' vs. LN2	0.0207	0.0142	0.0056	0.0152

By fitting R' to LN3 and S' to LN2, Kirby method can provide a reasonable description for the mixed population without knowing the complete rainfall and snowmelt flood series. In comparison with mixture model or joint model, it estimates 6 parameters ($\mu_{R'}$, $\sigma_{R'}$, $\mu_{S'}$, $\sigma_{S'}$, ρ, and P_R) by using fewer data (as listed in Table 3). Thus, the uncertainty of estimation (e.g. for 100-year flood) might be larger than other estimators. The probability density plots of R and R', and S and S' are displayed in Appendix A, from which one can reach the same conclusion as from Figure 3A-D.

3.5 Expected Moment Algorithm

Similar to the Kirby method, when the complete series of rainfall and snowmelt floods are not available (only the annual maximums), we can use the

expected moments algorithm (EMA) [Cohn et al. 1997, 2013; England et al. 2003a, 2003b] to estimate the distributions of both R and S, respectively. Assuming the causation of the peak flow in each year (t) is known (e.g. rainfall/snowmelt), the annual maxima in year t (R_t) can serve as the upper threshold for the annual maxima caused by the snowmelt (S_t) in the same year, and vice versa. In this case, R is a censored population, in which R' are observed, and the unobserved values have upper thresholds S' from the same year. Similarly, S' are observed values in population S, and the unobserved values have upper thresholds R' from the same year.

The expected moments algorithm (EMA) is a moments-based parameter estimation procedure that was adapted from the iterated least squares (ILS) method for fitting regression models to censored data [Schmee and Hahn, 1979; Cohn et al. 1997], which initially estimates the linear regression parameters by using censored data, then uses the estimates from the previous step to obtain a revised regression fit conditionally on the censoring thresholds until the iteration converges. EMA is also related to the expectation-maximization (EM) algorithm [Dempster, 1977], which is a widely used iterative approach to finding the maximum likelihood estimators when samples contain unobserved data.

For the data whose log-space values follow a normal distribution, the method of moments estimator and the maximum likelihood estimator are the same. Thus, EMA should generate the same quantile estimator as the one computed by EM algorithm. EMA consist of three major steps. Using the rainfall flood series R as the example:

Step 1. Initialization: get initial log-space sample moment estimates

$$\hat{\mu}_0 = \frac{\sum_{i=1}^{n_r} \ln(r_i)}{n_r} \qquad (3.25)$$

$$\hat{\sigma}_0^2 = \frac{\sum_{i=1}^{n_r}\left[\ln(r_i) - \hat{\mu}_0\right]^2}{n_r} \qquad (3.26)$$

Here r is the maxima of rainfall floods (R') that is also the annual maxima, and n_r is the number of rainfall events in the annual maximum series.

Step 2. Iteration: generate new sample moments

$$\hat{\mu}_{t+1} = \frac{\sum_{i=1}^{n_r}\ln(r_i) + \sum_{j=n_r+1}^{n} E\left[\ln(R_j) \mid \ln(R_j) \le \ln(s_j); \hat{u}_t, \hat{\sigma}_t\right]}{n} \qquad (3.27)$$

$$\hat{\sigma}_{t+1}^2 = \frac{\sum_{i=1}^{n_r}\left[\ln(r_i) - \hat{\mu}_{t+1}\right]^2 + \sum_{j=n_r+1}^{n} E\left\{\left[\ln(R_j) - \hat{\mu}_{t+1}\right]^2 \mid \ln(R_j) \le \ln(s_j); \hat{u}_{t+1}, \hat{\sigma}_t\right\}}{n} \qquad (3.28)$$

Here *t* represents the number of iterations, n is the population size of the annual maximum series (Q). R_j is not observed, but its upper threshold is known to be s_j. The conditional expected values at each iteration are determined by the moments (mean and variance) from the previous iteration.

Step 3. Convergence test: iterate Step 2 until parameter estimates converge,

$$\text{which means } \left|\mu_{t+1} - \mu_t\right| \le \varepsilon \quad \text{and} \quad \left|\sigma_{t+1}^2 - \sigma_t^2\right| \le \varepsilon \qquad (3.29)$$

$\varepsilon = 10^{-4}$ is adopted in Monte Carlo studies in Section 4.

The expected moments algorithm is based on the assumption of a fixed upper threshold for the rainfall or snowmelt event that is not the annual maxima, thus it is unreasonable to assume there is a strong cross-correlation between R and S, which means when the rainfall flood is large, the snowmelt flood is large as well. Once the moments for R and S are estimated by using Equation 3.25-3.29, one can employ the mixture model to get $F_Q(q)$.

3.6 3-Parameter Lognormal Estimator

When the information of flood causation is not available, an alternative is to fit the data to a single distribution. That is a traditional model people use in hydrologic studies. A single three-parameter lognormal distribution is considered to fit the annual maximum series, which often appears to have a lower threshold (see Figure 1). The moments for one single LN3 distribution are calculated following the Equation 24-28 in Section 3.4.3 by using the all the data from mixed population Q. Its probability plots is shown in Figure 4, which indicates that LN3 is a reasonable distribution for the North Saint Vrain Creek flood series.

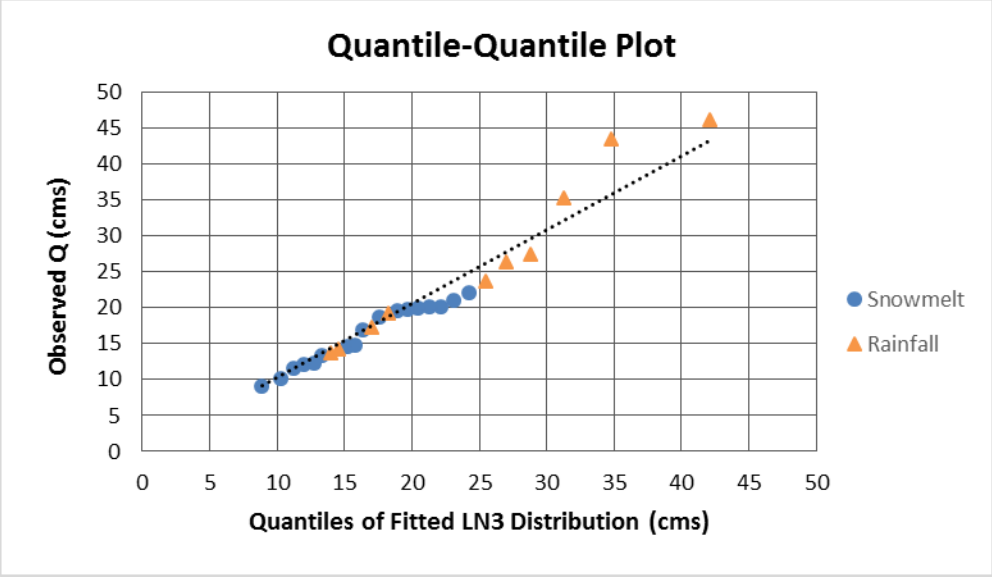


Figure 4 Quantile-quantile Plot of Three-parameter Lognormal Distribution based upon annual peak flows from 1926 to 1953 at North Saint Vrain Creek, CO

3.7 Summary of Estimators

Table 4 summarizes the six estimators developed in Sections 3.1-3.6 for annual flood risk estimation. The mixture estimator, which results from assuming R and S are independent, is a special case of the joint model estimator. The rainfall-only lognormal estimator is a special case of both the mixture and joint estimators when snowmelt

events are ignored. Kirby, EMA, and single LN3 use only the annual maximum series; Kirby and EMA also need to know whether annual maximum floods were snow or rainfall events.

Compared to the joint estimator, the Kirby estimator includes an extra parameter because it is necessary to use a LN3 distribution for rainfall events that were also annual maxima. Kirby also includes the weights or probabilities, P_R and P_S ($P_R + P_S = 1$) of rainfall and snowmelt events; in terms of parameters, however, the joint model includes a cross-correlation which is needed to determine P_R and P_S.

With the mixture and EMA estimators, P_R and P_S are computed from the distributions of R and S, assuming the two are independent and lognormal.

Table 4 Summary of 6 Estimators
for Computing Distribution of Annual Maximum Flood Examined in this Chapter

Estimators	Assume Independence	Observations	No. of Parameters
Use maximum series (R & S)			
Mixture	Yes	$N_R + N_S$	2+2=4
Joint	No	$N_R + N_S$	4+1=5
Rainfall-only	No	N_R	2
Use annual maximum series + flood causation			
Kirby	No	$N_{R'} + N_{S'} = N$	5+1=6
EMA*	Yes	$N_{R'} + N_{S'} = N$	2+2=4
Use annual maximum series only			
Single LN3**	No	$N_{R'} + N_{S'} = N$	3

* EMA fits the mixture model using only the AMS and knowledge of the source of each flood.
** LN3 fits a single flood distribution to the AMS.

4. Monte Carlo Study

A Monte Carlo study evaluated the relative performance of the six estimators introduced in Section 3. The log-space mean squared error (LMSE) is a reasonable performance criterion reflecting the impact of parameters estimation errors on flood risk management activities. Compared to the real-space MSE, it better resembles

actual losses associated with design flood estimation errors; in particular, it has the advantage that under-design errors receive greater weight than overdesign errors of equal magnitude [Fill, 1994; Fill and Stedinger, 1998]. LMSE was estimated using:

$$LMSE[\hat{q}_p] = \frac{1}{N}\sum_{i=1}^{N}\left[\ln(\hat{q}_p) - \ln(q_p)\right]^2 \qquad (4.1)$$

Each Monte Carlo simulation started with N = 10,000 replicates. However, when a sample had too few rainfall or snowmelt events, the sample was rejected. For all cases reported, at least 5,000 replicates were retained. Hydrologists are concerned with the extreme flood events, such as the 10-year and 100-year flood flows; thus, results for the 0.90 and 0.99 percentiles are presented.

4.1 Experiment

In the experiment, random joint lognormal samples were generated with the following characteristics:

Correlation (ρ) between log-space R and log-space S: 0, 0.2, 0.5, 0.9

Probability R > S, denoted P_R: 30%, 50%, 80% (thus P_S: 70%, 50%, 20%)

Sample sizes (n): 25, 50, and 100

CV-Ratio: in most experiments CV-Ratio = 2.48;

both 1.24 and 4.23 were also considered

For μ_{RL} in Table 6, different values of P_R are achieved by adjusting the median ratio $M_{R/S}$ in the equation to generate different μ_{SL} that achieve different P_R:

$$M_{R/S} = R_{0.5} / S_{0.5} = exp\,[\mu_{RL} - \mu_{SL}] \qquad (4.2)$$

$$\mu_{SL} = \mu_{RL} - \ln (M_{R/S}) \qquad (4.3)$$

Lu (Master's Thesis, 2013) shows that with the joint model both the log-space cross correlation (ρ) and median ratio have an influence on P_R. For CV-Ratio =2.48, when $M_{R/S} = 1$, $P_R = P_S = 0.5$ for all correlations; when $M_{R/S} = 1.5$, $P_R \approx 0.80$, which

varies between 0.73 and 0.86 with different values of ρ; when $M_{R/S} = 0.75$, $P_R \approx 0.30$, again it varies within the range from 0.23 to 0.33.

To have different cross correlations (ρ) between R and S, the log-space mean and standard deviation of S are scaled using the conditional normal distribution (Equation 3.7) function when R is given.

In addition, data sets with different coefficient of variation ratios (CV-ratios) for R and S are considered. CV is a measure of the relative standard deviation; thus, the CV-ratio (CV_R/CV_S) reflects the difference in variability between the populations.

The moments of annual maximum R and S series from three locations with different characteristics were employed in this effort to describe real catchments. Tables 5 and 6 list the location of each watershed, its drainage area, and its statistical characteristics of R and S. The three watersheds described in the table will be used as examples of realistic at-site P_R values, CVs, and the CV-ratios. For these watersheds, P_R varies from 27% to 53%. In practice P_R depends upon the distribution of the two sources of flood risk; for rain and snow it generally depends upon region, latitude, and watershed elevation, all of which affects whether precipitation is snow or rain, and their intensity.

Table 5 Watershed and Streamflow Information

Location Name	Drainage Area (km^2)	Elevation (m)	Average Peak Discharge (cms)
San Miguel River at Naturita, CO	2769	1644*	90
Coquihalla River, British Columbia	740	1700**	244
North Fork Big Thompson River, CO	214	1881*	17

* Gage datum; ** Median elevation

When the sample size is small, the Kirby method will not perform stably because the annual maximum series is split into R' and S' series, and one or the other (or both) sample will be very small. Because EMA uses the observed annual maximum to bound the unobserved maximum (i.e. the minimum of R and S each year), EMA is more stable. With the Kirby estimator, when the generated R' sample is

smaller than 10, an LN2 distribution is employed instead of LN3 for the R' distribution. *For all the estimators,* when either the R' or S' sample is smaller than 5, that sample is rejected. Thus, the randomness in the Monte Carlo simulation experiments is preserved conditionally for R' and S' sample sizes large enough for fitting a LN2 distribution to those samples. Furthermore, in unreported experiments, results for Kirby estimator are not reported when the sample size is 25 because either the R' or S' sample will be small.

Table 6 Original Moments and CV-Ratios for Annual Maximum Series

Location Name	μ_{RL}	σ_{RL}	μ_{SL}	σ_{SL}	Original P_R	CV_R/CV_S
San Miguel River at Naturita, CO	3.56	0.70	4.23	0.59	27%	1.24
Coquihalla River, British Columbia	5.11	0.60	5.05	0.26	53%*	2.48
North Fork Big Thompson River, CO	2.03	1.47	1.50	0.60	40%	4.23

* Estimated from $M_{R/S}$

4.2 Result and Discussion

4.2.1 Simulation results with CV ratio = 2.48

Figure 5 and 6 compare the performance among six estimators of the 10-year and 100-year flood with sample size of 50, which is a realistic length record for a hydrologic record.

For 10-year flood estimation (Figure 5), the LMSEs of the six estimators generally have only small differences. The exception is the rainfall-only LN2 estimator when snowmelt events dominate flood risk ($P_R = 0.3$). The mixture and joint estimators usually work better than other estimators. Those two estimators use the complete rainfall and snowmelt flood series (R and S), rather than just the mixed maximum series (Q), which can be more highly skewed and irregular. When the correlation ρ is 0.9 and $P_R = 0.3$, the mixture and EMA estimator are not as effective as for the low correlation cases, since they assume the independence between R and S, and rainfall is not as dominant. Generally, $\rho = 0.9$ is not a realistic value.

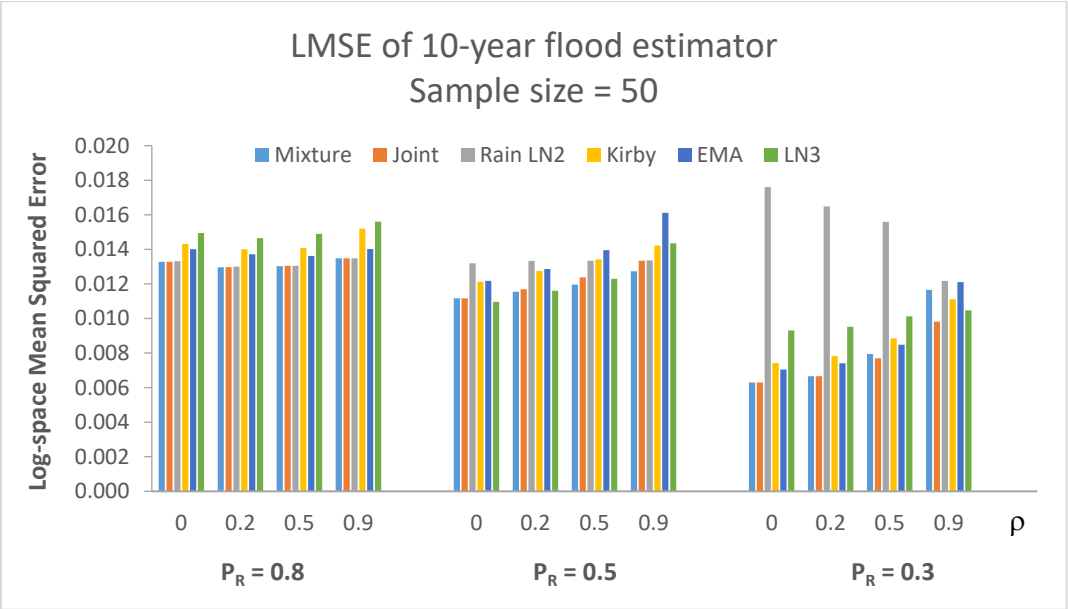


Figure 5 Demonstration of the effect of P_R on performance of the estimators, and modest effect of ρ. Figures shows log-space MSE of 10-Year Flood Estimator, CV-Ratio = 2.48

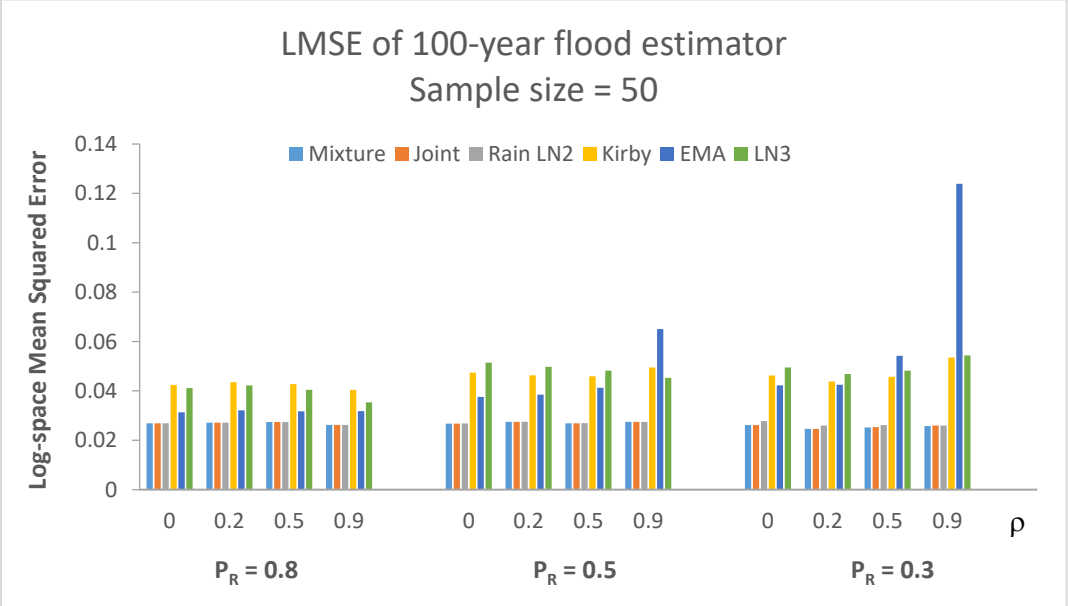


Figure 6 Demonstration of the effect of P_R on performance of the estimators, and modest effect of ρ. Figures shows log-space MSE of 100-Year Flood Estimator, CV-Ratio = 2.48

For 100-year flood estimation in Figure 6, the performance of mixture, joint, and rainfall-only estimators are essentially indistinguishable and much better than the

other estimators that use only the AMS. The rainfall-only lognormal estimator works well because the 99-percentile is determined by the rainfall events (see Figure 2).

If only the annual maximum series is available, EMA looks to be the best because it performs better than the Kirby and single LN3 estimators for $\rho \leq 0.5$ (and $P_R \geq 0.5$); $\rho = 0.9$ is unrealistic in many situations, such as floods in different seasons. The performance of Kirby and the single LN3 estimators is about the same, though Kirby works better than single LN3 when snowmelt events dominant ($P_R = 0.3$). $P_R = 0.3$ is a hard case for single-LN3 because the upper tail risk is really driven by the rainfall events, but they are a very small fraction of the annual maximum series.

4.2.2 Simulation results with multiple CV ratios

For different CV ratios, Figure 7 and 8 compare scaled LMSEs of 0.90 percentile and 0.99 percentile estimators. The results of n = 50 and $\rho = 0.2$ are reported. Appendices D report results of n = 25 and 100, and all the specific values of LMSE generated in Monte Carlo simulations.

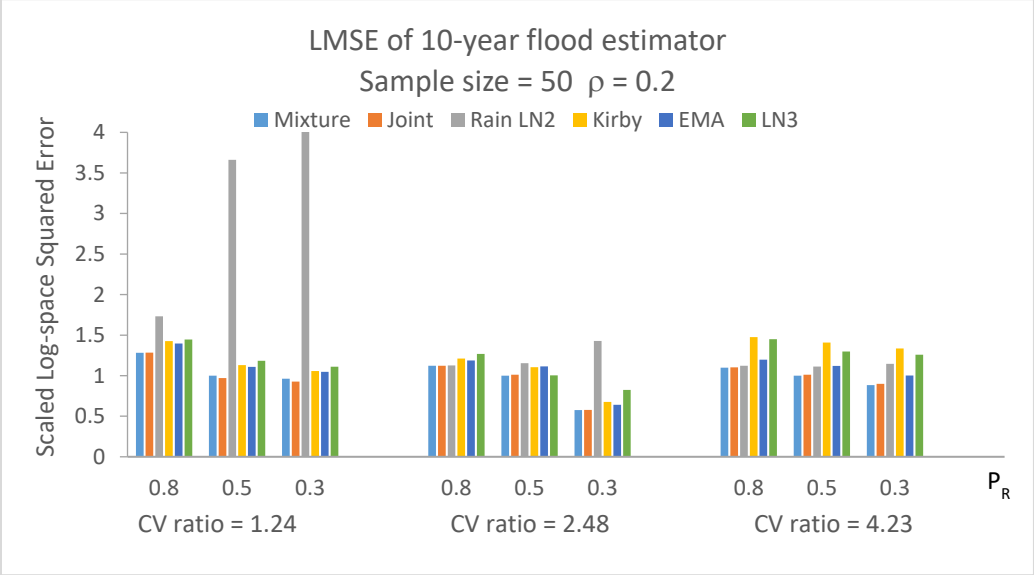


Figure 7 Demonstration of effects of CV-ratio on performance of the estimators.

Figures shows scaled LMSE of 10-year flood estimator with different CV-ratios for $\rho = 0.2$

*Scaled LMSE is 9.65 for $P_R = 0.3$, CV-ratio = 1.24. LMSE is truncated by 4.

In order to make the comparisons clear and because of the large effect of the CV-ratio on the scale of the results, LMSEs in Figure 7 and 8 have been scaled by the LMSE of mixture estimator (P_R = 0.5) for each CV ratio.

When CV ratio = 1.24, the rainfall-only estimator performs poor, especially with a higher percent of snowmelt events. For CV ratio = 1.24 and P_R = 0.3, rainfall does not dominate.

For 100-year flood estimation, estimators' performances are similar when CV ratios = 2.48 and 4.23. The rainfall-only estimator performs almost as well as mixture and joint estimators. The single LN3 and Kirby estimators perform about the same – both often poor, with Kirby a little better than the single LN3.

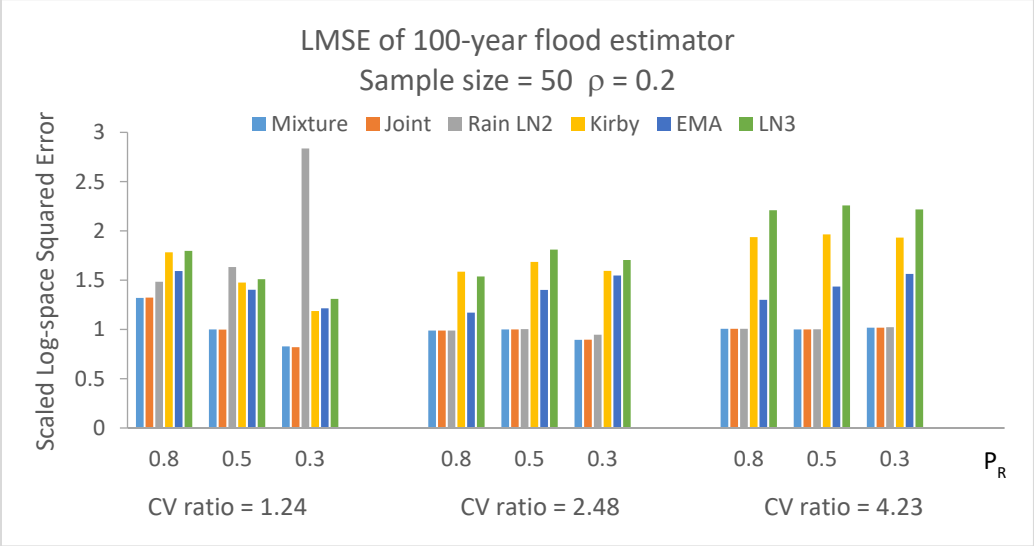


Figure 8 Demonstration of effects of CV-Ratio on performance of the estimators.
Figures shows scaled LMSE of 100-year flood estimator with different CV-ratios for ρ = 0.2

5. Conclusion and Recommendation

This paper considers different models that can be used to describe annual flood risk with mixed peak flow series (generated by rainfall and snowmelt events, or other

distinct processes) and their parameter estimators with different data sets. The analysis evaluated the three risk models, listed below, and different parameter estimators that go with each model (see Table 2).

Ideally, fitting the joint model would be the best for describing the mixed population when both the rainfall and snowmelt flood series records are complete, whether or not they are independent. The Monte Carlo study results indicate that, for all the estimators, correlation between the rainfall and snowmelt events has a modest effect on the mean squared errors of 10- and 100-year flood estimators for $\rho \leq 0.5$. Thus, the simple mixture estimator provides an accurate approximation of the annual maximum distribution for quantile estimation (10-year and 100-year flood) when the R-S correlation ρ is less than 0.5; this is the case for most rivers in the western U.S. [Elliott, et al., 1982]. If the rainfall events dominate determination of flood risk, the LN2 estimator of rainfall-only model should be adequate for describing the flood risk.

When available flood records include only the annual maximum flood series it is more difficult to develop models of the two series. In these cases, the EMA performances better than Kirby estimator for 100-year flood estimation when the R-S correlation $\rho \leq 0.5$; for 10-year flood estimation, the Kirby and EMA estimators' performance is about the same, though EMA is often a little better. Another advantage of EMA is that it generates complete models of the annual maxima for both rain and snowmelt, which can be augmented with a regional skew, historical information, or information on land use or climate change. Compared with alternative estimators, Kirby has the fewest observations per estimated parameter. If the number of events for either the snowmelt floods or the rainfall floods is small (less than 10) in the annual maximum series, the Kirby method is not stable.

To summarize, in situation where one source of floods provides the largest annual maxima, and a second source provides most of the small floods, it is reasonable

to split the annual maximum series into distinct physical processes and model them separately. When both rainfall and snowmelt floods records are available, the simple mixture estimator is recommended for flood frequency analysis, because correlation between rainfall and snowmelt floods has modest effects on the distribution of the annual maximum series. Among the three estimators that only use the annual maximum series, the EMA estimator is the best choice for 100-year flood estimation when the annual maximum series results from two distinct flood sources with different simple distributions. The Kirby and single LN3 estimators are less efficient. Still the single LN3 estimator is simpler than EMA, and is certainly appropriate in applications where mixtures are not a critical issue because one source of flooding dominates the flood risk, or the two sources have similar distributions.

References

Aitchison, J., and J. A. C. Brown (1957), *The Lognormal Distribution with special reference to its uses in economics,* Cambridge at the University Press.

Alila, Y., and A. Mtiraoui (2002), Implications of heterogeneous flood-frequency distributions on traditional stream-discharge prediction techniques, *Hydrol. Process, 16,* 1065-1084.

American Society of Civil Engineering (1996), *Hydrology Handbook,* 2nd Edition, ASCE, New York, 490-91.

Barr, R. Donald and E. Todd Sherrill (1999), Mean and Variance of Truncated Normal Distributions, The American Statistician, 53.4, 357-361.

Burges, S. J., D. P. Lettenmaier, and C. L. Bates (1975), Properties of the three-parameter lognormal probability distribution, *Water Resour. Res., 11*(2), 229-235.

Canfield, R. V., D. R. Olsen, R. H. Hawkins, and T. L. Chen (1980), *Use of Extreme Value Theory in Estimating Flood Peaks from Mixed Populations*, Reports, Paper 577. Utah Water Research Laboratory, Utah State University.

Charbeneau R., J. (1978), Comparison of the two- and three-parameter lognormal distributions used in streamflow synthesis, *Water Resour. Res., 14*(1), 149-150.

Cohen, A. C., Jr. (1951), Estimating parameters of logarithmic-normal distributions by maximum likelihood, *J. Amer. Statist. Ass., 46,* 206-212.

Cohn, T. A., W. L. Lane, and W.G. Baier (1997), An algorithm for computing moments-based flood quantile estimates when historical flood information is available, *Water Resources Research*, *33*(9), 2089-2096.

Cohn, T. A., J. F. England, C. E. Berenbrock, R. R. Mason, J. R. Stedinger, and J. R. Lamontagne (2013), A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series, *Water Resources Research, 49,* 5047–5058, doi:10.1002/wrcr.20392

Cudworth, A.G., Jr. (1989), *Flood Hydrology Manual,* Water Resources Technical Publication, U.S. Bureau of Reclamation, U.S. Dept. of the Interior, Denver, CO, pp. 205-207 and 219-220.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc. Ser. B, 39,* 1–22.

Elliott, J. G., R. D. Jarrett, and J. L. Ebling (1982), *Annual snowmelt and rainfall peak-flow data on selected foothills region streams, South Platte River, Arkansas River, and Colorado River basins, Colorado,* U.S. Geological Survey Open File Report 82-426

England, J. F., Jr., R. D. Jarrett, and J. D. Salas (2003a), Data-based comparisons of moments estimators using historical and paleoflood data, *Journal of Hydrology 278,* 172–196

England, J. F., Jr., J. D. Salas, and R. D. Jarrett (2003b), Comparisons of two moments-based estimators that utilize historical and paleoflood data for the log Pearson type III distribution, *Water Resour. Res., 39*(9), 1243, doi:10.1029/2002WR001791.

Escalante-Sandoval, C. (2007), A Mixed distribution with EV1 and GEV components for analyzing heterogeneous samples. *Ingeniería, investigación y tecnología, 8*(3), México

Escalante-Sandoval, C., and J. Raynal-Villasenor (2008), Trivariate generalized extreme value distribution in flood frequency analysis, *Hydrological Sciences, 53*(3), 550-567

Fill, H. D. (1994), *Improving Flood Quantile Estimates Using Regional Information.* Ph.D. Dissertation, Cornell University, Ithaca, New York.

Fill, H. D., and J. R. Stedinger (1998), Using regional regression within index flood procedures and an empirical Bayesian estimator, *Journal of Hydrology,* 210, 128–145

Hazen, A. (1914), Discussion on flood flows, *Trans. Amer. Soc. Civil Eng., 77,* 626-632.

Hirschboeck, K. K. (1987), Hydroclimatically – defined mixed distributions in partial duration flood series, in Singh, V.P., ed., *Hydrologic Frequency Modeling,* D. Reidel Publishing Company, 199_212.

Horton, R. E. (1914), Discussion on flood flows, *Trans. Amer. Soc. Civil Eng., 77,* 663-670.

Interagency Advisory Committee on Water Data [IACWD] (1982), *Guidelines for determining flood flow frequency, Bulletin #17B of the Hydrology Subcommittee,* U.S. Geological Survey, Reston, Virginia.

Jarrett, R. D., and J. E. Costa (1982), Multidisciplinary approach to the flood hydrology of foothill streams in Colorado, Johnson, A. J., and R.A. Clark eds., *International Symposium on Hydrology*: Bethesda, Md., American Water Resources Association, 565-569.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994), *Continuous Univariate Distributions,* Volume 1, Second Edition, Chapter 14 Lognormal Distributions, John Wiley & Sons, Inc., New York, NY.

Kidson, R. and Richards, K. S. (2005), Flood frequency analysis: assumptions and alternatives, *Progress in Physical Geography, 29*(3), 392-410.

Kite, G. W. (1977, 1988), *Frequency and Risk Analysis in Hydrology,* Water Resources Publications, Littleton, Colorado, USA.

Lamontagne, J. R., J. R. Stedinger, C. Berenbrock, A. G. Veilleux, J. C. Ferris, and D. L. Knifong (2012), *Development of regional skews for selected flood durations for the Central Valley Region, California, based on data through water year 2008*, U.S. Geological Survey Scientific Investigations Report 2012-5130, 60 p.

Lecce, S. A. (2000), Spatial variations in the timing of annual floods in the southeastern United States, *Journal of Hydrology 235*, 151–169

Lu, Jiajia (2013), *Statistical analysis of mixed population flood series*, Master's Thesis, School of Civil and Environmental Engineering, Cornell University.

Murphy, P. J. (2001a), Evaluation of mixed-population flood-frequency analysis, *J. Hydrol. Eng., 6*(1), 62-70

Murphy, P. J. (2001b), Estimating equation for mixed populations of floods in Massachusetts, *J. Hydrol. Eng., 6*(1), 72-74

Natural Environment Research Council (1975), *Flood Studies Report Volume I: Hydrological Studies*, Whitefriars Press Ltd., London, UK.

Parrett, C., Veilleux, A., Stedinger, J. R., Barth, N. A., Knifong, D. L., and Ferris, J. C. (2011), *Regional skew for California, and flood frequency for selected sites in the Sacramento–San Joaquin River Basin, based on data through water year 2006*, U.S. Geological Survey Scientific Investigations Report 2010–5260, 94 p.

Sangal, B. P., and A. K. Biswas (1970), The three-parameter lognormal distribution and its application in hydrology, *Water Resour. Res., 6*(2), 505-515.

Schmee, J., and G. J. Hahn (1979), A simple method for regression analysis with censored data, *Technometrics, 21*(4), 417-432.

Singh, V. P., S. X. Wang, and L. Zhang (2005), Frequency analysis of nonidentically distributed hydrologic flood data, *Journal of Hydrology 307*, 175–195.

Smith, A. James, G. Villarini, And M. L. Baeck (2011), Mixture distributions and the hydroclimatology of extreme rainfall and flooding in the eastern United States, *Journal of Hydrometeorology*, Volume 12, 294-309.

Stedinger, J. R. (1980), Fitting log normal distributions to hydrologic data, *Water Resources Research, 16*(3), 481-490.

Stedinger, J. R., and Cohn, T. A. (1986), Flood frequency analysis with historical and paleoflood information, *Water Resources Research, 22*(5), 785-793.

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou (1993), *Frequency analysis of extreme events*, Chapter 18, Handbook of hydrology, McGraw-Hill, New York.

Stedinger, J. R. (2000), *Inland Flood Hazards,* Chapter 12, Cambridge University Press, Cambridge, UK.

U.S. Army Engineer District, Corps of Engineers (1958), Research Note No. 1, *Frequency of New England Floods.*

U.S. Army Corps of Engineers (1982), *Mixed-population frequency analysis,* Hydrologic Engineering Center, Institute for Water Resources, Davis, CA.

Watt, W.E. (ed., 1989), *Hydrology of Floods in Canada: A Guide to Planning and Design,* Associate Committee on Hydrology, National Research Council, Ottawa, pp. 25-28, 69-71.

Waylen, P., and M-K. Woo (1982), Prediction of Annual Floods Generated by Mixed Processes, *Water Resources Research, 18,* 1283-1286.

Wilson, E. B., and J. Worcester (1945), The Normal Logarithmic Transform, *Rev. Econ. Statist., 27*(1), 17-22.

Appendices for Chapter 2

For the sake of brevity, some of the details related to the analyses in Chapter 2 have been placed in this appendix. This Appendix for chapter 2 has four sections, denoted 2.A through 2.D. They provide an explanation of numerical methods used to compute the CDF of the maximum for the joint model and to compute the R' and S' distributions, additional figures displaying the PDF for R' and S' which are modeled directly with the Kirby method, details of the Expected Moment Algorithm (EMA) computation, and the sets of parameters used in the Monte Carlo simulation study reported in Chapter 2, and additional results for n = 25, 100, and $\rho = 0.7$.

Appendix 2.A – *Numerical Integration for Joint Model*

Section 3.2 in Chapter 2 discussed the need for numerical integration of the Joint model to compute the probability that a given flood level is exceeded. This appendix introduces the specific numerical method – composite Simpson's rule.

The numerical integration addresses evaluation of the following equation, which yields the cumulative probability function for the annual maximum Q.

$$F_Q(q) = \int_{-\infty}^{q}\int_{-\infty}^{q} f_{R|S}(r \mid s)\, f_S(s)\, dr\, ds = \int_{-\infty}^{q}\left[\int_{-\infty}^{q} f_{R|S}(r \mid s)\, dr\right] f_S(s)\, ds = \int_{-\infty}^{q} F_{R|S}(q \mid s)\, f_S(s)\, ds \tag{2.A-1}$$

The initial bivariate integral considers the probability of all *r*-*s* pairs both less than *q*. The second middle formula is an intermediate step in obtaining the third univariate integral. Assuming the conditional distribution of R given S is available (as it is for the join lognormal distribution), the original bivariate integral formula has been converted into a single univariate integration. Furthermore, because $F_{R|S}(r|s)$ for most of its critical range should not be very sensitive to s, the last univariate integral should be easier to evaluate than one where the integral changed wildly with s.

The composite Simpson's rule is adopted dividing the finite length integration interval (0, q) into n=20,000 segments each of width $h=q/n$, corresponding to $m=n/2=10000$ pairs of segments. Simpsons' Rule works with pairs of segments. This yields:

$$F_Q(q) = \int_{-\infty}^{q} F_{R|S}(q \mid s) f_S(s) ds$$
$$\cong \frac{h}{3}\left[F_{R|S}(q \mid s_0) f_S(s_0) + 4\sum_{i=1}^{m} F_{R|S}(q \mid s_{2i-1}) f_S(s_{2i-1}) + 2\sum_{j=1}^{m-1} F_{R|S}(q \mid s_{2j}) f_S(s_{2j}) + F_{R|S}(q \mid s_{2m}) f_S(s_{2m}) \right]$$

(2.A-2)

Because that the integration was done in q space, the lower bound was zero. Equation 2.A-2 was used to compute the results reported in chapter 2. Accuracy was checked by using different values of h.

Appendix 2.B – *Probability Density Plots for R' and S' in Kirby Method*

Section 3.4 in Chapter 2 discussed the distributions of R' and S'. The R' and S' probability density functions (PDFs) were derived with the integral in Equation 3.17 and 3.19 using the parameters of the joint models with four cross-correlations. The PDFs are plotted in Figure 2.B-1 and 2.B-2 below for the 4 cases. The LN2 and LN3 PDFs in figures 2.B-1 and 2.B-2 are for 2-parameter and 3-parameter lognormal distributions that have the mean and variance for R' and S'. The moments of R' and S' were computed numerically using the R' and S' PDFs for each case.

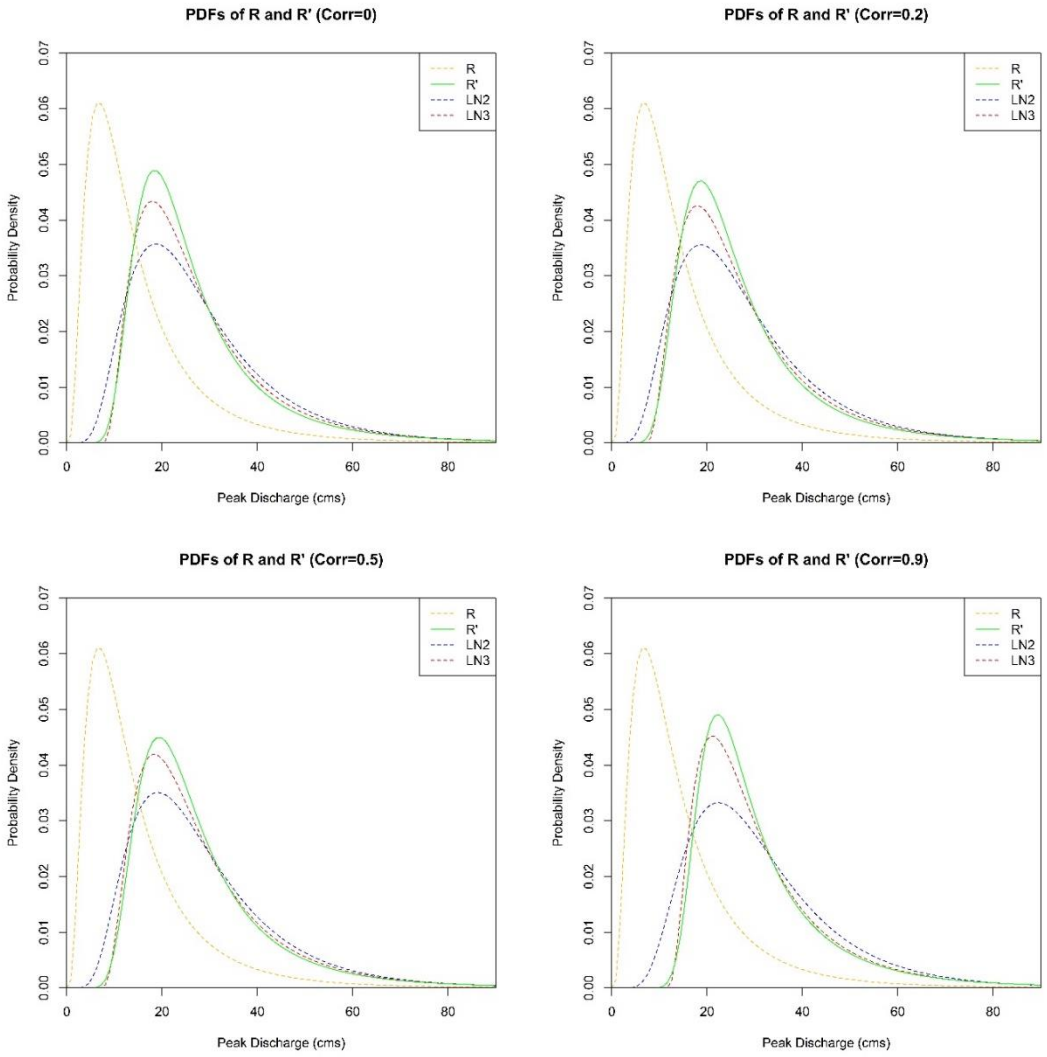


Figure 2.B-1 PDFs of R, R', and LN2 & LN3 Distributions with Moments of R'

A comparison of the PDFs for each case supports the conclusion in Section 3.4: LN2 is a reasonable distribution for describing the sub-population of snowmelt floods that are also annual maxima; compared with the 2-parameter lognormal distribution, the 3-parameter lognormal is more reasonable for describing the sub-population of rainfall floods that are also annual maxima because it is almost as if they had a lower bound greater than zero. The difference between the R' and R distributions are very apparent. On the other hand, for $\rho \leq 0.5$, there is very little

difference between the S and S’ distribution; thus, if S has a 2-parameter lognormal distribution, S’ has very close to a 2-parameter lognormal distribution.

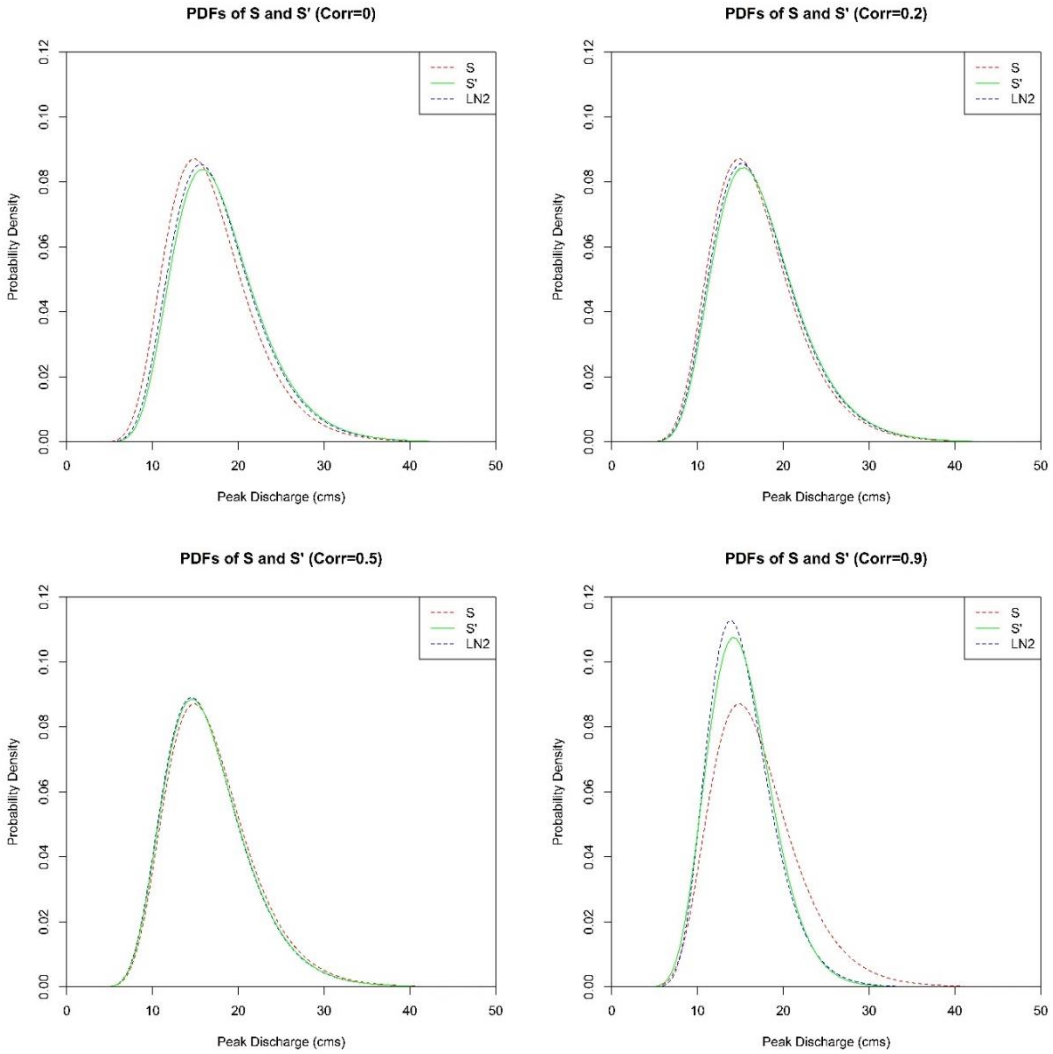


Figure 2.B-2 PDFs of S, S’, and LN2 Distributions with Moments of S’

Appendix 2.C – *Expected Moments Algorithm (EMA)*

In Section 3.5 of Chapter 2, the three iterated steps of EMA have been introduced. In Step 2 (Equation 3.27 and 3.28), the conditional expected values of the log-space rainfall floods (R) were used for updating the new moments of ln(R). This appendix shows how we calculated the conditional expected values of log-space R.

The population of rainfall floods (R) is assumed to have a lognormal distribution, thus ln(R) has a normal distribution. R_j is a rainfall flood known to be less than its upper threshold s_j. The conditional expected values for censored observations R_j know to be less than s_j are calculated as follow (Barr and Sherrill, 1999; Cohn, et al. 2013):

We use the key relationship for standard normal random variables that

$$E\left(Z \mid \mathrm{Z} \le \mathrm{z}\right) = -\frac{\phi(z)}{\Phi(z)} \qquad \text{(2.C-1)}$$

and

$$E\left(Z^2 \mid Z \le z\right) = 1 - z\frac{\phi(z)}{\Phi(Z)} \quad \text{(2.C-2)}$$

Thus, for censored observations, EMA employs the approximations that

$$E\left[\ln\left(R_j\right) \mid \ln\left(R_j\right) \le \ln\left(s_j\right); \hat{u}_t, \hat{\sigma}_t\right] = \hat{u}_t + \hat{\sigma}_t E\left[Z \mid Z \le z\right] = \hat{u}_t - \hat{\sigma}_t \frac{\phi(z)}{\Phi(z)}$$

(2.C-3)

and

$$E\left\{\left[\ln\left(R_j\right) - \hat{\mu}_{t+1}\right]^2 \mid \ln\left(R_j\right) \le \ln\left(s_j\right); \hat{\mu}_{t+1}, \hat{\sigma}_t\right\} = \hat{\sigma}_t^2 E\left[Z^2 \mid \mathrm{Z} \le z\right] = \hat{\sigma}_t^2\left[1 - z\frac{\phi(z)}{\Phi(z)}\right]$$

(2.C-4)

Here μ and σ are the moments for the rainfall flood series (R); z is a standard normal variate and Z is a standard normal random variable; ϕ and Φ are the PDF and CDF of the normal distribution, respectively; t represents the iteration number.

Appendix 2.D – *Monte Carlo Study*

1. Parameters Adopted

The original parameters used in the Monte Carlo study (Section 4, Chapter 2) are listed in Table 2.D-1. To obtain different values of the median ratio $M_{R/S}$, we simply adjust the log-space mean of the snowmelt series (μ_{SL}) using equation:

$$\mu_{SL} = \mu_{RL} - \ln (M_{R/S}) \qquad (2.D\text{-}1)$$

The original values of P_R are listed in Table 2.D-1. Both the CV ratio and the correlation affect the value of P_R (Lu, Master's Thesis, 2013). Most cases reported use CV-ratio=2.48. Parameter sets with different cross-correlations but the same median ratio $M_{R/S}$ and CV-ratios of 2.48 have similar P_R values (see Table 2.D-2). Thus, in the figures with results, different sets of median ratios are labels by their approximate P_R value. The adjusted μ_{SL} values are listed in Table 2.D-3.

Table 2.D-1 Original Parameters

Location Name	μ_{RL}	σ_{RL}	μ_{SL}	σ_{SL}	Original P_R
San Miguel River at Naturita, CO	3.56	0.70	4.23	0.59	27%
Coquihalla River, British Columbia	5.11	0.60	5.05	0.26	53%
North Fork Big Thompson River, CO	2.03	1.47	1.50	0.60	40%

Table 2.D-2 P_R with different median ratios and cross-correlations all CV-ratios = 2.48

Correlation	Median Ratio ($M_{R/S}$)		
ρ (log-space)	1.50	1.00	0.75
0	0.73	0.50	0.33
0.2	0.75	0.50	0.32
0.5	0.78	0.50	0.29
0.9	0.86	0.50	0.23

Table 2.D-3 Log-space Mean of Snowmelt Series (μ_{SL}) after Adjustment

Location Name	$P_R \approx 0.8$	$P_R \approx 0.5$	$P_R \approx 0.3$
San Miguel River at Naturita, CO	3.15	3.56	3.85
Coquihalla River, British Columbia	4.70	5.11	5.40
North Fork Big Thompson River, CO	1.62	2.03	2.32

2. Study Results

Figures 2.D-1 – 2.D-4 present results of Monte Carlo studies for sample size n = 25 and n = 100. Results for $P_R = 0.8$ with n = 25 are omitted because samples often have too few snowmelt events to fit a S distribution.

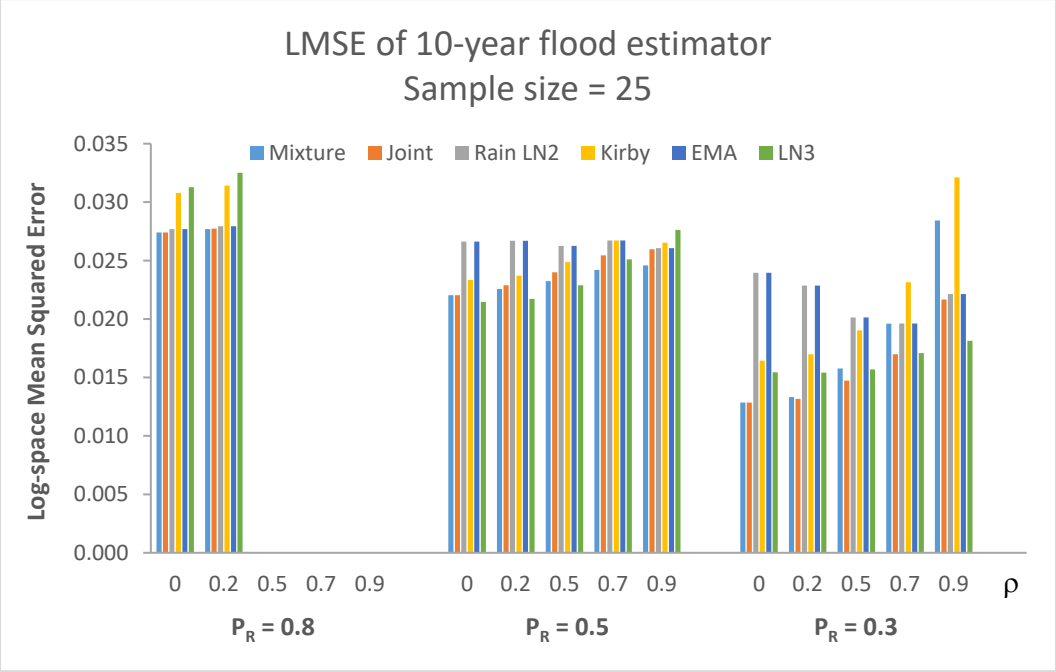


Figure 2.D-1 Log-space MSE of 10-Year Flood Estimator, n = 25, CV ratio = 2.48. Results for $P_R = 0.8$ with $\rho \geq 0.5$ are omitted because samples have too few snowmelt events. Samples that have an insufficient number of R' and S' values (≤ 5) were dropped.

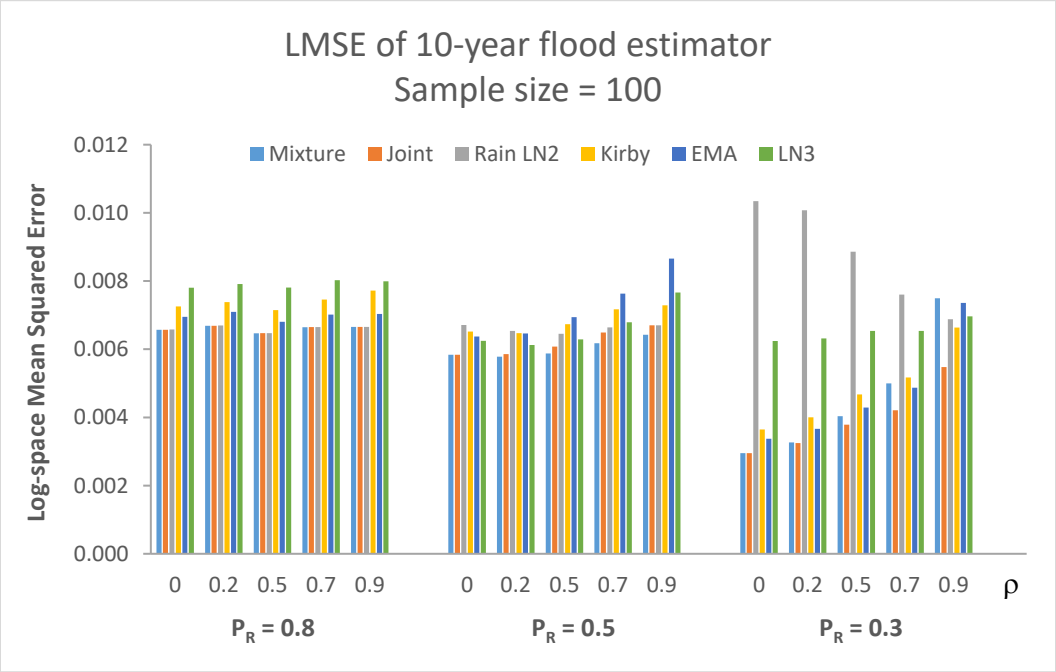


Figure 2.D-2 Log-space MSE of 10-Year Flood Estimator, n = 100, CV ratio = 2.48

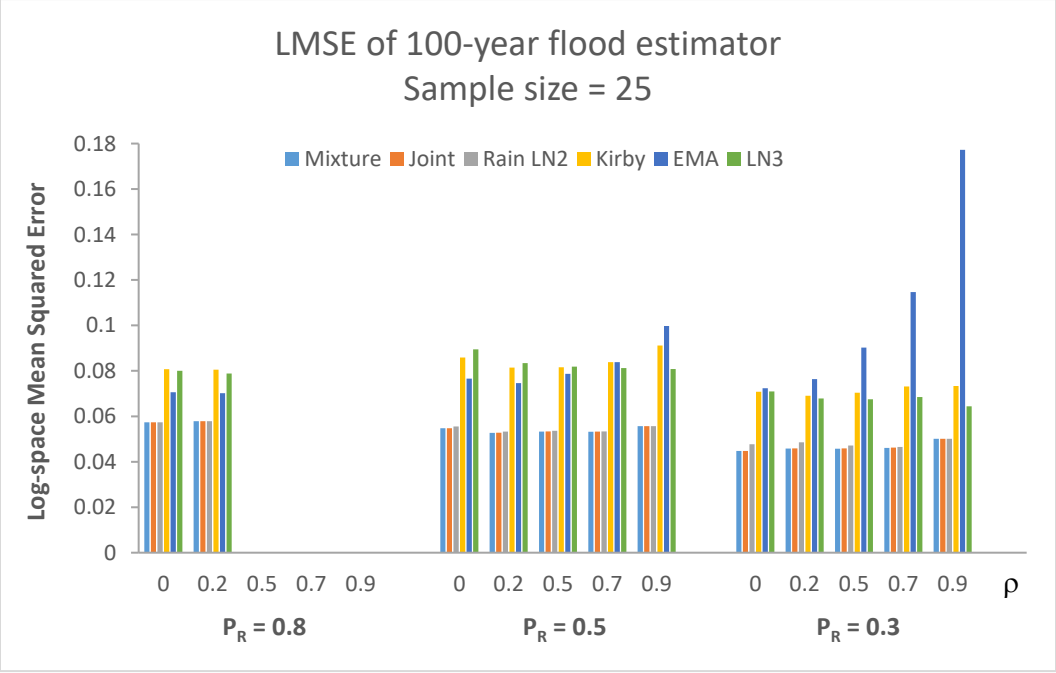


Figure 2.D-3 Log-space MSE of 100-Year Flood Estimator, n = 25, CV ratio = 2.48.

Results for $P_R = 0.8$ with $\rho \geq 0.5$ are omitted because samples have too few snowmelt events.

Samples that have an insufficient number of R' and S' values (≤ 5) were dropped.

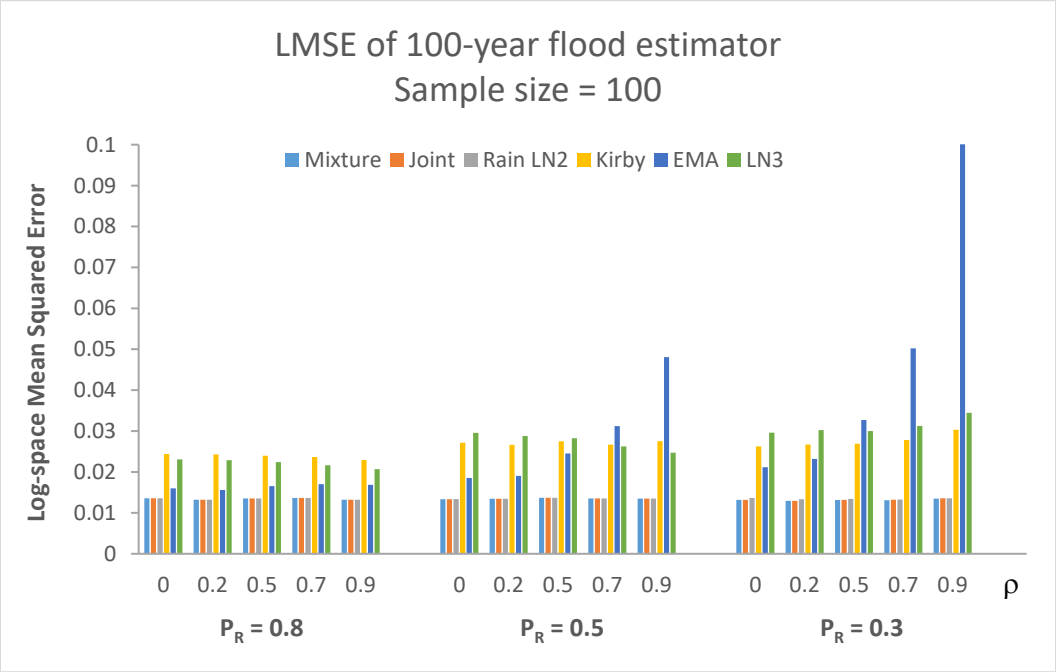


Figure 2.D-4 Log-space MSE of 100-Year Flood Estimator, n = 100, CV ratio = 2.48

Results in Figure 5 and 6 for n = 50, and in Figures 2.D-1 to 2.D-4 corresponding to n = 25 and 100, generate similar conclusions.

For 10-year flood estimation, mixture and joint estimators often perform best, except when $\rho \geq 0.7$ and $P_R = 0.3$. The rainfall-only estimator is a good alternative when $P_R = 0.8$. The performance of EMA is affected by both P_R and the sample size. When n = 25 and rainfall doesn't dominant the large flood risk, EMA doesn't work well for 10-year flood estimation; when n = 100, EMA performs better than Kirby with low correlations ($\rho < 0.5$).

For 100-year flood estimation are discussed in Section 4.2. The performance of mixture, joint, and rainfall-only estimators are essentially indistinguishable and much better than the other estimators that use only the annual maximum series. If only the annual maximum series is available, EMA looks to be the best for $\rho \leq 0.5$; however, $\rho \geq 0.7$ should be rare in practice when R and S correspond to different seasons.

The performances of the Kirby estimator and the single LN3 estimator are about the same, though Kirby works better than single LN3 when snowmelt events dominant ($P_R = 0.3$) and sample size is large (n = 100).

So, as noted in Chapter 2, in situation where one source of floods provides the largest annual maxima, and a second source provides most of the small floods, it is reasonable to split the annual maximum series into distinct physical processes and model them separately. When both rainfall and snowmelt floods records are available, the simple mixture estimator is adequate. Among the three estimators that only use the annual maximum series, the EMA estimator is the best choice for 100-year flood estimation when the annual maximum series results from two distinct flood sources with different simple distributions.

CHAPTER 3

FLOOD FREQUENCY ANALYSIS IN THE CONTEXT OF CLIMATE CHANGE

Abstract

A fundamental assumption in classical flood frequency analysis is that geomorphologic characteristics, and hydroclimatic and hydrologic statistics for a watershed, are stable over time. This allows hydrologists to assume available data represent a single time-invariant population of extreme events. Increasingly, people are concerned about climate change and climate variability. Thus, they challenge the validity of the assumption of hydrologic stationary. *So the question is, what should be done?*

This study evaluates different flood frequency estimation methods that can incorporate climate change. Several estimators employ time-varying parameters that can capture trends; others ignore trends, use a limited flood window, or adopt a safety factor (e.g. 25% increase). A Monte Carlo re-sampling study considers log-Pearson type III (LP3) estimators of the 100-year-flood some 25 years beyond the end of a 100-year flood record. Though modest trends (within ±0.25% per year) in a 100-year annual peak-flow record are not statistically detected 76% of the time, the 100-year flood will be significantly underestimated/overestimated if such trends are neglected. Estimating a trend parameter for the mean and especially the variance increases the variance of flood-quantile estimators; whereas it decreases the bias that results when trends are neglected. Clearly, flood risk management in an uncertain world will be a challenge.

Table of Contents

1. Introduction

A fundamental assumption in classical flood frequency analysis (FFA) is that the geomorphologic and climatic characteristics within the watershed remain stable over time. This allows the hydrologist to assume that the data represent a single population, and we can use a time-independent distribution to describe the extreme events over the period of record and into the future. However, this premise may not be valid due to various anthropogenic influences and climate change. Climate change has been a concern in hydrologic studies for decades [Bulletin 17B; IACWD, 1982]. A change in the climate can alter the distribution of many of the factors affecting floods (e.g., precipitation, snow cover, soil moisture content, sea level, glacial lake conditions, and vegetation). Thus, it may change the characteristics of flood distributions [IPCC, 2012, p.175]. Many studies have discussed if stationarity is still a reasonable assumption in flood frequency analysis [see Salas and Obeysekera, 2014; Serinaldi and Kilsby, 2014]. Others have applied statistical methods to test for trends in annual maximum series [Lins and Slack, 1999; Jain and Lall, 2001; Koutsoyiannis and Montanari, 2007; Hirsch and Ryberg, 2012; Mentaschi, et al. 2016; Hirsch and Archfield, 2016].

The purpose of this paper is to consider statistical methods for identifying and accounting for nonstationarity (trends) in annual peak flow data. A first step is to acknowledge the difference between climate variability and climate change [Stedinger and Griffis, 2008]. By neglecting the serial correlation in flood series (or the cross correlations among concurrent floods in multi-site tests), trend analyses can overestimate the statistical significance of computed trends [Hamed, 2008; Khaliq et al., 2009]. Particularly in a long record, long-term persistence can lead to greatly overstating the statistical significance of observed trends [Cohn and Lins, 2005].

The claim in *Science* of the death of stationarity by Milly et al. [2008] caused much debate [Lins and Cohn, 2011; Matalas, 2012; Koutsoyiannis and Montanari, 2014; Montanari and Koutsoyiannis, 2014; Milly et al. 2015; Serinaldi and Kilsby, 2015]. For analyzing persistence and variability in hydroclimatic records, stationarity will most likely be the paradigm for some time; while climate change does result in nonstationarity, so do urbanization and development factors. Change in flood risk is not new. Engineering facilities such as dikes and reservoirs alter flows, and land use changes affect flood volumes. As a result, the assessment of causes of changes in floods is complex and difficult [IPCC, 2012, p.175]. Regarding flood risk management, a broad range of anthropogenic influences should be considered when evaluating flood risk, in addition to climate change [Vogel and Walter, 2011].

A critical issue is the adjustment needed to incorporate climate change into policy-making based on flood-risk analyses. Rosner et al. [2014] adopt the concept of economic opportunity cost to calculate the "expected regret", rather than computing net benefits by using a decision tree; there the probabilities of two different choices (over engineered or under prepared) correspond to type I and type II error in the trend test.

This paper reviews existing literature, discusses several proposed flood frequency analysis methods, and includes a Monte Carlo evaluation of the performance of different methods.

2. FFA in a Changing World

Many hydrologists believe that anthropogenic climate change has not appreciably impacted the magnitude or frequency of fluvial floods at this time [IPCC, 2012, p.13]. Still, ongoing flood risk studies try to address the nonstationarity issue in frequency analysis. If there is no statistically significant trend detected, it seems

reasonable to use the entire flood record for a site under the assumption of stationarity with traditional methods. Otherwise, the impacts of climate change need to be incorporated into flood risk assessment, recognizing how flood risk is evolving over time.

Projections of the 100-year flood under climate change can be derived with a number of different analyses. These include coupled analysis: Greenhouse gas emission scenario - General circulation model (GCM) - Hydrological model (e.g. precipitation-runoff model) - Flood frequency analysis with a trend. The implementation of that model chain requires a large effort with high complexity, including the difficulty climate models have in representing current and future precipitation. In this paper, we consider several simple statistical procedures for extrapolating flood risk.

2.1 Ignore any Change/Trend: Stationarity

The Log-Pearson type III (LP3) distribution is recommended by the national guidelines for Determining Flood Flow Frequency (Bulletin 17B) -- the likelihood of climate change is generally dismissed in those guidelines [IACWD, 1982, p.6].

Assuming the annual maximum series Q follows a LP3 distribution, the logarithm of floods Q has a Pearson type III (P3) distribution:

$$\ln(Q) = X \sim P3[\mu, \sigma, \gamma]$$

Here μ, σ, and γ are the fixed location, scale, and shape parameters of the P3 distribution.

2.2 LP3 with Time-dependent Parameters

Many papers have discussed the possibility and treatment of trends [Olsen et al., 1999; Cohn and Lins, 2005; Hirsh and Ryberg, 2012; Hirsch and Archfield, 2016].

The most widely discussed method to incorporate climate change into statistical models is via simple, often linear, models of the change in the parameters of the probability distribution of floods with time [Stedinger and Griffis, 2011; Prosdocimi et al., 2014]. Studies have applied this method to the log-normal, GEV, or log-Pearson type III (LP3) distributions [Renard, et al. 2006; El Adlouni et al., 2007; El Adlouni and Ouarda, 2007; Cunderlik et al. 2007; Ribatet et al. 2009; Ouarda and El Adlouni. 2011; Vogel and Walter, 2011; Rootzen and Katz, 2013; Rosner et al., 2014; Salas and Obeysekera, 2014]. To be consistent with Bulletin 17B [IACWD 1982], this paper uses the LP3 distribution, and models the distribution of ln-*Q* as a function of time with:

$$\ln\left(Q_t\right) = X_t \sim P3\left[\mu(t), \sigma(t), \gamma\right] \qquad (1)$$

$$\mu_t = \mu_0 + \beta_1 t \qquad (2)$$

$$\ln\left(\sigma_t^2\right) = \ln\left(\sigma_0^2\right) + \beta_2 t \qquad (3)$$

Here t is the elapsed time between the current year and the first year in a flood record; μ_0 and σ_0 are the initial location and scale parameters for ln(Q) [Coles, 2001, pp.105-108; El Adlouni et al., 2007]. The change in the peak flow distribution is a combined effect from the change in the log-space mean (Eqn.2) and variance (Eqn.3). To avoid having a negative estimated log-space variance, a logarithm transformation is employed in Eqn.3 so that the logarithm of the variance of X_t is modeled. Eqn.3 can be written as:

$$\sigma_t^2 = \sigma_0^2 \exp\left(\beta_2 t\right) \qquad (4)$$

2.3 Short-term Window

Another approach to FFA is to employ only a limited recent set of recent floods (e.g. 30 years) for frequency analysis – the implicit statement is that floods are representative of a given climate state and samples from a different climate state

should not be considered [Raff et al., 2009]. Obtaining reliable extreme events estimators is likely to be problematic if record lengths are limited. Especially when in the context of climate change, the selection of time scale is usually subjective. On the other hand, sometimes it is suggested that hydrologists use the available recorded instrumental record along with "historical" and paleo-flood records that reflect the long-term climate pattern to increase the range of available information [Stedinger and Cohn, 1986].

2.4 Safety Factor

One simple solution is to increase the design flood by a specified amount, a so called "safety factor", to account for uncertainty [Olsen, 2006]. Some research in European countries refined this concept by using the projected changes in peak flow from GCMs. For instance, one federal state in Germany, the UK, and two river basin authorities in Belgium have developed guidelines that increase the design flood by 15%, 20%, and 30% with the respective time horizons of 2050, 2085, and 2100, according to their selected climate system scenarios [Maden et al., 2013]. However, decision makers would be hesitant to spend additional funds without strong evidence supporting increased flood magnitudes in the future. This is also the reason that a risk-based approach is needed for planning levee or dam designs that includes uncertainty.

2.5 Precipitation-related Models

An alternative to statistical models based on annual maximum flood series, is a physically based rainfall-runoff simulation. In order to explore the physical-causal basis for changes in peak flows, adding the precipitation information and watershed characteristics to flood frequency analysis is worth considering [Bloschl, 2006]. Based on the IPCC special report of the risks of extreme events, the overall most consistent

trends toward heavier precipitation events are found in North America (likely increase over the continent). Compared to the low confidence in significant trends in fluvial floods, there is medium confidence that anthropogenic influence has contributed to intensification of extreme precipitation at the global scale [IPCC, 2012, chap.3, p.13]. Thus, the concern is that the precipitation produced by intense events will increase yielding a corresponding increase in high flows [Mallakpour and Villarini, 2015; Hirsh and Archfield, 2016].

The regional regression is one method to connect the 100-year peak streamflow with the mean annual precipitation, the drainage area, and the watershed slope [Capesius and Stephens, 2009]. For instance, the generalized least-squares (GLS) regression equation (based on 141 stations) in the Colorado-mountains hydrologic-region yielded the model:

$$Q_{100} = 10^{-0.46} A^{0.75} S^{0.14} P^{1.35} \qquad (5)$$

where Q_{100} is the 100-year peak streamflow (cubic feet per second, cfs), A is the drainage area (mi^2), P is the mean annual precipitation (inch), and S is the mean watershed slope (%). For the case of climate change, a 100-year flood based on at-site gauged data can be adjusted using this linear model, in which an adjusted mean annual precipitation (P) could be obtained from corrected GCM projection. Such a flow adjustment would implicitly include land-cover changes that would result from a wetter or drier climate, to the extent the original data set captures those relationships. As mentioned before, the problem of uncertainties in the GCM downscaling (from global to regional) process has not been well addressed under the nonstationary assumption, which will result in less accurate precipitation forecasts [see Benestad et al., 2007; Gutmann, et al., 2014]. Because equations such as (5) are not highly accurate, in this case the equation should be used to estimate the change in Q_{100} computed with a gaged record that would results from a change in P.

3. Worlds with Trend in the Mean and Variance of Flood Series

The focus of this paper is a Monte Carlo study that evaluates the performance of flood frequency analyses in a world where there is a trend in the log-space mean flow, the log-space variance, or both. The following sections describe development of parameters sets describing possible trends, and the different combinations of parameters considered in the Monte Carlo analysis.

3.1 Models of Trend in Mean and Variance of X

A simulation experiment considering climate change should consider reasonable forms and rates for any changes. Many studies have considered a change in the mean, or the mean of the logarithms. However, people believe climate change not only affects the moisture content of storms and their intensity, but also results in a regional shift in the causes or character of storms, such as winter rain or summer thunderstorms versus snowmelt [Stedinger and Griffis, 2011]. Thus, it is appropriate to allow both the mean and variance of the log-space annual maximum series to vary with time.

It is not easy to determine if a change has occurred in the distribution of annual peaks during the period of record. However, correctly forecasting what change will occur in the future is even more difficult. For this study, we consider the simple linear trend models in Eqn. 2 and 3 with rates based on past North American observations.

Regression with the following two models (based upon Eqn. 2 and 4) were used to estimate historical trends in the mean and variance:

$$X_t = \alpha_1 + \beta_1 t + \varepsilon_{1,t} \qquad (6)$$

$$\left(X_t - \hat{\mu}_t\right)^2 = \alpha_2 \exp\left(\beta_2 t\right) + \varepsilon_{2,t} \qquad (7)$$

Here X_t is the logarithm of a recorded flow at the gauged station; $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are random errors; $\hat{\mu}_t$, equals to $\hat{a}_1 + \hat{b}_1 t$, is the estimated expected value of X in year t

provided by Eqn. 6, when substituted into Eqn. 2. α_1 in Eqn. 6 is the same as μ_0 in Eqn. 2, and α_2 in Eqn. 7 is the same as σ_0^2 in Eqn.3 and 4. The value of β_1 for Eqn.6 were estimated using ordinary least squares (OLS), which works well if the error $\varepsilon_{1,t}$ are independent and identically distributed [Stedinger and Griffis, 2011]. The value of β_2 for Eqn. 7 was estimated using nonlinear least squares. Eqn. 7 was based on Eqn. 4 instead of Eqn. 3 to estimate the β_2 because the use of the logarithms of $(X - \hat{\mu}_t)^2$ are not well behaved given that the mode of $(X - \hat{\mu}_t)$ is zero.

Trend computations used the 472 HCDN-2009 (Hydro-Climatic Data Network) stations. These stations were screened to exclude sites where human activities affect the natural flow. Thus, the streamflow series retained primarily reflects prevailing meteorological conditions for specified years. The record lengths range from 50 to over 110 years.

Figure 1 displays the computed trend parameters $\hat{\beta}_1$ in the logarithms of annual peak flows (percent change per year). Figure 2 shows computed trends $\hat{\beta}_1$ in the variance of the logarithm of annual peak flows (percent change per year).

The 472 annual maximum series exhibit little serial correlation. Of the 472 sites, 86 (or 18.2 % of the series) have β_1 (trend in log-space mean) that are statistically significant from zero at the 5% level and 33 (7.0 % of series) of the β_2 (trend in log-space variance) values are significant at the 5% level. Clearly there appears to be historical trends in both the mean and the variance at some sites. The number of stations that show positive or negative trends is higher than expected (5%). However, it is not reasonable to conclude those are significant trends, because we have not taken the spatial correlation (cross-correlation) and the serial correlation into account [Vogel, et al. 2011].

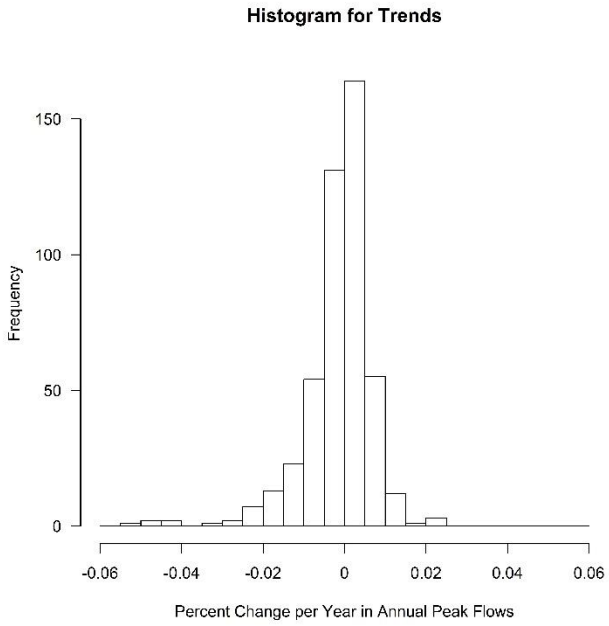


Figure 1: Estimated Trend Magnitudes in Log-space Mean of Annual Maximum Series

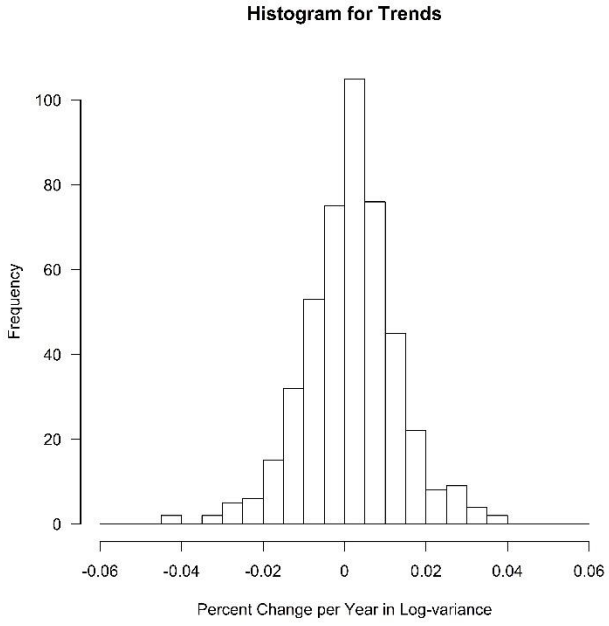


Figure 2: Estimated Trend Magnitudes in Log-space Variance of Annual Maximum Series

Consider an analysis of the computed trends. For each site *i*, the estimated trend $\hat{\beta}_i$ equals the sum of true trend $\tilde{\beta}_i$ (signal) and a β-estimation error $\tilde{\xi}_i$.

$$\hat{\beta}_i = \tilde{\beta}_i + \tilde{\xi}_i \qquad i = 1, 2, 3, \ldots, m \qquad (8)$$

where m = 472 is the total number of sites. The true trend $\tilde{\beta}_i$ in Eqn. 8 is a fixed value for each *i*.

We employ linear least squares to estimate the $\hat{\beta}_1$ in Eqn.6, the variance of that estimator is:

$$Var\left[\tilde{\xi}_i\right] = Var\left[\hat{\beta}_i\right] = \frac{\sigma_{\varepsilon_i}^2}{\sum_{j=1}^{n_i}\left(x_{i,j} - \bar{x}_i\right)^2} = \frac{\sigma_{\varepsilon_i}^2}{\left(n_i - 1\right)S_{X_i}^2} \qquad j = 1, 2, 3, \ldots, n_i \qquad (9)$$

where $x_{i,j}$ is the log-space peak flow of site *i* in year *j*; n_i is the record length of site *i*; and ε is the error term in the linear regression (Eqn. 6) [Devore, 2014, pp.512].

Eqn. 7 employs the non-linear least squares with Levenberg-Marquardt numerical algorithm to calculate $\hat{\beta}_2$ and its standard error, because the residuals are highly skewed. Among the 472 stations, using nonlinear least squares for Eqn. 7, for 10 sites $\hat{\beta}_2$ had very large standard errors due to one or more extreme values at the beginning or end of the series. To reduce the impacts of these extreme values, we employ a robust non-linear regression using iterated reweighted least squares [Green, 1984; Gentle, 2007, p.233; Ruckstuhl, 2016] with the Gauss-Newton algorithm. In addition, one site was dropped for estimating $\hat{\beta}_2$ due to its extreme large standard error, which is unlikely to occur under the assumption that the trend in the log-variance is monotonic and consistent in the future.

The variance of $\hat{\beta}$ across all 472 records ($Var[\hat{\beta}]$) equals the variance of $\tilde{\beta}$ ($Var[\tilde{\boldsymbol{\beta}}]$) plus the average of the variance of the error term $\tilde{\xi}_i$ ($\frac{1}{m}\sum_{i=1}^{m} Var[\tilde{\xi}_i]$), That is,

$$Var\left[\hat{\beta}\right] = Var\left[\tilde{\beta}\right] + \frac{1}{m}\sum_{i=1}^{m} Var\left[\tilde{\xi}_i\right] \quad (10)$$

These numbers for the trend computations for the mean and variance are summarized in Table 1.

Table 1. Analysis for the two trend computations across all 472 sites*

Regression	$E[\hat{\beta}]$ Across sites	$Var[\hat{\beta}]$ Across sites	Avg. across sites of variance of ξ	$Var[\tilde{\beta}]$ Across sites	% Explained by ξ
Eqn. 6, β_1	-1.44E-03**	7.50E-05	3.25E-05	4.25E-05	43%
Eqn. 7, β_2	1.34E-03**	1.34E-04	1.14E-04	2.03E-05	85%

* The analysis of β_2 is based on 471 sites

** Mean of β_1 and β_2 estimators are both significantly different from 0 with type I error of 5%

In Figure 1, 43% of the variability of $\hat{\beta}_1$ comes from the noise ($\tilde{\xi}_1$); in Figure 2, 85% of the variability of $\hat{\beta}_2$ comes from the noise ($\tilde{\xi}_2$). Thus, most of the observed variability in $\hat{\beta}_2$ is due to random error $\tilde{\xi}_2$, rather than variability that appears to be due to $\tilde{\beta}_2$. However, the estimated variability in the trends in the variance $\tilde{\beta}_2$ appears to be smaller than the variance of $\tilde{\beta}_1$ for the mean ($2.03\text{x}10^{-5}$ versus $4.25\text{x}10^{-5}$); what is critical is that the estimators of the trend in the variances $\hat{\beta}_2$ have a larger estimation error than do the estimated trends in the means $\hat{\beta}_1$ ($1.14\text{x}10^{-4}$ versus $3.25\text{x}10^{-5}$).

3.2 Cases Considered in Monte Carlo Study

Based on the analysis of estimated trends reported in Section 3.1, the variance of $\tilde{\beta}_1$ is $4.25\text{x}10^{-5} = (6.5\text{x}10^{-3})^2$ and the variance of $\tilde{\beta}_2$ is $2.03\text{x}10^{-5} = (4.5\text{x}10^{-3})^2$. In our Monte Carlo study, we assume the expected trends $\text{E}[\tilde{\beta}]$ is zero and consider the trends in log-space mean $\tilde{\beta}_1$ and variance $\tilde{\beta}_2$ as large as ±2 times their standard deviation across sites which yields roughly for $\tilde{\beta}_1$ and $\tilde{\beta}_2$ an interval of -1% to 1% change per year.

Based on the configurations of different trend magnitudes in the log-space mean and log-space variance, the Monte Carlo study includes four cases:

Case 1: Positive trends in log-mean (μ_x)

Case 2: Positive trends in log-mean (μ_x) and log-variance (σ^2_x)

Case 3: Negative trends in log-mean (μ_x)

Case 4: Negative trends in log-mean (μ_x) and log-variance (σ^2_x)

For cases I and III, we selected 0, ±0.25%, ±0.5%, and ±1% per year as the trend in the log-mean (μ_x); in cases II and IV, the equivalent trends are introduced in the log-variance (σ^2_x). To estimate the 100-year flood in the future, 25 and 50 years hence, the trends in mean and variance are assumed consistent across the past 100 years and the next 50 years.

When there is no trend in σ^2_x, even when there is a trend in the mean μ_x, the coefficient of variation (CV) of the annual maximum flood series (Q_t) is constant; that means the mean and standard deviation in real-space change with the same ratio in each year t. When there are trends in σ^2_x, the real-space coefficient of variation and real-space skewness of Q varies with time [Griffis and Stedinger, 2007a].

Table 2 lists the differences between the initial 100-year flood and the projected 100-year flood for different cases (I and II) and trend magnitudes with a modest trend of 0.25%, and an extreme trend of 1%. When there were 1% trends in both the log-space mean and variance, the original 100-year flood (the 1% event at T = 0) becomes the 3.59% event at the end of 25 years, and the 9.23% event at the end of 50 years.

Table 2. Ratio of the projected 100-year flood to the initial 100-year flood and the AEP of the initial 100-year flood in the projected year after T years

Trends (%)	Q (T=25)	Q (T=50)	AEP (T=25)	AEP (T=50)
0, 0	1.00	1.00	1.00%	1.00%
0.25, 0	1.06	1.13	1.24%	1.53%
0.25, 0.25	1.11	1.24	1.43%	2.02%
1.00, 0	1.27	1.63	2.27%	4.96%
1.00, 1.00	1.54	2.48	3.59%	9.23%

3.3 Random Sample Generation

The Monte Carlo study employed N = 10,000 n-year records, where n was 100. They were generated by re-sampling the annual maximum series from the station at Driftwood Bridge Sinnemahoning Creek, Sterling Run, PA [HCDN-2009, USGS gage number: 01543000]. Its record ran from 1914 to 2016, exactly 103 years. The observations approximately follow the LP3 distribution with log-space (natural base) mean = 9.06 (log-cfs), log-space variance = 0.34, and log-space coefficient of skewness = +0.36, without any unusual observations. So as to obtain representative results this station was selected because the peak-flow record was well behaved, with a skew near zero and the data when plotted indeed looked LP3. The lag-1 autocorrelation of the whole series is just -0.084, which is not significantly different from zero with type I error of 5%. Even if the original series had a trend, it is lost in the process of randomly drawing observations from the historical record. Figure 3 show a probability plot for the annual peak flows in log-space with exceedance probability in normal scale. A slight positive skew in the log-flows is evident.

The trends selected in Section 3.2 were added to both the mean and variance of the re-sampling data set (X'_t) by shifting and scaling the log-space historical observations. Equation 11 through 13 below provide an algorithm for computing the new annual maximum series with a designated trend, X_t:

$$\frac{X_t - \mu_t}{\sigma_t} = \frac{X'_t - \bar{X}'}{S_{X'}} \qquad (11)$$

This can be written:

$$X_t = \mu_t + \frac{\sigma_t}{S_{X'}}\left(X'_t - \bar{X}'\right) \qquad (12)$$

If there is no trend in the variance ($S_X = \sigma_t$), then:

$$X_t = (\mu_t - \bar{X}') + X'_t \qquad (13)$$

Here $\bar{X}'$ and $S_{X'}$ are the mean and standard deviation of the original data series X'_t; μ_t and σ_t are the designated mean and standard deviation for X_t, for each t. These 100 re-sampled observations with added trends were considered as the flood records in the Monte Carlo study.

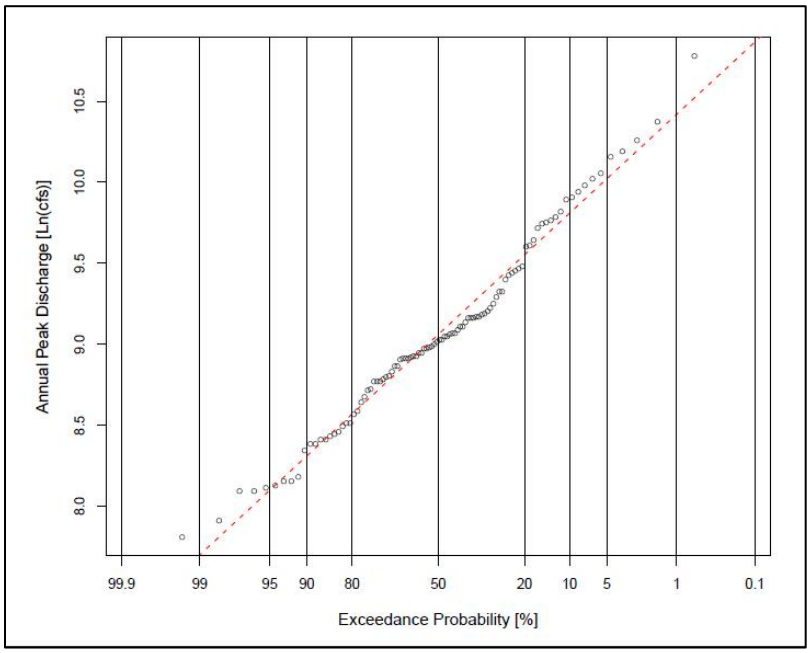


Figure 3: Probability Plot of Annual Maximum Series

Driftwood Bridge Sinnemahoning Creek, Sterling Run, PA 1914-2013

4. Estimators

Section 2 described four methods that can be used to project flood risk. The methods employ: (1) up-to-date records assuming stationarity, (2) annual maximum series with known or estimated time-dependent flood distribution parameters, (3) only the most recent 30 years of record, and (4) European safety factors. First the paper considers estimation of 100-year flood ($Q_{100,T}$) some T = 25 or 50 years in the future; second it considers the annual exceedance probability ($AEP_{0.01,T}$) of the 100-year flood estimator, $\hat{Q}_{100,T}$, for T = 25 or 50 years.

Method 1 assumes the flood series is stationary. the 100-year flood estimator is calculated by employing the sample mean, variance, and coefficient of skewness of X_t to fit a LP3 distribution.

$$\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{t=0}^{n-1} X_t \tag{14}$$

$$\hat{\sigma}^2 = S_X^2 = \frac{1}{n-1}\sum_{t=0}^{n-1}\left(X_t - \bar{X}\right)^2 \tag{15}$$

$$\hat{\gamma} = G_X = \frac{n}{(n-1)(n-2)}\sum_{t=0}^{n-1}\left(X_t - \bar{X}\right)^3 \tag{16}$$

Method 2 employs three estimators: "Trend_0" assumes the true trends are known, which is unrealistic in most hydrologic studies and practices; "Trend_1" assumes we don't know the exact trend in the mean and estimates that trend using linear regression based upon equation 6. It also assumes the variance has a fixed value; "Trend_2" estimates a trend in both the mean and variance (see Appendix 3.C).

When computing the "Trend_2" estimator, the trend β_2 in the log-space variance was not estimated using Eqn.7 to avoid having a negative or an extremely large estimated variance for a 25 or 50 years projection. In the Monte Carlo study, we estimate α_2 and β_2 directly with non-linear least squares based upon:

$$\left(X_r - \hat{\mu}_r\right)^2 = \alpha_2 \frac{r\exp\left(0.5\beta_2 r\right) + \exp\left(-0.5\beta_2 r\right)}{\exp\left(0.5\beta_2 r\right) + r\exp\left(-0.5\beta_2 r\right)} + \varepsilon_{2,r} \quad \text{r = t+1 = 1, 2, …, n} \tag{17}$$

When β_2 is 0 in equation 17, the predicted variance is the initial variance α_2. With Eqn.17, when β_2 is positive and t is large, $E[(X_t - \hat{\mu}_t)^2] = \alpha_2 r$, which is linear and won't increase as dramatically as the model in Eqn.7; when β_2 is negative and t is large, $E[(X_t - \hat{\mu}_t)^2] = \alpha_2 / r$, which while always positive, approaches zero slowly. For $|\beta_2| \leq 1\%$ and $0 \leq t \leq 150$, the variance models in equation 7 and 17 are almost indicial.

The "safety factors" with 100-year flood estimator are ±25% for the 25 years projection and ±30% for the 50 years projection. Thus, $\hat{Q}^*_{100,t} = \pm 1.25\ \hat{Q}_{100,t}$ and

$\hat{Q}^{*}_{100,t} = \pm 1.3\ \hat{Q}_{100,t}$. Where $\hat{Q}_{100,t}$ is estimated assuming stationarity and the traditional sample moments (Eqn.14-16).

To fit an LP3 distribution requires an estimator of the log-space skewness coefficient. With "Trend_0", "Trend_1", and "Trend_2", the estimate of the skewness coefficients was the de-trended observations (see Appendix 3.C); for other estimators, all the available data (30 or 100 years) is used to compute the skewness coefficient.

The estimators of 100-year flood were calculated using Eqn.18 below. Here the parameters of the LP3 distribution are estimated by the method of moment (MOM) of the log series X_t.

$$\hat{Q}_{100,t} = \exp\left(\hat{X}_{100,t}\right) = \exp\left[\hat{\mu}_t + \hat{\sigma}_t K_{0.99}\left(\hat{\gamma}\right)\right] \qquad (18)$$

In Eqn.18, $K_{0.99}(\hat{\gamma})$ is the frequency factor, which is the 99 percentile of a standard P3 variate with skew coefficient $\hat{\gamma}$, mean zero, and variance 1 [Stedinger et al., 1993, p.18.20]. See Appendix 3.C for the details of the computation of $\hat{\mu}_t$ and $\hat{\sigma}_t$ for each estimator.

5. Criteria

The traditional concern is the precision of estimator $\hat{Q}_{100,t}$. Another concern is the true probability with which the design flow is exceeded [Stedinger, 1980; Stedinger, 1997; Beard, 1997 and 1998]; thus, when the distribution of Q is known, one should consider:

$$A\tilde{E}P_{0.01,t} = \Pr\left(Q_t \geq \hat{Q}_{100,t}\right) \qquad (19)$$

To compare the performance of the six estimators from both points of view, the mean squared error (MSE) of log $\hat{Q}_{100,t}$ and mean square deviation (MSD) of log $A\tilde{E}P_{0.01,t}$ from the target were calculated using Equations 20 and 21:

$$LMSE(\hat{Q}_{100,t}) = \frac{1}{N}\sum_{i=1}^{N}\left[\ln\left(\hat{Q}_{100,t}\right)_i - \ln\left(Q_{100,t}\right)\right]^2 \qquad (20)$$

$$LMSD(\tilde{AEP}_{0.01,t}) = \frac{1}{N}\sum_{i=1}^{N}\left[\ln\left(\tilde{AEP}_{0.01,t}\right)_i - \ln(0.01)\right]^2 \quad (21)$$

Where LMSE or LMSD represents the MSE or MSD of the natural logarithm of $\hat{Q}_{100,t}$ or $\tilde{AEP}_{0.01,t}$; The true value of $\tilde{AEP}_{0.01,t}$ is 0.01 by construction, however $\tilde{AEP}_{0.01,t}$ may be highly variable.

6. Results

6.1 Precisions of 100-year flood estimators

Figures 4 through 7 display the boxplots of the common logarithm (base 10) of the estimated 100-year floods ($\hat{Q}_{100}$) for the 25-year projection with our 4 cases. Results for the 50-year projection are included in the Appendix 3.B. The Monte Carlo simulation used N = 10,000 replicates. In the boxplots, whiskers extend 1.5 times the interquartile range beyond the top and bottom of the box (which are at the sample quartiles). Any "outliers" (observations beyond the whiskers) are plotted individually. With such a large number of replicates, there are almost always some outliers; however, the pattern and number of outliers will vary with different samples, whereas the median, sample quartiles, and whiskers are very stable.

Tables of the LMSE and the squared bias of $\hat{Q}_{100}$ in log-space are included in Appendix 3.A. They allow evaluation of the bias and the contribution of the bias to the total LMSEs.

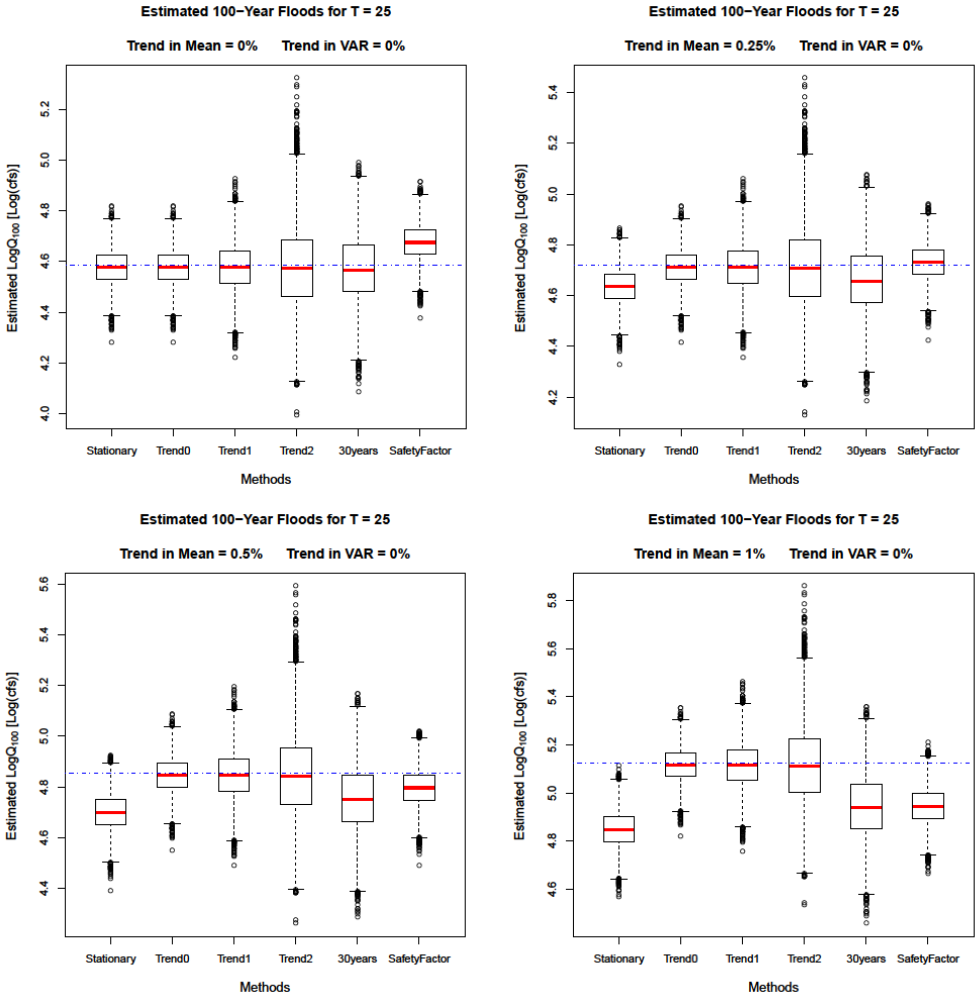
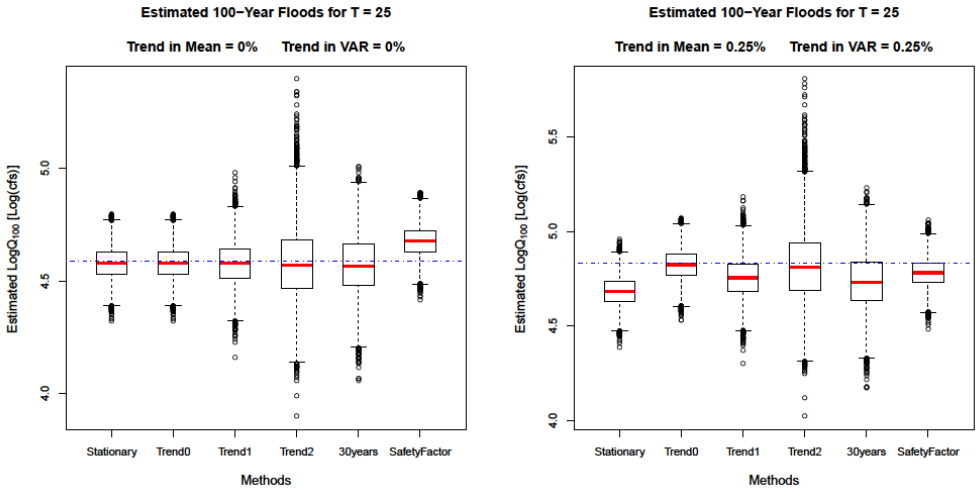


Figure 4: Boxplots of estimated Q_{100} for T = 25 (Case 1-Positive trends in μ_x)



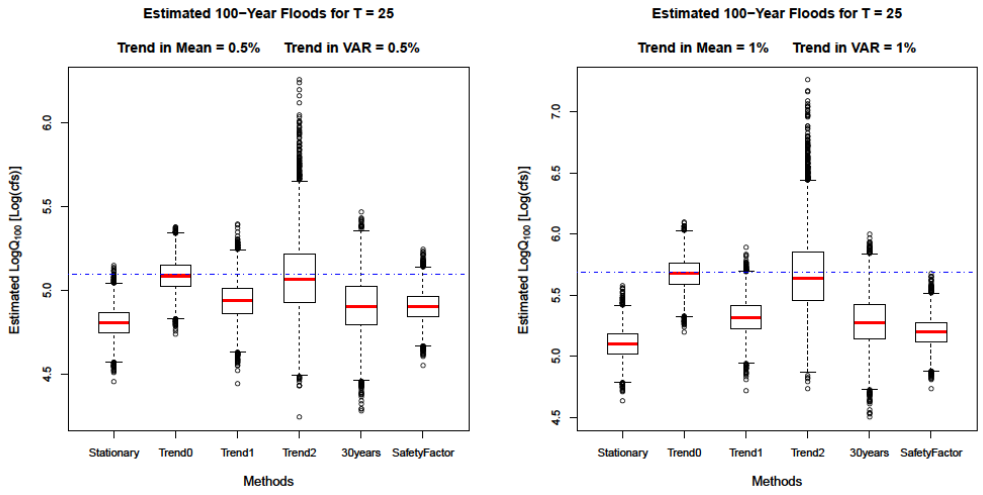


Figure 5: Boxplots of estimated Q_{100} for T = 25 (Case 2-Positive trends in μ_x and σ_X^2)

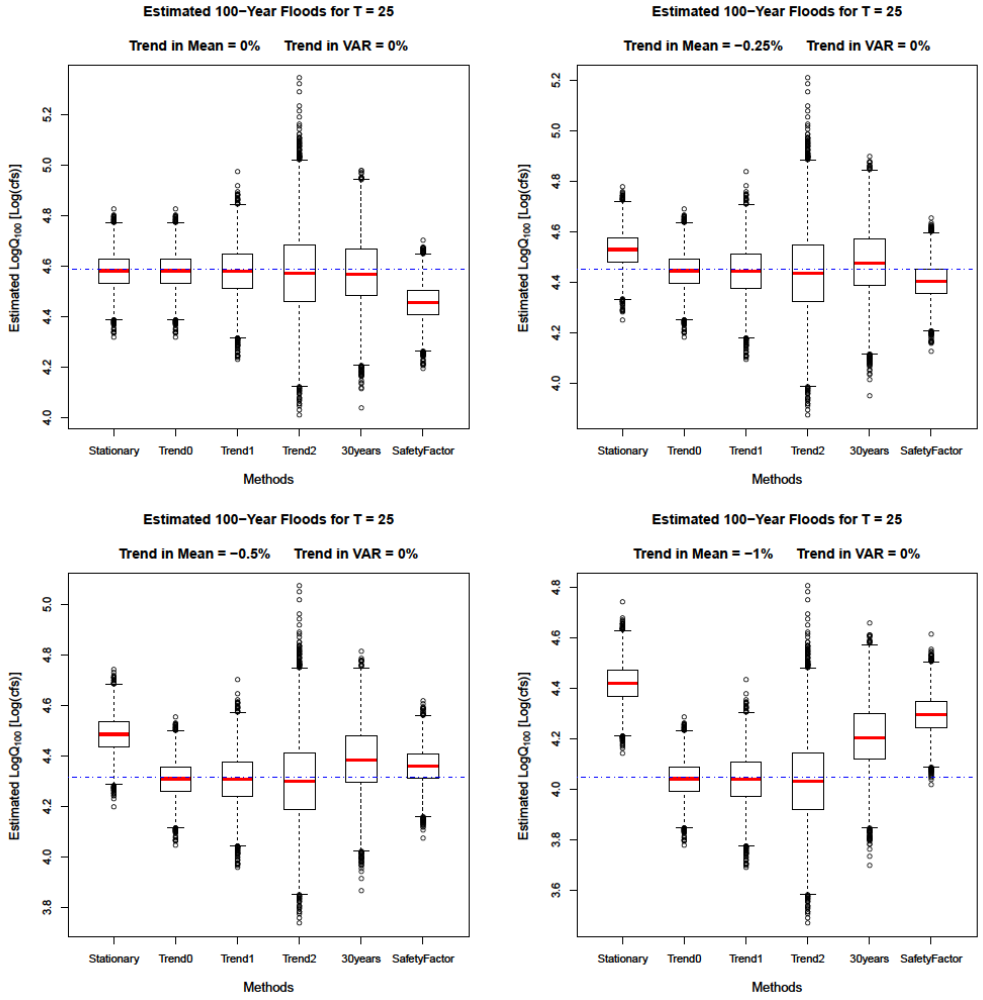


Figure 6: Boxplots of estimated Q_{100} for T = 25 (Case 3-Negative trends in μ_x)

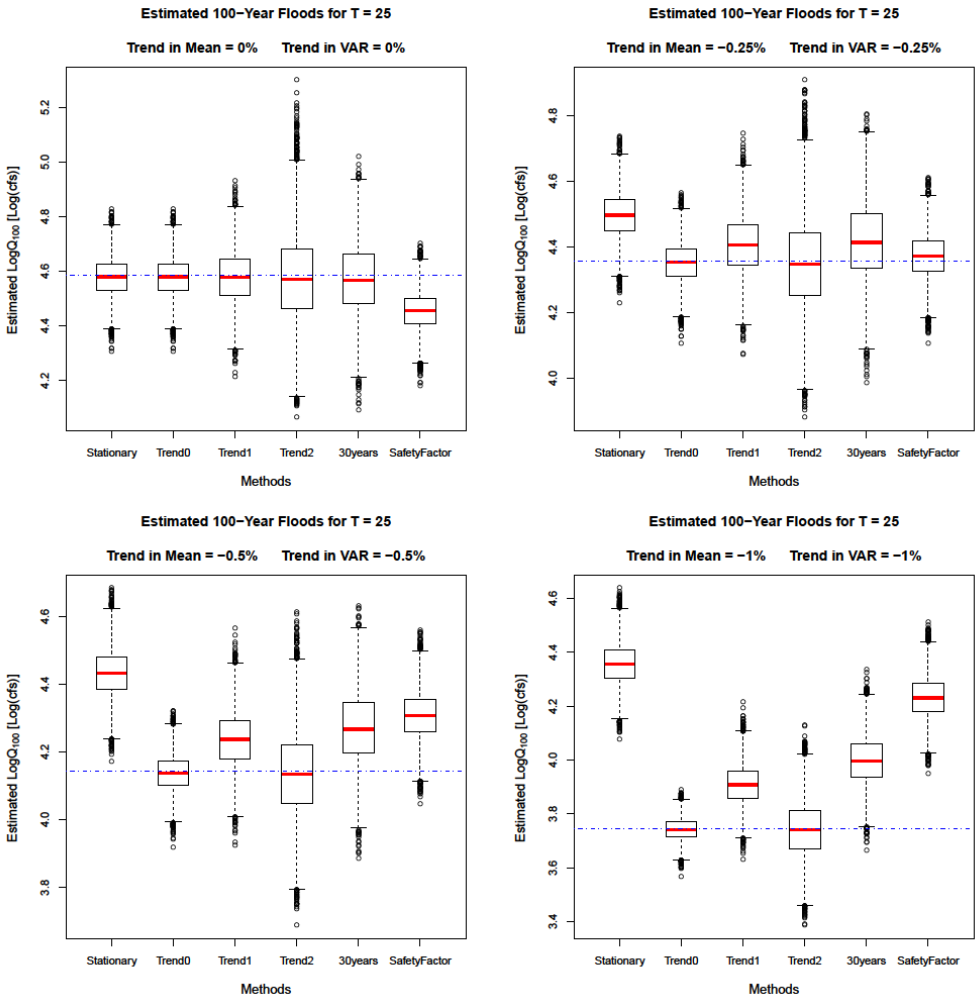


Figure 7: Boxplots of estimated Q_{100} for T = 25 (Case 4-Negative trends in μ_x and σ_X^2)

In Figures 4-7 display the distribution of quantile estimators. “Stationary”, “Trend_0” and “Safety Factor” employ the estimated mean, standard deviation and a skewness coefficient from the data. They have relatively few outliers and thus the distributions of the estimators have “thin tails” in all four cases. “30-year Record” is like “Stationary” except it uses only the latest 30 years of record, and thus has a larger variance and more outliers.

"Stationary" estimator will significantly underestimate the 100-year flood with the positive trend models, and overestimate the 100-year flood with the negative trend models. In those cases, bias often dominates the LMSE.

Trend_0 is included to allow a comparison with Trend_1 and Trend_2, that include estimation of a trend in the mean, and the mean and variance, respectively. Because "Trend_0" uses the magnitudes of any trend in the mean and variance, it has the smallest LMSEs for $\hat{Q}_{100}$ in all 4 cases.

It is interesting that generally Trend_1 does not do much worse than "Stationarity" when there are no trends, whereas it can do much better when there is a trend in the mean. Trend_1 does about as well as Trend_0 when these is only a trend in the mean. On the other hand, "Trend_2" has a large variance and a large number of outliers: the distribution has very heavy tails. This was expected because of the need to estimate the trend parameter for the variance. So, while Trend_2 is relatively unbiased in all cases considered, in terms of LMSE, it is generally inferior to Trend_1 for most cases except for $\beta_1 = \beta_2 = -1\%$ (see Figure 4 through 7, Figure 3.B-1 through 3.B-4).

The "30-year Record" estimator employs only the most recent 30 years' records, but it doesn't perform well. It is still biased because the estimation is based on past data rather than future observations without an adjustment for trends; while this estimator's bias is less than the "Stationary" estimator that uses all 100 years, its variance is three times larger due to the limited record employed.

The safety factor method with a factor of ±0.25 only works effectively when $\beta_1 = 0.25\%$, $\beta_2 = 0$, or when $\beta_1 = \beta_2 = -0.25\%$.

Overall, "Trend_0" is an ideal situation that is not really possible and always does the best. Comparing the LMSEs of the quantile estimators in the figures and Appendix 3.A, employing "Stationary" when the magnitude of the trends are small (within ±0.25% per year) generally does well. "Trend_1" is often recommended

except when there are suspected trends in both mean and perhaps the variance. It really does relatively well except for the case $\beta_1 = \beta_2 = \pm 1\%$. While, "Trend_2" will be a feasible alternative for the extreme situations $\beta_1 = \beta_2 = \pm 1\%$, because it estimates a slope parameter for the variance β_2, it has a relatively large variance.

Table 3 lists the proportion of the samples (among 10,000 replicates) whose $\widehat{\boldsymbol{\beta}}_{\mathbf{2}}$ is significantly different zero with a two-side α = %5 test. For cases 2 and 4, when the trends are small (±0.25%), only 5% to 7% of the 10,000 random samples exhibited trends in the variance that were statistically significant, which implies "Trend_1" would be selected as the appropriate estimator for those records.

Table 3. Relative frequency of samples that have significant $\widehat{\boldsymbol{\beta}}_{\mathbf{2}}$ values (t-test with type I error is 5%)

Trends (%)	Case 1	Case 2	Case 3	Case 4
0.00	0.05	0.04	0.05	0.05
0.25	0.05	0.07	0.05	0.07
0.50	0.05	0.15	0.05	0.14
1.00	0.05	0.41	0.05	0.42

6.2 Deviations from anticipated AEP

Figure 8-11 provide column charts of the LMSD of the estimated annual exceedance probability $A\tilde{E}P_{0.01,T}$ for 100-year flood estimator $\hat{Q}_{100,T}$ in T = 25 years. It yields relatively consistent conclusions with the LMSE of $\hat{Q}_{100}$. The LMSD of $A\tilde{E}P_{0.01,T}$ in T = 50 years are shown in Appendix 3.B. The results and conclusions were essentially the same.

The difference in the configuration of trends (cases 1-4) affect the LMSE of $\hat{Q}_{100}$ and the LMSD of $A\tilde{E}P_{0.01}$. The LMSE and LMSD calculated using "Trend_0", "Trend_1", and "Trend_2" keep constant if there is no trend in σ^2_x (Figures 8 and 10,

Tables 3.A-1 and 3.A-3), but the LMSE and LMSD of "Trend_1" will increase with trends when trends appear in both μ_x and σ^2_x (Figures 9 and 11, Tables A2 and A4).

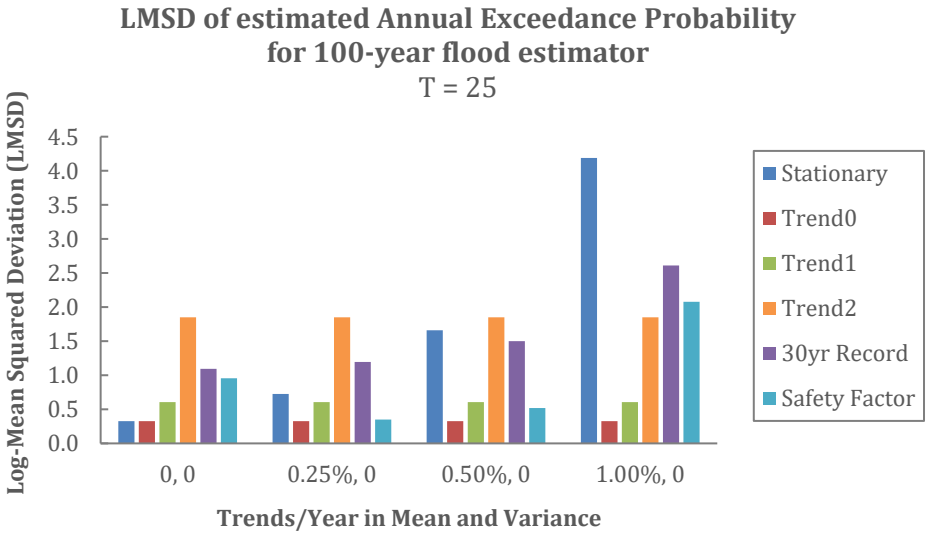


Figure 8: LMSD of Annual Exceedance Probability for 1% Event

(Case 1-Positive trends in μ_x)

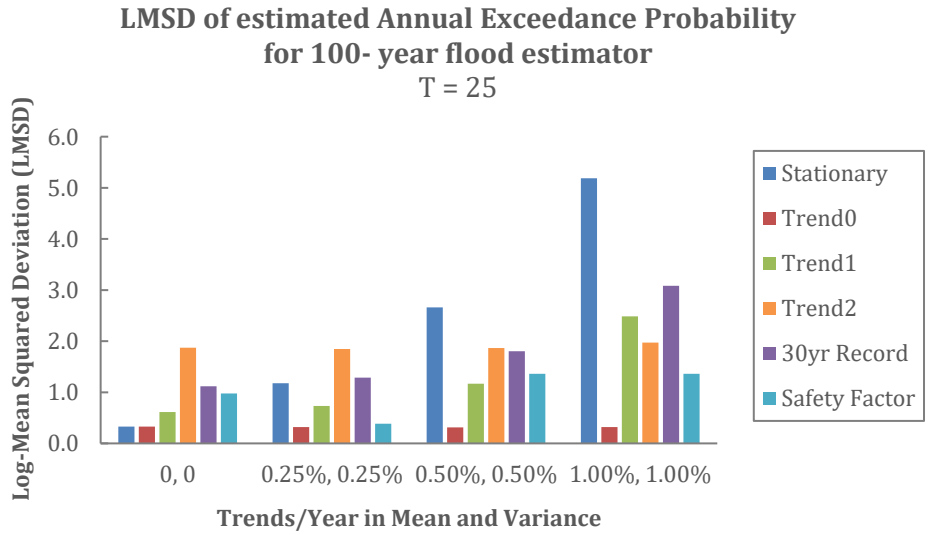


Figure 9: LMSD of Annual Exceedance Probability for 1% Event

(Case 2-Positive trends in μ_x and σ_X^2)

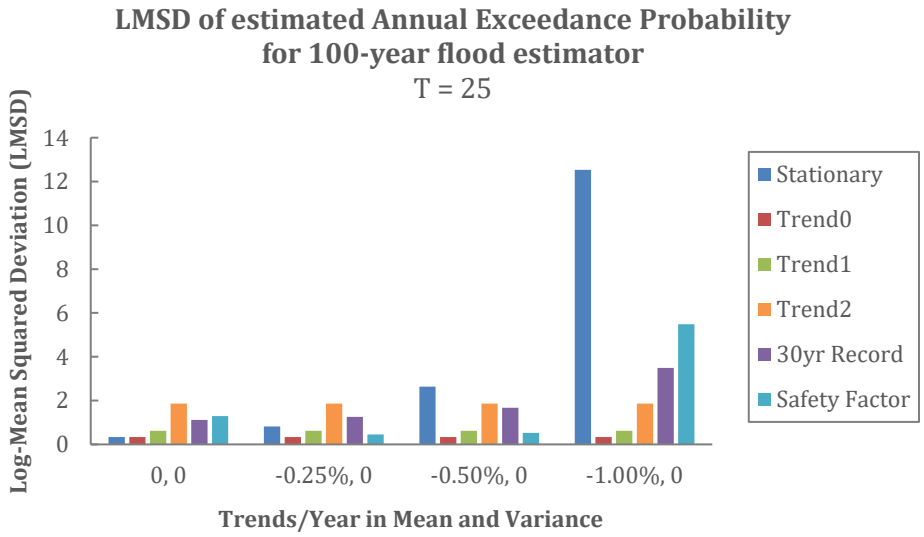


Figure 10: LMSD of Annual Exceedance Probability for 1% Event

(Case 3-Negative trends in μ_x)

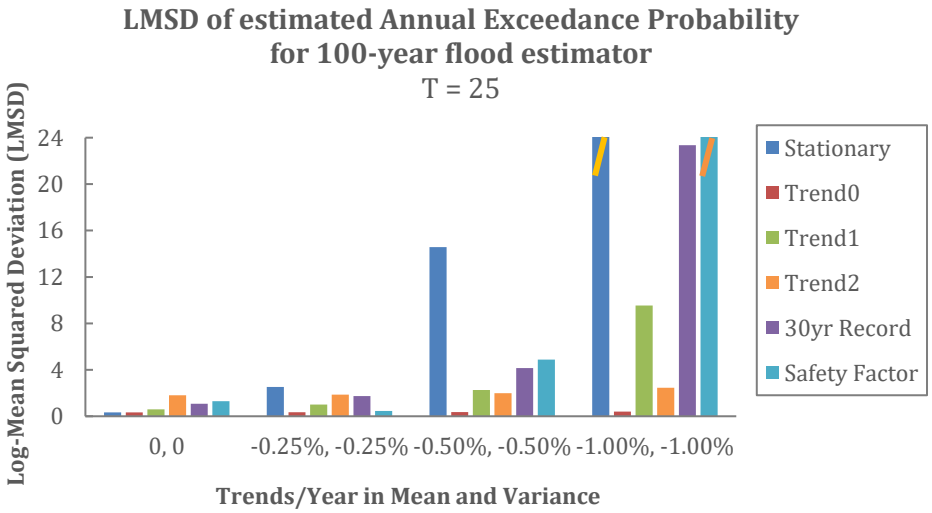


Figure 11: LMSD of AEP for 1% Event (Case 4-Negative trends in μ_x and σ_X^2)

*When $\beta_1 = \beta_2 = -1.00\%$, the LMSD for "Stationary" and "Safety Factor" are much larger than 24

6.3 Type II Error

For cases 1-4, wherein trends in the mean and variance are most modest (within ±0.5% per year), such small trends might not be detected by statistical trend tests. Consider the probability that one concludes there is no trend (in the mean) based on the generated 100-year records when the trends do exist. Assuming the estimator of the trend in an annual maximum series follows a normal distribution and obtaining the standard error (SE ≈ 0.002 for all selected trends) for the estimator $\hat{\beta}_1$ in Eqn. 6 from a Monte Carlo simulation with 100,000 replicates, we list the type II errors of the z-test (with $\alpha = 0.05$) in Table 4 for selected cases. Over 3/4 of the time, the modest trend (±0.25% change per year) in the mean with a 100-year record won't be detected; however, the 100-year flood will be significantly underestimated/overestimated if such trends are neglected.

Table 4. Type II errors of trend detection for log-mean flood $\widehat{\beta}_1$

	Mean	**Var.**	**Mean**	**Var.**	**Mean**	**Var.**
Trend (per year)	**±0.25%**	**0**	**0.25%**	**0.25%**	**-0.25%**	**-0.25%**
Type II error	76.1%		78.3%		73.6%	
Trend (per year)	**±0.5%**	**0**	**0.5%**	**0.5%**	**-0.5%**	**-0.5%**
Type II error	29.5%		40.6%		19.6%	

7. Conclusion

This paper considers several simple approaches to flood frequency analysis when records may have trends, and evaluates their performance via a Monte Carlo re-sampling study. Unfortunately, probability of failing to detect a modest trend (i.e. the trend is not statistically significant) in annual maximum series is high (see Table 4); however, the 100-year flood estimator could in expectation appreciably underestimate/overestimate the true value if such trends are neglected.

When the trend magnitudes are known (based upon GCM or regional analyses) and employed to update the parameters of LP3 distribution year by year, the 100-year flood estimators (Trend_0) have the smallest log-space MSE. In practice, it is difficult to know the exact magnitude of trends. The alternative is to estimate trends based on the record. Trend_1 estimates the trend in the mean of the annual maximum log-series; it performs well for most cases except the extreme scenario ($\beta_1 = \beta_2 = \pm 1\%$). Trend_2 estimates the trend in both the mean and the variance of the log-space peak-flow, resulting in a large variance in the 100-year flood estimator and the largest MSE when trends were modest (within ±0.5% per year): estimation of a trend in the variance is difficult, and should only be done with caution. Use of a shorten record (30 years in our case) is not a good choice because of the shorter record and the past observations don't accurately reflect the trends in the future without an adjustment. Using a safety factor does well when a safety factor happens to have about the right value; it is not recommended. Thus, the selection of a best estimator depends on the specific cases, which is essentially a trade-off between the variance and the bias of the estimators.

A critical step is to obtain accurate estimates of the magnitudes of any trend for the future annual maximum flood series. Besides relying on the data from flood records, we naturally think of a regional estimator or getting information from atmospheric models that represent the change in climate state more closely. As mentioned before, incorporating precipitation records or long-term climatic indexes (such as temperature) into flood frequency analysis remains an opportunity for extreme quantile estimation whose promise has yet to be captured. Alternatively, reliable downscaling methods that are needed to employ GCMs results for a specific region could be applied. Clearly, flood risk management in an uncertain world will be a challenge.

References

Al-Futaisi, A. and J. R. Stedinger (1999). Hydrologic and economic uncertainties and flood-risk project design. *J. Water Resour. Plann. Manage.,* 125(6): 314-324.

Beard, L. R. (1997). Estimating flood frequency and average annual damage. *Journal of Water Resources Planning and Management,* 123(2): 84-88

Beard, L. R. (1998). Expected probability and annual damage estimators. *Journal of Water Resources Planning and Management*, 124: 365-366

Benestad, R. E., I. Hanssen-Bauer, and E. J. Forland (2007). An evaluation of statistical models for downscaling precipitation and their ability to capture long-term trends. *Int. J. Climatol*. 27: 649-665

Bloschl, G. (2006). Statistical Upscaling and Downscaling in Hydrology. *Encyclopedia of Hydrological Sciences.1:9*. Published Online.

Capesius, J. P. and Stephens, V. C. (2009). Regional regression equations for estimation of natural streamflow statistics in Colorado. Technical Report 2009-5136, U.S. Geological Survey.

Cohn, T. A. and Lins, H. F. (2005). Natures style: Naturally trendy. *Geophysical Research Letters*, 32.5: L23402.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values.* Springer Series in Statistics. Springer-Verlag, London.

Cunderlik J., Jourdain V., Ouarda T.B.M.J. et Bobée B. (2007). Local non-stationary flood-duration-frequency modelling. *Can. Water Resour. J. / Rev. Can. Ressour. Hydr.,* 32 (1): 43-58.

Devore, Jay. L. (2014). *Probability and Statistics for Engineering and the Sciences, Ninth Edition.* Chapter 12, Simple Linear Regression and Correlation. Cengage Learning, Boston, MA

El Adlouni S., Ouarda T.B.M.J., Zhang X., Roy R. et Bobée B. (2007). Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resour. Res.,* 43 (3) : W03410. doi: 10.1029/2005WR004545

El Adlouni S. et Ouarda T.B.M.J. (2008). Comparaison des méthodes d'estimation des aramètres du modèle GEV non stationnaire / Comparison of methods for estimating the parameters of the non-stationary GEV model. *Rev. Sci. Eau,* 21(1): 35-50.

Greene, W. H. (2012), *Econometric Analysis (7th Edition)*, Chapter 20: Serial Correlation, pp. 909-912. Prentice Hall, Upper Saddle River, NJ.

Griffis, V. W. and Stedinger, J. R. (2007). Incorporating climate change and variability into Bulletin 17B LP3 model. *In World Environmental and Water Resources Congress.* ASCE.

Gutmann, E., T. Pruitt, M. P. Clark, L. Brekke, J. R. Arnold, D. A. Raff, and R. M. Rasmussen (2014). An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.*, 50, doi: 10.1002/2014WR015559

Hamed, K. H. (2008). Trend detection in hydrologic data: the Mann Kendall trend test under the scaling hypothesis. *Journal of hydrology*, 349.3:350-363.

Hirsch, R. M. and Ryberg, K. R. (2012). Has the magnitude of floods across the USA changed with global CO_2 levels? *Hydrological Sciences Journal*, 57.1:1-9.

Hirsch, R. M. and S. A. Archfield (2016). Not higher but more often. *Natural Climate Change*, Vol 5, 198-199.

IACWD (1982). Bulletin 17b of the hydrology subcommittee: Guidelines for determining flood flow frequency. Technical report, U.S. Geological Survey.

IPCC (2012). *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK, and New York NY, USA.

Jain, S. and Lall, U. (2001). Floods in a changing climate: does the past represent the future? *Water Resources Research*, 37.12:3193-3205.

Khaliq, M. N., Ouarda, T. B. M. J., Gachon, P., Sushama, L., and A. St-Hilaire (2009). Identification of hydrological trends in the presence of serial and cross correlations: A review of selected methods and their application to annual flow regimes of Canadian rivers. *Journal of Hydrology*, 368.1:117-130.

Koutsoyiannis, D. and Montanari, A. (2007). Statistical analysis of hydroclimatic time series: Uncertainty and insights. *Water Resour. Res.*, 43.5: W05429.

Koutsoyiannis, D., and A. Montanari (2014), Negligent killing of scientific concepts: The stationarity case, *Hydrol. Sci. J.*, 60, 1174–1183, doi:10.1080/02626667.2014.959959.

Lins, H. F., and T. A. Cohn (2011), Stationarity: Wanted dead or alive? *J. Am. Water Resour. Assoc.*, 47, 475–480.

Lins, H. F. and Slack, J. R. (1999). Streamflow trends in the United States. *Geophysical Research Letters*, 26.2:227-230.

Madsen, H., Lawrence, D., Lang, M., Martinkova, M., and Kjeldsen, T. (2013). *A Review of Applied Methods in Europe for Flood-Frequency Analysis in a Changing Environment*. FLOODFREQ COST Action ES0901. Center for Ecology & Hydrology.

Mallakpour, I. and G. Villarini (2015), The changing nature of flooding across the central United States, *Nature Climate Change*, 5, 250–254.

Matalas, N. C. (2012), Comment on the announced death of stationarity, *Water Resour. Plann. Manage.*, 138, 311–312.

Mentaschi, L. et al. (2016). The transformed-stationary approach: a generic and simplified methodology for non-stationary extreme value analysis. *Hydrol. Earth Syst. Sci.*, 20, 3527–3547

Milly, P.C.D., et al. (2008). Stationarity is dead: wither water management? *Science.* 319, 573-574

Milly, P.C.D., et al. (2015). On critiques of "Stationarity is dead: wither water management?" *Water Resour. Res.*, 51, 7785-7789, doi:10.1002/2015WR017408.

Montanari, A., and D. Koutsoyiannis (2014), Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resour. Res.*, 50, 9748–9756, doi:10.1002/2014WR016092.

Olsen, J. R., J. R. Stedinger, N. C. Matalas, and E. Z. Stakhiv (1999), Climate variability and flood frequency estimation for the Upper Mississippi and Lower Missouri Rivers, *J. of the Amer. Water Resources Association*, 35(6), 1509-1524.

Olsen, J. R. (2006). Climate change and floodplain management in the United States. *Climate Change*, 76:407-426.

Ouarda, T. B., and S. El Adlouni (2011), Bayesian nonstationary frequency analysis of hydrological variables, *Journal of the American Water Resources Association*, 47(3), 496–505, doi:10.1111/j.1752-1688.2011.00544.x.

Prosdocimi, I., Kjeldsen, T. R., and Svensson, C. (2014). Non-stationarity in annual and seasonal series of peak flow and precipitation in the UK. *Natural Hazards and Earth System Sciences*, 14:1125-1144.

Raff, D. A., Pruitt, T., and Brekke, L. D. (2009). A framework for assessing flood frequency based on climate projection information. *Hydrol. Earth Syst. Sci.*, 13:2119-2136.

Renard, B., M. Lang, and P. Bois (2006), Statistical analysis of extreme events in a non-stationary context via a Bayesian framework: case study with peak-over-threshold data, *Stochastic Environmental Research and Risk Assessment*, 21(2), 97–112, doi:10.1007/s00477-006-0047-4.

Ribatet M., Ouarda T.B.M.J., Sauquet E. et Grésillon J.M. (2009). Modeling all exceedances above a threshold using an extremal dependence structure: Inferences on several flood characteristics. *Water Resour. Res.*, 45: W03407. doi: 10.1029/2007wr006322

Rootzen, H. and Katz, R. W. (2013). Design life level: Quantifying risk in a changing climate. *Water Resour. Res.*, 49.9:5964-5972.

Rosner, A., Vogel, R. M., and Kirshen, P. H. (2014). A risk based approach to flood management decisions in a nonstationary world. *Water Resour. Res.*, 50.3:1928-1942.

Salas, J. D. and Obeysekera, J. (2014). Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *Journal of Hydrologic Engineering*, 19.3:554-568.

Serinaldi, F., and C. G. Kilsby (2015). Stationarity is undead: Uncertainty dominates the distribution of extremes. *Advances in Water Resources*, 77:17-36

Stedinger, J. R. (1980). Fitting Log Normal Distributions to Hydrologic Data. *Water Resour. Res.*, 16(3): 481-490

Stedinger, J. R., and T. A. Cohn (1986). Flood frequency analysis with historical and paleoflood information. *Water Resour. Res.*, 22(5): 785-793

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou (1993). *Frequency analysis of extreme events.* Handbook of hydrology, Chapter 18. McGraw-Hill, Inc., New York, U.S.

Stedinger, J. R. (1997). Expected probability and annual damage estimators. *Journal of Water Resources Planning and Management*, 123(2): 125-135

Stedinger, J. R. and Griffis, V. W. (2008). Flood frequency analysis in the United States: Time to update. *Journal of Hydrologic Engineering*, 13.4:199-204.

Stedinger, J. R. and Griffis, V. W. (2011). Getting from here to where? Flood frequency analysis and climate. *JAWRA Journal of the American Water Resources Association*, 47.3:506-513.

U.S. Geological Survey (2011), GAGE II: Geospatial Attributes of Gages for Evaluating Streamflow, https://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml

Vogel, Richard M., C. Y. and Walter, M. (2011). Nonstationarity: Flood magnification and recurrence reduction factors in the United States. *JAWRA Journal of the American Water Resources Association*, 47.3:464-474.

Zuelling, R. E., and T. S. Schmidt (2012), Characterizing invertebrate traits in wadeable streams of the contiguous US: differences among ecoregions and land uses, *Freshwater Science,* 31(4):1042–1056. DOI: 10.1899/11-150.1

Appendices for Chapter 3

This appendix contains results for the Monte Carlo study in Chapter 3 that are not included in the Chapter 3. Results for T = 25 years are summarized by figures in Chapter 3. This appendix contains a number of tables that support those figures. In particular Appendix 3.A contains tables for results for a 25-year projection. Appendix 3.B contains figures and tables for results for a 50 year projection (T = 50) when estimating Q_{100} and $AEP_{0.01}$ for that future year.

The tables and figures use the following notation:

$\hat{Q}_{100}$ is the estimator of the 100-year flood;

$\hat{X}_{100}$ is the estimator of the logarithm of 100-year flood, $\hat{X}_{100} = \ln\left(\hat{Q}_{100}\right)$;

$\tilde{AEP}_{0.01}$ is the estimator of the annual exceedance probability for the 1% event.

Based on the choices of different trend magnitudes in log-space mean and variance, four cases are addressed in the Monte Carlo study:

Case 1: Positive trends in log-mean (μ_x)

Case 2: Positive trends in log-mean (μ_x) and log-variance (σ_t^2)

Case 3: Negative trends in log-mean (μ_x)

Case 4: Negative trends in log-mean (μ_x) and log-variance (σ_t^2)

Appendix 3.A – *Monte Carlo Study Results for 25-year Projection*

Table 3.A-1 through 3.A-4 report the mean of $A\tilde{E}P_{0.01}$, the log-space mean squared values for $A\tilde{E}P_{0.01}$ and $\hat{Q}_{100}$, as well as the squared bias of $\hat{X}_{100}$.

Table 3.A-1: Monte Carlo Study Results Case 1 (Positive trends in μ_x), T=25

Estimators	**Mean** $A\tilde{E}P_{0.01}$	**LMSD(** $A\tilde{E}P_{0.01}$**)**	**LMSE(** $\hat{Q}_{100}$**)**	**Bias²(** $\hat{X}_{100}$**)**
Trend in mean = 0, Trend in variance = 0				
Stationary	1.23%	0.3265	0.0263	0.0003
Trend_0	1.23%	0.3265	0.0263	0.0003
Trend_1	1.39%	0.6045	0.0487	0.0004
Trend_2	2.02%	1.8493	0.1440	0.0003
30yr Record	1.77%	1.0944	0.0896	0.0014
Safety Factor	0.56%	0.9566	0.0679	0.0419
Trend in mean = 0.25%, Trend in variance = 0				
Stationary	2.23%	0.7262	0.0647	0.0384
Trend_0	1.23%	0.3265	0.0263	0.0003
Trend_1	1.39%	0.6045	0.0487	0.0004
Trend_2	2.02%	1.8493	0.1440	0.0003
30yr Record	2.40%	1.1944	0.1065	0.0182
Safety Factor	1.05%	0.3503	0.0270	0.0007
Trend in mean = 0.50%, Trend in variance = 0				
Stationary	3.70%	1.6591	0.1550	0.1276
Trend_0	1.23%	0.3265	0.0263	0.0003
Trend_1	1.39%	0.6045	0.0487	0.0004
Trend_2	2.02%	1.8493	0.1440	0.0003
30yr Record	3.21%	1.5012	0.1421	0.0534
Safety Factor	1.83%	0.5194	0.0453	0.0180
Trend in mean = 1.00%, Trend in variance = 0				
Stationary	8.17%	4.1869	0.4301	0.3996
Trend_0	1.23%	0.3265	0.0263	0.0003
Trend_1	1.39%	0.6045	0.0487	0.0004
Trend_2	2.02%	1.8493	0.1440	0.0003
30yr Record	5.50%	2.6105	0.2659	0.1760
Safety Factor	4.37%	2.0767	0.1978	0.1673

Table 3.A-2: Monte Carlo Study Results Case 2 (Positive trends in μ_x and σ^2_x), T=25

Estimators	**Mean** $\tilde{AEP}_{0.01}$	**LMSD(** $\tilde{AEP}_{0.01}$**)**	**LMSE(** $\hat{Q}_{100}$**)**	**Bias²(** $\hat{X}_{100}$**)**
Trend in mean = 0, Trend in variance = 0				
Stationary	1.22%	0.3274	0.0263	0.0002
Trend_0	1.22%	0.3274	0.0263	0.0002
Trend_1	1.39%	0.6138	0.0493	0.0004
Trend_2	2.08%	1.8721	0.1468	0.0004
30yr Record	1.80%	1.1185	0.0921	0.0015
Safety Factor	0.55%	0.9756	0.0692	0.0432
Trend in mean = 0.25%, Trend in variance = 0.25%				
Stationary	2.96%	1.1779	0.1462	0.1148
Trend_0	1.21%	0.3186	0.0349	0.0003
Trend_1	2.06%	0.7314	0.0884	0.0303
Trend_2	2.04%	1.8468	0.1959	0.0006
30yr Record	2.80%	1.2873	0.1628	0.0505
Safety Factor	1.59%	0.3840	0.0447	0.0134
Trend in mean = 0.50%, Trend in variance = 0.50%				
Stationary	5.28%	2.6603	0.4770	0.4361
Trend_0	1.21%	0.3140	0.0468	0.0004
Trend_1	2.91%	1.1674	0.2000	0.1302
Trend_2	2.02%	1.8644	0.2672	0.0009
30yr Record	3.94%	1.8037	0.3265	0.1864
Safety Factor	3.24%	1.3628	0.2321	0.1912
Trend in mean = 1.00%, Trend in variance = 1.00%				
Stationary	10.1%	5.1858	1.8876	1.8141
Trend_0	1.20%	0.3186	0.0880	0.0004
Trend_1	5.04%	2.4842	0.8299	0.7259
Trend_2	2.03%	1.9724	0.5174	0.0020
30yr Record	6.33%	3.0842	1.0869	0.8652
Safety Factor	7.38%	3.8456	1.3363	1.2628

Table 3.A-3: Monte Carlo Study Results Case 3 (Negative trends in μ_x), T=25

Estimators	**Mean** $\tilde{AEP}_{0.01}$	**LMSD(** $\tilde{AEP}_{0.01}$**)**	**LMSE(** $\hat{Q}_{100}$**)**	**Bias²(** $\hat{X}_{100}$**)**
Trend in mean = 0, Trend in variance = 0				
Stationary	1.23%	0.3328	0.0268	0.0003
Trend_0	1.23%	0.3328	0.0268	0.0003
Trend_1	1.40%	0.6242	0.0505	0.0003
Trend_2	2.14%	1.8612	0.1481	0.0007
30yr Record	1.80%	1.1170	0.0919	0.0014
Safety Factor	3.14%	1.2958	0.1191	0.0926
Trend in mean = -0.25%, Trend in variance = 0				
Stationary	0.62%	0.8216	0.0586	0.0318
Trend_0	1.23%	0.3328	0.0268	0.0003
Trend_1	1.40%	0.6242	0.0505	0.0003
Trend_2	2.14%	1.8612	0.1481	0.0007
30yr Record	1.30%	1.2569	0.0943	0.0038
Safety Factor	1.68%	0.4500	0.0388	0.0120
Trend in mean = -0.50%, Trend in variance = 0				
Stationary	0.28%	2.6380	0.1798	0.1518
Trend_0	1.23%	0.3328	0.0268	0.0003
Trend_1	1.40%	0.6242	0.0505	0.0003
Trend_2	2.14%	1.8612	0.1481	0.0007
30yr Record	0.93%	1.6736	0.1173	0.0266
Safety Factor	0.82%	0.5233	0.0383	0.0104
Trend in mean = -1.00%, Trend in variance = 0				
Stationary	0.04%	12.5335	0.7707	0.7393
Trend_0	1.23%	0.3328	0.0268	0.0003
Trend_1	1.40%	0.6242	0.0505	0.0003
Trend_2	2.14%	1.8612	0.1481	0.0007
30yr Record	0.46%	3.4905	0.2287	0.1368
Safety Factor	0.14%	5.4862	0.3587	0.3274

Table 3.A-4: Monte Carlo Study Results Case 4 (Negative trends in μ_x and σ^2_x), T=25

Estimators	**Mean** $\tilde{AEP}_{0.01}$	**LMSD(** $\tilde{AEP}_{0.01}$**)**	**LMSE(** $\hat{Q}_{100}$**)**	**Bias²(** $\hat{X}_{100}$**)**
Trend in mean = 0, Trend in variance = 0				
Stationary	1.23%	0.3296	0.0265	0.0003
Trend_0	1.23%	0.3296	0.0265	0.0003
Trend_1	1.38%	0.5955	0.0481	0.0004
Trend_2	2.03%	1.8035	0.1414	0.0005
30yr Record	1.74%	1.0806	0.0883	0.0011
Safety Factor	3.14%	1.2941	0.1189	0.0927
Trend in mean = -0.25%, Trend in variance = -0.25%				
Stationary	0.32%	2.5249	0.1260	0.1010
Trend_0	1.23%	0.3414	0.0202	0.0002
Trend_1	0.87%	1.0031	0.0528	0.0119
Trend_2	2.10%	1.8696	0.1084	0.0003
30yr Record	0.97%	1.7371	0.0893	0.0184
Safety Factor	1.07%	0.4604	0.0259	0.0009
Trend in mean = -0.50%, Trend in variance = -0.50%				
Stationary	0.04%	14.570	0.4731	0.4467
Trend_0	1.24%	0.3573	0.0155	0.0001
Trend_1	0.51%	2.2641	0.0822	0.0467
Trend_2	2.20%	1.9936	0.0853	0.0002
30yr Record	0.45%	4.1475	0.1441	0.0855
Safety Factor	0.20%	4.8770	0.1713	0.1449
Trend in mean = -1.00%, Trend in variance = -1.00%				
Stationary	0.00%*	155.14	2.0134	1.9821
Trend_0	1.25%	0.4017	0.0093	0.0000*
Trend_1	0.14%	9.5396	0.1697	0.1417
Trend_2	2.52%	2.4500	0.0567	0.0001
30yr Record	0.04%	23.357	0.3820	0.3399
Safety Factor	0.00%*	91.769	1.2862	1.2548

* Values are extremely small, but not exactly zero

As discussed in Section 6, trends in the log-space variance generally increase the LMSD for $\tilde{AEP}_{0.01}$ and the LMSE for $\hat{Q}_{100}$. For "Stationary", "30-yr Record", and "Safety Factor" estimators, bias usually is the major factor that determines the efficiency of $\hat{Q}_{100}$. Variance is the main concern for "Trend_1" in case 1 and case 3, and for "Trend_2" in all four cases. In case 2 and case 4, both the bias and the variance contribute to the LMSE of "Trend_1" estimator. Still, "Trend_0" has the smallest LMSD for $\tilde{AEP}_{0.01}$ and LMSE for $\hat{Q}_{100}$, and has mean values of $\tilde{AEP}_{0.01}$ that are

closest to 1% in almost every scenario. This is an ideal situation that is unlikely to happen in practice.

Appendix 3.B - *Monte Carlo Study Results for 50-year Projection*

This appendix reports the Monte Carlo study results for a 50-year projection. Figure 3.B-1 through 3.B-4 show the boxplots of the common logarithm (base 10) of the 100-year flood estimators for 4 cases; Figure 3.B-5 through 3.B-8 display and compare the LMSD of the $A\tilde{E}P_{0.01}$; Table 3.B-1 through 3.B-4 list the specific values of LMSD for $A\tilde{E}P_{0.01}$ and LMSE for $\hat{Q}_{100}$, the mean of $A\tilde{E}P_{0.01}$, and the squared bias of $\hat{X}_{100}$.

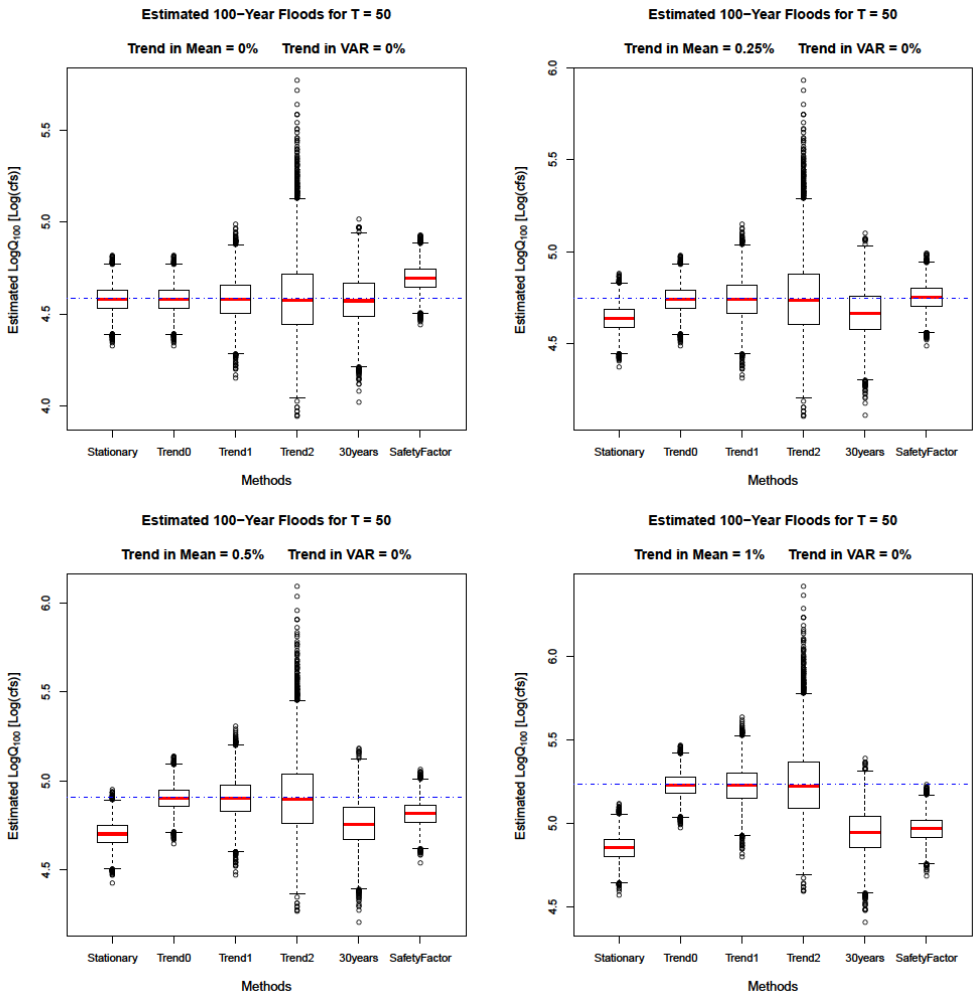


Figure 3.B-1: Boxplots of estimated Q_{100} for T = 50 (Case 1–Positive trends in μ_x)

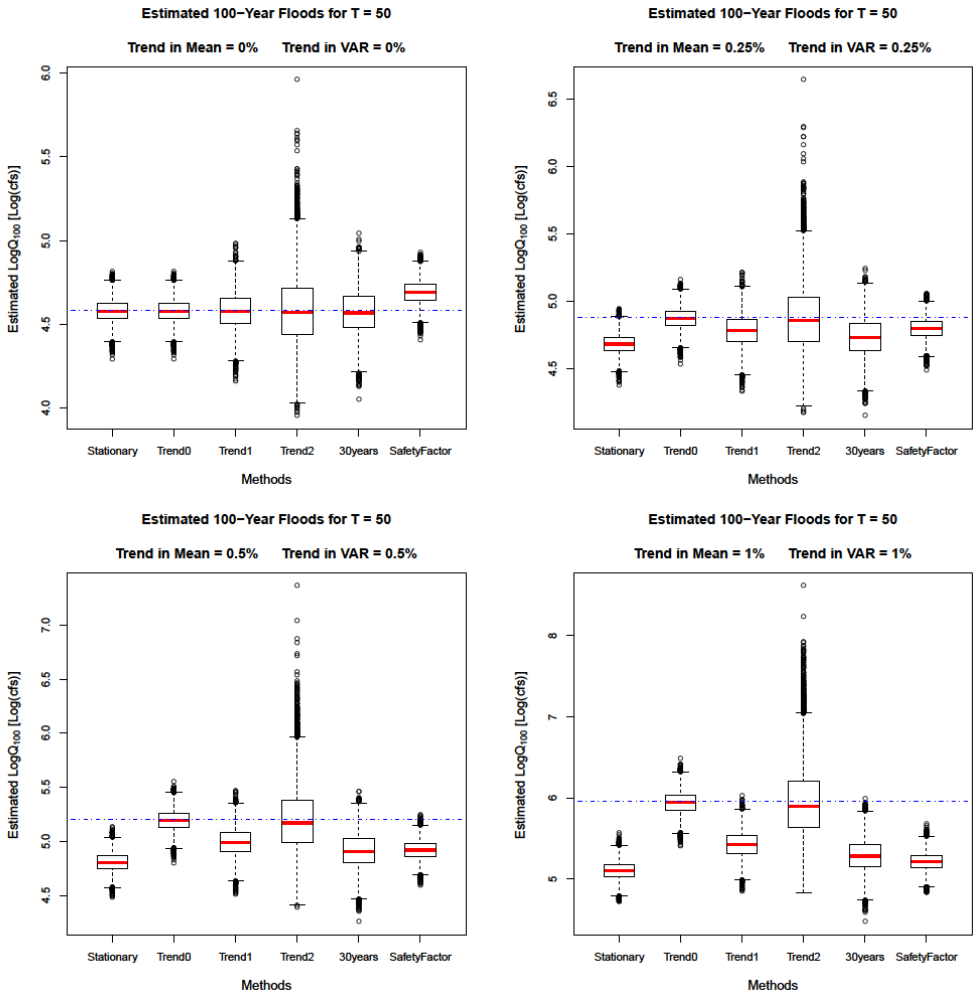


Figure 3.B-2: Boxplots of estimated Q_{100} for T = 50 (Case 2–Positive trends in μ_x and σ_t^2)

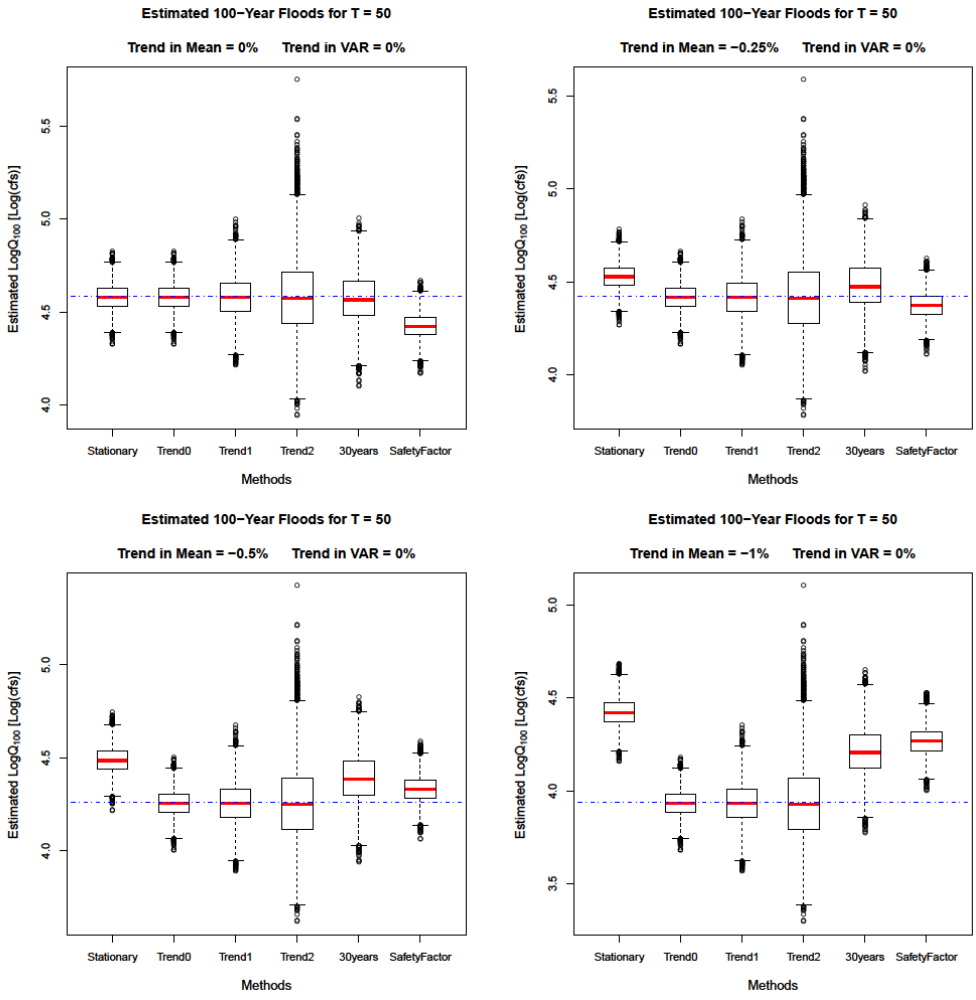
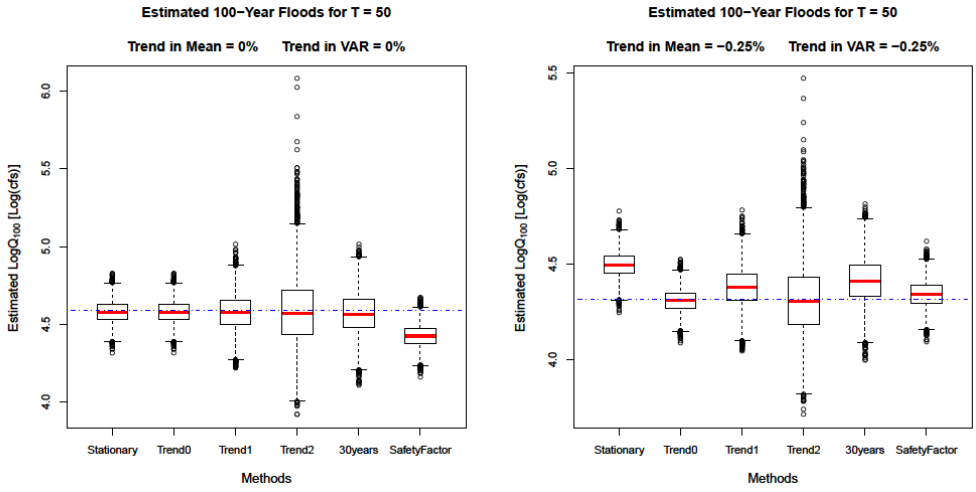


Figure 3.B-3: Boxplots of estimated Q_{100} for T = 50 (Case 3–Negative trends in μ_x)



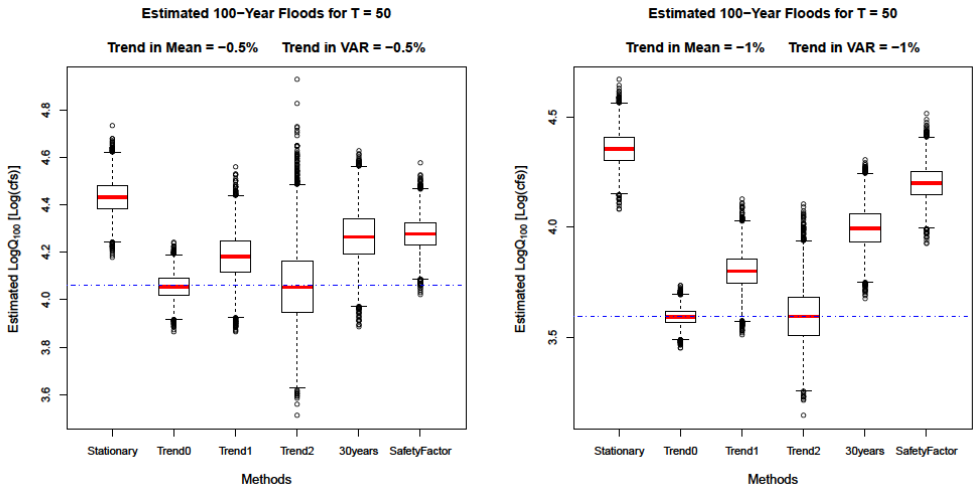


Figure 3.B-4: Boxplots of estimated Q_{100} for T = 50 (Case 4–Negative trends in μ_x and σ_t^2)

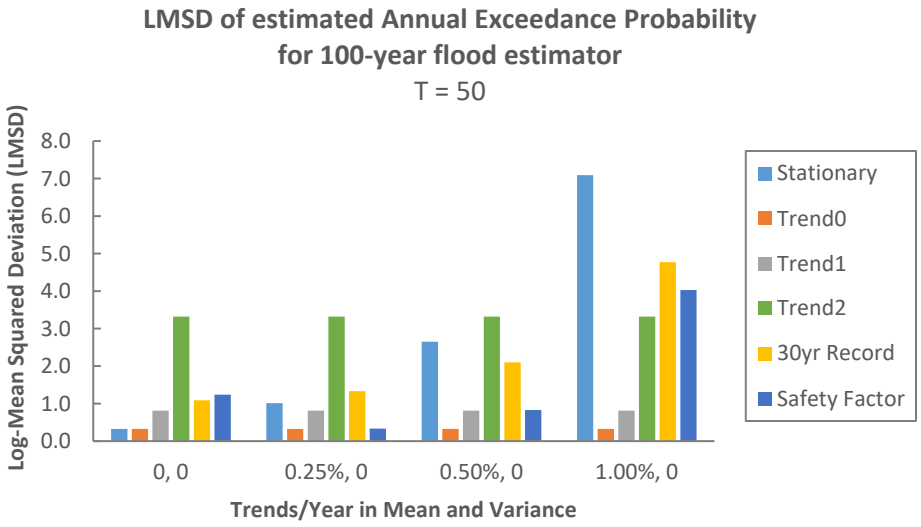


Figure 3.B-5: LMSD of Annual Exceedance Probability for 1% Event

(Case 1-Positive trends in μ_x)

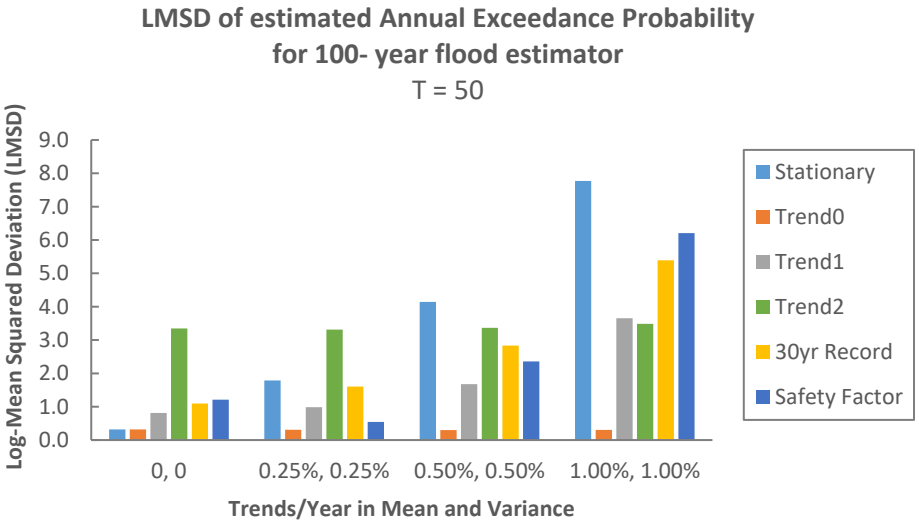


Figure 3.B-6: LMSD of Annual Exceedance Probability for 1% Event

(Case 2-Positive trends in μ_x and σ_t^2)

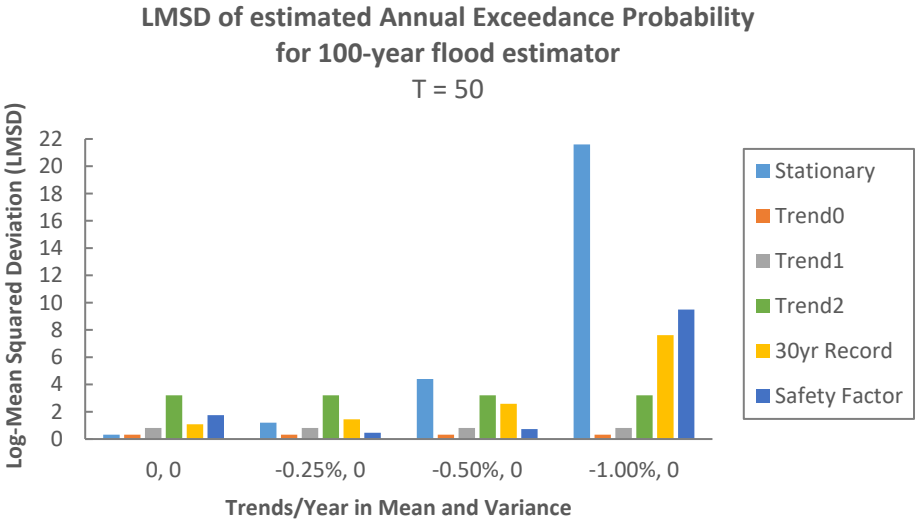


Figure 3.B-7: LMSD of Annual Exceedance Probability for 1% Event

(Case 3-Negative trends in μ_x)

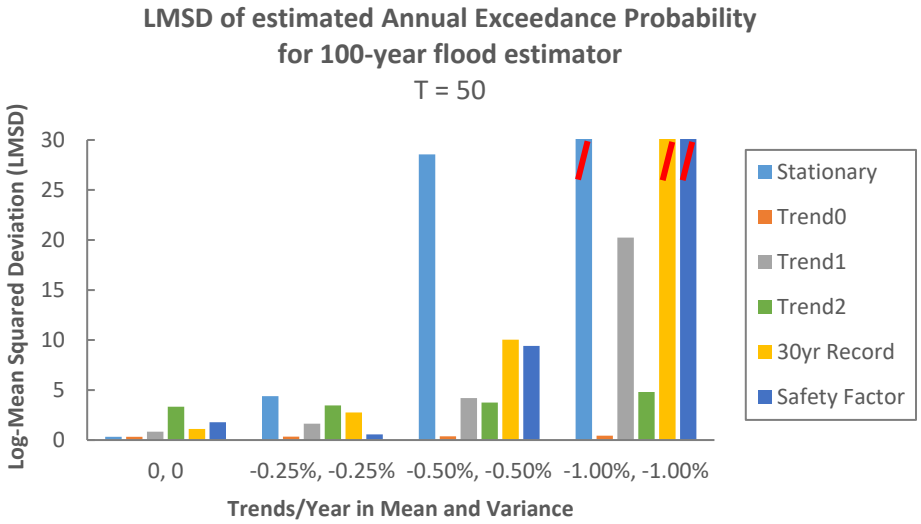


Figure 3.B-8: LMSD of Annual Exceedance Probability for 1% Event

(Case 4-Positive trends in both μ_x and σ_t^2)

*For $\beta_1 = \beta_2 = -1.00\%$, LMSDs of "Stationary", "30yr Record", and "Safety Factor" are larger than 30

The Monte Carlo results considering 25-year projection and 50-year projection yield similar conclusions as summarized in Section 6. "Trend_1" estimator is recommended except the extreme scenario ($\beta_1 = \beta_2 = \pm 1\%$); "Trend_2" estimator is a reasonable alternative when $\beta_1 = \beta_2 = \pm 1\%$.

Table 3.B-1: Monte Carlo Study Results Case 1 (Positive trends in μ_x), T=50

Estimators	**Mean** $\tilde{AEP}_{0.01}$	**LMSD(** $\tilde{AEP}_{0.01}$**)**	**LMSE(** $\hat{Q}_{100}$**)**	**Bias²(** $\hat{X}_{100}$**)**
Trend in mean = 0, Trend in variance = 0				
Stationary	1.22%	0.3258	0.0262	0.0002
Trend_0	1.22%	0.3258	0.0262	0.0002
Trend_1	1.49%	0.8121	0.0653	0.0003
Trend_2	2.53%	3.3210	0.2439	0.0000*
30yr Record	1.72%	1.0882	0.0887	0.0007
Safety Factor	0.48%	1.2366	0.0870	0.0610
Trend in mean = 0.25%, Trend in variance = 0				
Stationary	2.70%	1.0104	0.0916	0.0653
Trend_0	1.22%	0.3258	0.0262	0.0002
Trend_1	1.49%	0.8121	0.0653	0.0003
Trend_2	2.53%	3.3210	0.2439	0.0000*
30yr Record	2.83%	1.3308	0.1235	0.0353
Safety Factor	1.13%	0.3350	0.0264	0.0000
Trend in mean = 0.50%, Trend in variance = 0				
Stationary	5.31%	2.6501	0.2572	0.2298
Trend_0	1.22%	0.3258	0.0262	0.0002
Trend_1	1.49%	0.8121	0.0653	0.0003
Trend_2	2.53%	3.3210	0.2439	0.0000*
30yr Record	4.49%	2.0995	0.2090	0.1205
Safety Factor	2.40%	0.8291	0.0745	0.0471
Trend in mean = 1.00%, Trend in variance = 0				
Stationary	15.1%	7.0891	0.8041	0.7734
Trend_0	1.22%	0.3258	0.0262	0.0002
Trend_1	1.49%	0.8121	0.0653	0.0003
Trend_2	2.53%	3.3210	0.2439	0.0000*
30yr Record	10.2%	4.7708	0.5268	0.4368
Safety Factor	7.85%	4.0266	0.4115	0.3808

* Values are extremely small, but not exactly zero

Table 3.B-2: Monte Carlo Study Results Case 2 (Positive trends in μ_x and σ_t^2), T=50

Estimators	**Mean** $\tilde{AEP}_{0.01}$	**LMSD(** $\tilde{AEP}_{0.01}$**)**	**LMSE(** $\hat{Q}_{100}$**)**	**Bias²(** $\hat{X}_{100}$**)**
Trend in mean = 0, Trend in variance = 0				
Stationary	1.23%	0.3178	0.0256	0.0003
Trend_0	1.23%	0.3178	0.0256	0.0003
Trend_1	1.51%	0.8125	0.0657	0.0004
Trend_2	2.62%	3.3475	0.2477	0.0000
30yr Record	1.76%	1.0974	0.0899	0.0011
Safety Factor	0.48%	1.2090	0.0852	0.0599
Trend in mean = 0.25%, Trend in variance = 0				
Stationary	3.89%	1.7842	0.2414	0.2108
Trend_0	1.22%	0.3067	0.0359	0.0004
Trend_1	2.48%	0.9834	0.1304	0.0531
Trend_2	2.52%	3.3142	0.3505	0.0000*
30yr Record	3.54%	1.6049	0.2240	0.1137
Safety Factor	1.96%	0.5449	0.0693	0.0387
Trend in mean = 0.50%, Trend in variance = 0				
Stationary	7.93%	4.1442	0.8854	0.8452
Trend_0	1.21%	0.3008	0.0510	0.0006
Trend_1	3.76%	1.6782	0.3361	0.2436
Trend_2	2.46%	3.3671	0.5089	0.0000*
30yr Record	5.87%	2.8329	0.6040	0.4659
Safety Factor	4.77%	2.3560	0.4718	0.4316
Trend in mean = 1.00%, Trend in variance = 0				
Stationary	16.7%	7.7720	3.9541	3.8813
Trend_0	1.20%	0.3056	0.1087	0.0009
Trend_1	7.14%	3.6531	1.6359	1.4989
Trend_2	2.44%	3.4851	1.0963	0.0001
30yr Record	11.1%	5.3924	2.6054	2.3857
Safety Factor	12.4%	6.2039	2.9891	2.9163

* Values are extremely small, but not exactly zero

Table 3.B-3: Monte Carlo Study Results Case 3 (Negative trends in μ_x), T=50

Estimators	**Mean** $\tilde{AEP}_{0.01}$	**LMSD(** $\tilde{AEP}_{0.01}$**)**	**LMSE(** $\hat{Q}_{100}$**)**	**Bias²(** $\hat{X}_{100}$**)**
Trend in mean = 0, Trend in variance = 0				
Stationary	1.22%	0.3245	0.0260	0.0002
Trend_0	1.22%	0.3245	0.0260	0.0002
Trend_1	1.49%	0.8167	0.0656	0.0003
Trend_2	2.55%	3.2086	0.2390	0.0000*
30yr Record	1.76%	1.0823	0.0888	0.0012
Safety Factor	3.84%	1.7524	0.1640	0.1382
Trend in mean = 0.25%, Trend in variance = 0				
Stationary	0.49%	1.2030	0.0846	0.0585
Trend_0	1.22%	0.3245	0.0260	0.0002
Trend_1	1.49%	0.8167	0.0656	0.0003
Trend_2	2.55%	3.2086	0.2390	0.0000*
30yr Record	1.03%	1.4511	0.1038	0.0162
Safety Factor	1.71%	0.4555	0.0393	0.0132
Trend in mean = 0.50%, Trend in variance = 0				
Stationary	0.17%	4.4071	0.2926	0.2655
Trend_0	1.22%	0.3245	0.0260	0.0002
Trend_1	1.49%	0.8167	0.0656	0.0003
Trend_2	2.55%	3.2086	0.2390	0.0000*
30yr Record	0.59%	2.5791	0.1725	0.0846
Safety Factor	0.66%	0.7297	0.0523	0.0251
Trend in mean = 1.00%, Trend in variance = 0				
Stationary	0.01%	21.601	1.2616	1.2312
Trend_0	1.22%	0.3245	0.0260	0.0002
Trend_1	1.49%	0.8167	0.0656	0.0003
Trend_2	2.55%	3.2086	0.2390	0.0000*
30yr Record	0.17%	7.6149	0.4766	0.3874
Safety Factor	0.07%	9.4902	0.5973	0.5669

* Values are extremely small, but not exactly zero

Table 3.B-4: Monte Carlo Study Results Case 4 (Negative trends in μ_x and σ_t^2), T=50

Estimators	**Mean** $\tilde{AEP}_{0.01}$	**LMSD(** $\tilde{AEP}_{0.01}$**)**	**LMSE(** $\hat{Q}_{100}$**)**	**Bias²(** $\hat{X}_{100}$**)**
Trend in mean = 0, Trend in variance = 0				
Stationary	1.23%	0.3259	0.0263	0.0003
Trend_0	1.23%	0.3259	0.0263	0.0003
Trend_1	1.53%	0.8323	0.0673	0.0005
Trend_2	2.68%	3.3298	0.2477	0.0001
30yr Record	1.81%	1.1158	0.0918	0.0017
Safety Factor	3.89%	1.7779	0.1666	0.1407
Trend in mean = 0.25%, Trend in variance = 0				
Stationary	0.19%	4.3925	0.2001	0.1755
Trend_0	1.24%	0.3427	0.0190	0.0002
Trend_1	0.86%	1.6396	0.0787	0.0215
Trend_2	2.84%	3.4675	0.1805	0.0000*
30yr Record	0.68%	2.7587	0.1267	0.0532
Safety Factor	0.94%	0.5565	0.0285	0.0039
Trend in mean = 0.50%, Trend in variance = 0				
Stationary	0.01%	28.565	0.7648	0.7388
Trend_0	1.25%	0.3653	0.0140	0.0002
Trend_1	0.43%	4.1912	0.1286	0.0791
Trend_2	3.06%	3.7416	0.1356	0.0000*
30yr Record	0.17%	10.029	0.2892	0.2284
Safety Factor	0.09%	9.4151	0.2789	0.2529
Trend in mean = 1.00%, Trend in variance = 0				
Stationary	0.00%*	350.27	3.1016	3.0704
Trend_0	1.28%	0.4303	0.0078	0.0001
Trend_1	0.07%	20.246	0.2600	0.2214
Trend_2	3.75%	4.7960	0.0828	0.0000*
30yr Record	0.00%*	81.405	0.8960	0.8520
Safety Factor	0.00%*	204.75	1.9788	1.9476

* Values are extremely small, but not exactly zero

Appendix 3.C – *Estimators in Monte Carlo Study*

Chapter 3 introduced several estimators that can be used when a record has "trend" and one wishes to describe flood risk at some time in the future. This Appendix will describe in detail the estimators of the parameters developed to implement each of the methods considered in the Monte Carlo study.

The logarithms of the available record for flood frequency analysis are denote by X_t, t = 0, 1, 2, …, N-1. In the MC study, N = 100. These values are obtained by re-sampling the gauged flow record available at Sinnemahoning Creek, Sterling Run, PA (gage number: 01543000), and then adding a trend appropriate for each case. Here a trend in the mean or variance understood to be described by the equations (see Section 2):

$$\mu_t = \mu_0 + \beta_1 t \qquad \text{(3.C-1)}$$

$$\sigma_t^2 = \sigma_0^2 \exp\left(\beta_2 t\right) \qquad \text{(3.C-2)}$$

μ_0 is the mean value at t = 0, and σ_0^2 is the variance at t = 0.

1. Stationarity

When assuming the annual maximum series is stationary, the 100-year flood estimator is calculated by employing the sample mean, variance, and coefficient of skewness of X_t to fit a LP3 distribution.

$$\bar{X} = \frac{1}{N}\sum_{t=0}^{N-1} X_t \qquad \text{(3.C-3)}$$

$$S_X^2 = \frac{1}{N-1}\sum_{t=0}^{N-1}\left(X_t - \bar{X}\right)^2 \qquad \text{(3.C-4)}$$

$$G_X = \frac{N}{\left(N-1\right)\left(N-2\right)}\sum_{t=0}^{N-1}\left(X_t - \bar{X}\right)^3 \qquad \text{(3.C-5)}$$

These can be used in equation 3.C-6 to obtain an estimator of the quantile.

$$X_t = \hat{\mu}_t + \hat{\sigma}_t K\left(\hat{\gamma}\right) \qquad \text{(3.C-6)}$$

2. Trend_0

"Trend_0" assumes the slopes β_1 and β_2 in equations (3.C-1) and (3.C-2) are known, corresponding to the trend in the log-space mean and variance. The concern is flood risk some T years beyond the end of the record, which is itself N years in length. In order to estimate the future mean $\hat{\mu}_{N+T-1}$ and variance $\hat{\sigma}^2_{N+T-1}$ for T=25 or 50, the initial mean μ_0 and variance σ_0^2 (at t = 0) should be estimated. As mentioned in Section 3, Eqn.3.C-1 and 3.C-2 are modified to include an error term to serve as the basis of the estimator of trends in the log-space mean and variance to become

$$X_t = \alpha_1 + \beta_1 t + \varepsilon_{1,t} \qquad \text{(3.C-7)}$$

$$\left(X_t - \hat{\mu}_t\right)^2 = \alpha_2 \exp\left(\beta_2 t\right) + \varepsilon_{2,t} \qquad \text{(3.C-8)}$$

Here α_1 and α_2 are the same as μ_0 and σ_0^2 in Eqn.3.C-1 and 3.C-2. The estimator α_1 is computed first so that $\hat{\mu}_t$ is available for computing α_2 based upon Eqn.3.C-8 which requires $\hat{\mu}_t$.

Equation 3.C-8 is based on the assumption expressed in 3.C-2 that $E[(X_t - \hat{\mu}_t)^2] = \sigma_t^2$ which should equal $\sigma_0^2 \exp\left(\beta_2 t\right)$. Unfortunately, $E[(X_t - \hat{\mu}_t)^2]$ is a very imprecise estimator of the variance for each t that is based upon only a single observation. Fortunately, using Eqn.3.C-8, all of the observations are employed to estimate α_2. If $\beta_2 = 0$, use of Eqn.3.C-8 will essentially yield the traditional sample variance estimator.

For "Trend_0", the slopes β_1 and β_2 are known. Then α_1 and α_2 corresponding to μ_0 and σ_0^2, are estimated as:

$$\hat{\alpha}_1 = \frac{1}{N}\sum_{t=0}^{N-1}\left(X_t - \beta_1 t\right) \qquad \text{(3.C-9)}$$

$$\hat{\alpha}_2 = \frac{1}{N-1}\sum_{t=0}^{N-1}\left[\left(X_t - \hat{\mu}_t\right)^2 \exp\left(-\beta_2 t\right)\right] \qquad \text{(3.C-10)}$$

Equation 3.C-9 is the classical least-squares estimator, corresponding to the MLE for normal errors. Equation 3.C-10 is the classical least-squares estimator, if the

equation 3.C-8 is reformulated by dividing by $exp(\beta_2 t)$ for each t. Such a reformulation is likely to make the errors homoscedastic given that one would expect the variance of the errors in Eqn.3.C-8 to either increase or decrease with changes in the variance σ_0^2.

For "Trend_0" the mean and variance in year t are estimated as:

$$\hat{\mu}_t = \hat{\alpha}_1 + \beta_1 t \qquad (3.C\text{-}11)$$

$$\hat{\sigma}_t^2 = \hat{\alpha}_2 \exp(\beta_2 t) \qquad (3.C\text{-}12)$$

For a 25-year or 50-year projection, the mean and variance are calculated using Eqn.3.C-11 and 3.C-12 with t = N+T-1.

The coefficient of skewness used for computing the 100-year flood estimator is the skew of the de-trended X_t series, which is the skew of $(X_t - \hat{\mu}_t)/\hat{\sigma}_t$, because we assume there is no trend in the skewness coefficient.

3. Trend_1

"Trend_1" allows for a trend in the mean but not the variance. Standard least-squares estimators are employed for the initial mean value α_1 and the log-space trend β_1; see Eqn.3.C-7. The mean value for any t is estimated as:

$$\hat{\mu}_t' = \hat{\alpha}_1 + \hat{\beta}_1 t \quad (3.C\text{-}13)$$

The mean value for a 25-year or 50-year projection is calculated using Eqn.3.C-13 with t = T+N-1.

"Trend_1" assumes the variance is a constant. The log-space variance estimator is the residual variance for the regression for Eqn.3.C-7:

$$s_{\varepsilon_1}^2 = \frac{1}{N-2}\sum_{t=0}^{N-1}\left(X_t - \hat{\mu}_t'\right)^2 \qquad (3.C\text{-}14)$$

The skew coefficient of "Trend_1" estimator is computed using $(X_t - \hat{\mu}_t')/s_{\varepsilon_1}$, t=0, …, N-1.

4. Trend_2

"Trend_2" considers a trend in both the mean and the variance. The mean value for the 25-year or 50-year projection is the same as the mean in Trend_1. Trend_2 needs additional estimators of α_2 and β_2. The estimator in equation 3.C-10 employs the estimator of β_2 to compute the estimator α_2 using weights $exp(-\hat{\beta}_2 t)$; when β_2 is estimated, the use of an estimated parameter to define the weights can cause problems.

Referring back to Eqn.3.C-2. The trend in the log-space variance is estimated directly with nonlinear least squares based on the equation:

$$\left(X_r - \hat{\mu}_r^{'}\right)^2 = \alpha_2 \frac{r\exp\left(0.5\beta_2 r\right) + \exp\left(-0.5\beta_2 r\right)}{\exp\left(0.5\beta_2 r\right) + r\exp\left(-0.5\beta_2 r\right)} + \varepsilon_{2,r} \quad r = t+1 = 1, 2, \ldots, N \quad (3.C\text{-}15)$$

The variance for a 25-year or 50-year projection is calculated using Eqn. 3.C-15 with r = N+T.

The skew of the de-trended X_t is the skew of $(X_t - \hat{\mu}_t')/\hat{\sigma}_t'$. Here $\hat{\mu}_t'$ and $\hat{\sigma}_t'$ are estimated from Eqn.3.C-7 and 3.C-15 with t = 0, …, N-1 (r = 1, …, N).

5. 30 Years Record

This method also assumes stationarity but only uses the record from the most recent 30 years: $X' = \{X_{N-30}, X_{N-29}, \ldots, X_{N-1}\}$. The mean, variance, and skew are calculated using that 30-year record.

6. Safety Factor

The mean, variance, and skew are exactly the same as the moments in "Stationarity" method. The only difference is that a factor is added to the real-space 100-year flood estimator as described in the chapter.

CHAPTER 4

FLOOD FREQUENCY ANALYSIS IN THE CONTEXT OF CLIMATE CHANGE II: APPLICATION OF LOG-PEARSON TYPE 3 DISTRIBUTION

Abstract

Classical flood frequency analysis considers fitting a probability distribution with fixed parameters to the annual maximum flood series. This assumes the watershed characteristics and meteorological statistics are constant over the period of record and on to the year for which a flood risk estimate is needed; essentially the assumption is that the annual maximum series is stationary. Recently, hydrologists have been concerned about the impacts of climate change, which may make the annual maximum series non-stationary. A common proposal is to estimate the trends in the mean (sometimes also the variance) of floods over the historical period and employ the estimated trends to project the flood-risk distribution's parameters into the future. This paper evaluates via a Monte Carlo study with the log-Pearson type 3 (LP3) distribution several methods for such dynamic flood frequency analysis.

With a sample size of 40 and small trends (within ±0.25% per year), a stationarity estimator is about the best among realistic methods across all cases considered regardless of the skewness coefficient; for larger trends one should use the time-varying LP3 parameters that incorporate the trend in the log-mean (Trend_1). With the sample size of 100, Trend_1 generally does about as well as can be done with reasonable trends; only for ±1% per year or more-extreme-values in both the mean and variance was it advantageous to estimate the trends in both the log-mean and log-variance of the annual maximum series (Trend_2). Clearly, with the larger sample size, Trend_1 is the robust choice if trends are anticipated. The results show that the

LP3 coefficient of skewness has considerable effects on the precision of quantile and exceedance probability estimators in a dynamic world.

Table of Contents

1. Introduction

Flood frequency analysis usually fits a probability distribution to peak-flow records that consist of the annual maximum flows. Hydrologists estimate the risk of extreme floods (e.g. 100-year flood) by extrapolating with the fitted distribution to get quantiles with the target exceedance probabilities. However, when long-term climate change becomes a concern, the time invariance (i.e. stationary) assumption of the method may not be valid. Climate change can result in changes in the mean and/or the variability of floods that persists for an extended period, decades or longer [IPCC, 2012, p.29]. Researchers have explored if climate change will significantly affect the water cycle, and affect the distributions of flood peaks [Lins and Slack, 1999; Jain and Lall, 2001; Koutsoyiannis and Montanari, 2007; Hirsch and Ryberg, 2012; Mentaschi, et al. 2016; Hirsch and Archfield, 2016]. Despite the complex and controversial answers to that question, the national Guidance for Determining Flood Flow Frequency (Bulletin 17B) [IACWD, 1982] in United States and its upcoming update (Bulletin 17C, 2017 Draft) both keep the time invariance assumption; however, Bulletin 17C encourages hydrologists to incorporate climate change when there is sufficient scientific evidence to support quantification of its impacts.

The log-Pearson type III (LP3) distribution is recommended in Bulletin 17B. Its characteristics, parameters, and application have been discussed in Bobée [1975], Bobée and Robitaille [1977], Bobée and Ashkar [1991], Vogel, et al., [1993], Griffis et al. [2004], and Griffis and Stedinger [2007ab, 2009].

The probability density function of a LP3 distribution can be written:

$$f_Q(q) = \frac{1}{q|\beta|\Gamma(\alpha)}\left[\frac{\ln(q)-\tau}{\beta}\right]^{\alpha-1}\exp\left[-\frac{\ln(q)-\tau}{\beta}\right] \qquad (1)$$

where Q is the annual maximum flood; α, β, and τ, are natural parameters for the distribution corresponding to shape and scale parameters for the $\ln(q)$ distribution,

whereas τ is a lower bound for ln(q). Alternatively, the distribution can be defined in terms of the mean, standard deviation, and skewness coefficient of either Q or X = ln(Q). Here we denote the logarithm of the annual maximum series as X and work with the moments of X.

The parameters discussed in this chapter include the log-space mean

$$\mu_X = \tau + \alpha \qquad (2a)$$

that determines the log-space location; the log-space standard deviation

$$\sigma_X^2 = \alpha\beta^2 \qquad (2b)$$

that determines the log-space scale; and the log-space coefficient of skewness

$$\gamma_X = 2\ sign(\beta)/\sqrt{\alpha} \qquad (2c)$$

that determines the shape of the probability density function for X. These are the parameters used in many flood frequency investigations, particularly in the United States [IACWD, 1982; Bobee and Ashkar, 1991; Griffis and Stedinger, 2007a]. In that spirit, we write that (neglecting the subscripts):

$$\ln(Q) = X \sim P3[\mu, \sigma, \gamma] \qquad (3a)$$

Considering possible climate change impacts, a commonly adopted model for X employs time-varying distribution parameters in the flood frequency analysis [Coles, 2001, pp.105-108; El Adlouni et al., 2007]. In this case, Eqn. 3a becomes:

$$\ln(Q) = X \sim P3[\mu(t), \sigma(t), \gamma] \qquad (3b)$$

reflecting the possible change in the mean and the standard deviation.

This paper describes how dynamic descriptions of μ and σ might be developed, and the impact of estimation of time trends in those LP3 parameters on the error in estimated quantiles. Procedures of choosing the regional coefficient of skewness and calculating the at-site coefficient of skewness are introduced in Section 3. Section 4 covers the cases and scenarios included in the Monte Carlo study. Section 5 discusses the details of the estimators employed in the simulation. Section 6

provides an analysis and summary of the results. At last, Section 7 provides the conclusion and recommendations.

2. Consideration of Climate Change Impacts

A simple linear model using time as a predictor provides an estimate of the trend in the mean (or/and variance) of the annual maximum series. The trend describes how the mean (or/and variance) change with time (e.g. percent change per year). The performance of flood risk estimators that attempt to correctly represent such a simple linear trend should do as well or better than would estimators for more complex situations.

Equations of the trend model were discussed in Chapter 3; the two models to be fit to the data to describe trends in the mean and variance of X_t are:

$$X_t = \alpha_1 + \beta_1 t + \varepsilon_{1,t} \qquad (4)$$

$$\left(X_t - \hat{\mu}_t\right)^2 = \alpha_2 \exp\left(\beta_2 t\right) + \varepsilon_{2,t} \qquad (5)$$

Here α_i and β_i with subscripts correspond to the trend model for the mean (i=1) or the variance (i=2); whereas α and β without subscripts refer to the LP3 parameters in Eqn. 1; X_t is the logarithm of a recorded flow at the gauged station; β_1 and β_2 describe the trends in the mean and variance of X_t; $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are random errors; $\hat{\mu}_t$ equals $\hat{a}_1 + \hat{b}_1 t$, and is the estimated expected value of X_t in year t. The value of β_1 in Eqn. 4 was estimated using ordinary least squares (OLS) regression, which works well if the errors $\varepsilon_{1,t}$ are normal, independent and identically distributed [Greene, 2012, p.52]. In our case for modest log-space skews, OLS should work well for the mean regression. The value of β_2 for Eqn. 5 was estimated using nonlinear least-absolute-value regression (as described in Section 5), because the residuals for Eqn. 5 would be very non-normal.

In addition, the study also considers 3 methods that do not include dynamic parameters; they are: (a) keep the assumption of stationarity and use all of annual maximum series; (b) keep the assumption of stationarity but use only the most recent 30 years of data; (c) add a safety factor to the estimated quantile to reflect climate change impacts. The third method has been recommend for flood risk estimation in Europe [Maden et al., 2013]. The third method incorporates an inconsistency: if parameters are changing over time, then in estimating the model's parameters such trends should be taken into account in the estimation of the parameters.

A Monte Carlo study evaluated these methods in Chapter 3 based on the re-sampling of a real record for Sinnemahoning Creek, Sterling Run, PA (HCDN-2009, USGS gage number: 01543000), which ran from 1914 to 2016, exactly 103 years. To further explore the performance of different methods under the assumption of non-stationarity, this chapter generates random samples from LP3 distributions with selected coefficients of skewness and two systematic record lengths. In addition, regional skew information (regional skew and its standard error) is employed in the simulations to represent what would be done in practice with Bulletin 17B or 17C.

3. Parameter Selection

To properly describe the annual maximum series, there are reasonable ranges for the log-space skew of a LP3 distribution. This range is generally dominated by the value of real-space skew γ_Q, and restricted by the upper bound on the coefficient of variation (Cv) and the scale parameter β (β > −1) in the probability density function. A realistic range for the log-space skew γ_X for flood series around the world would be (-1,1), depending on the value of σ_X [Griffis and Stedinger, 2007a, Fig. 10; also Landwehr et al. 1978; Weaver, et al., 2009, p.110, Fig. A2].

Our investigation, following Griffis, Stedinger, and Cohn [2004], considers LP3 coefficient of skewness of –1, –0.5, 0.2, 0, -0.2, +0.5, and +1 as the value of the log-space regional skew G. Following Griffis and Stedinger [2009] (also see Chowhury and Stedinger, 1991), we generate the at-site skew randomly based on the precision of regional skew estimator, whose estimation error is believed to be on the order of 0.1 [Tasker and Stedinger, 1986; Gruber and Stedinger, 2008; Parrett, et al. 2010, p.89, Fig. B1; Veilleux et al., 2011].

The approach employed to generate the population skew is:

1. If the log-space regional skew is negative, generate a random at-site skew using a Pearson type 3 (P3) distribution with lower bound -1.4, variance 0.1, and mean G.

2. If the log-space regional skew is positive, generate a random at-site skew using a P3 distribution with upper bound of 1.4, variance 0.1, and mean G.

3. If the log-space regional skew is zero, generate a random at-site skew using a normal distribution with variance 0.1 and mean 0.

If the generated at-site skew is smaller than -1.4 or larger than 1.4, that values is dropped and another generated.

4. To improve the efficiency of quantile estimation, Bulletin 17B suggests combining the regional skew information with the at-site coefficient of skewness estimator. In our Monte Carlo study, the log-space skew for each annual maximum series is combined with a regional skew G to get a weighted skewness estimator $\tilde{G}$ for the site using

$$\tilde{G} = \frac{\hat{\gamma}\left(MSE_G\right) + G\left(MSE_{\hat{\gamma}}\right)}{MSE_G + MSE_{\hat{\gamma}}} \qquad (6)$$

$\tilde{G}$ is the weighted skew used by estimators with each sample in the Monte Carlo study. The mean squared error of the regional skew estimator MSE_G is equal to

the variance of log-space regional skew 0.1. Griffis and Stedinger [2009] derived equations for estimating $MSE_{\hat{\gamma}}$. The equations are listed in Appendix A (Eqn. A1-A4).

5. If the weighted skew $\tilde{G} > 1.4$, set the value to 1.4. If the weighted skew $\tilde{G} < -1.4$, set the value to –1.4.

For cases with time varying parameters, the distribution mean in year t was generated using $\mu_t = \mu_0 + \beta_1 t$, t = 0, …, T-1; the variance was generated using $\ln\left(\sigma_t^2\right) = \ln\left(\sigma_0^2\right) + \beta_2 t$ or equivalently $\sigma_t^2 = \sigma_0^2 \exp\left(\beta_2 t\right)$. This corresponds to Eqn. 5 without the errors introduced when one substitutes $(X_t - \hat{\mu}_t)^2$ for the variance σ_t^2.

4. Monte Carlo Study

This Monte Carlo study will evaluate four methods:

(1) Use up-to-date records assuming stationarity

(2) Use annual maximum series with estimated time-varying flood distribution parameters.

(3) Use the most recent 30 years of the record.

(4) Use European safety factors with (1), the stationary method.

Based on the analysis of trends in 473 Sites in Chapter 3, this study includes trend magnitudes of 0, ±0.25%, ±0.5%, and ±1%, as the trends in the mean, or in both the mean and variance. The simulation considers 5 different cases:

Case 1: Positive trends in log-mean (μ_x)

Case 2: Equal positive trends in log-mean (μ_x) and log-variance (σ_X^2)

Case 3: Negative trends in log-mean (μ_x)

Case 4: Equal negative trends in log-mean (μ_x) and log-variance (σ_X^2)

Case 5: Positive trends in log-mean (μ_x) and a negative trend in log-variance (σ_X^2) whose magnitude is half the trend in the mean.

Case 1 and 3 have trend only in the mean. Case 2 and Case 4 have trends in the log-space mean and variance, which have the same rate of increase (or decrease). The detailed analysis of the trend selection was discussed in Chapter 3, section 3. Equal trend magnitudes were selected with a vision that floods were getting larger or smaller in every respect, as opposed to cases 1 and 3 where the log-space variance was constant, corresponding to a fixed real-space coefficient of variation. Case 5 reflects a vision that as an area becomes wetter with larger floods, typically floods have a smaller coefficient of variation: thus the log-space variance might decline, and vice versa. Observed positive-negative mean and variance trend pairings among 473 HCDN-2009 stations (record lengths range from 50 to over 110 years) had 45 (-,-), 128 (+,+), 107(+,-) and 93 (-,+) pairings (see Table 1 below), though most trend parameters were not statistically significant. Case 5 reflects a positive mean trend and a negative variance trend, which was the third most frequent pairing.

Table 1. Number of sites corresponding to the configuration of trends in the mean and the variance from 473 HCDN-2009 (Hydro-Climatic Data Network) stations

	Positive trends in the mean	**Negative trends** in the mean
Positive trends in the variance	128	93
Negative trends in the variance	107	145

To compare 100-year flood estimators $\hat{Q}_{100,t}$, the Monte Carlo study computed the mean squared error (MSE) of the logarithm of $\hat{Q}_{100,t}$ (i.e. the MSE of $\hat{X}_{100,t}$). Here t equals the length of the record plus the number of years we project the risk (25 years or 50 years) in the future; the logarithm mean square error is calculated as:

$$LMSE(\hat{Q}_{100,t}) = \frac{1}{N}\sum_{i=1}^{N}\left[\ln\left(\hat{Q}_{100,t}\right)_i - \ln\left(Q_{100,t}\right)\right]^2 / \sigma_{X,t}^2 \qquad (7)$$

In Eqn. 7, $Q_{100,t}$ is the true value of the 100-year flood in year t. By dividing by the log-space variance in year t, the mean squared error of X here is dimensionless. Thus, the value of the log-space mean and standard deviation have no impact on the LMSE. If a trend existed in the parameters, then σ_X^2 corresponds to the variance in the year for which a flood quantile estimator is needed. In this simulation, we choose $\mu_X = 9$, and $\sigma_X = 0.5$ (ln-cfs) as the natural log-space mean and standard deviation, corresponding approximately to Sinnemahoning Creek, Sterling Run, PA (USGS gage number: 01543000).

Furthermore, we evaluate the MSE associated with the annual exceedance probability, $A\hat{E}P_{0.01,t}$, which is defined as:

$$A\hat{E}P_{0.01,t} = 1 - F_Q\left(Q_{100,t} \mid \hat{\mu}_t, \hat{\sigma}_t, \tilde{\mathrm{G}}\right) \qquad (8)$$

where F_Q is the cumulative probability function of the fitted LP3 distribution. $\hat{\mu}_t$, $\hat{\sigma}_t$, and $\tilde{G}$ are the estimated location, scale, and shape parameters. Our MSE for the logarithm of $A\hat{E}P_{0.01,t}$ is calculated as:

$$LMSE(A\hat{E}P_{0.01,t}) = \frac{1}{N}\sum_{i=1}^{N}\left[\ln\left(A\hat{E}P_{0.01,t} + 10^{-4}\right)_i - \ln\left(1.01\times10^{-2}\right)\right]^2 \qquad (9)$$

here 1% is the exceedance probability for $Q_{100,t}$. In Eqn. 9, the small value 10^{-4} is added to both terms. It avoids the problem of AEP = 0 when the estimated upper bound is less than Q_{100}. Also, it is a more robust metric that gives limited weight to estimated $A\hat{E}P_{0.01,t} < 10^{-4}$. Once an AEP estimator for the 100-year events is less than 10^{-4}, practically it does not matter if the estimate is really 10^{-5}, 10^{-6}, or 10^{-8}. Relative to the target of 10^{-2}, $A\hat{E}P_{0.01,t} < 10^{-4}$ are essentially zero.

Our Monte Carlo study also explores the impacts of different sample sizes on the precision of $\hat{Q}_{100,t}$ and $A\hat{E}P_{0.01,t}$. Results include samples of 40 and 100 years.

5. Estimators

The Monte Carlo study considers six estimators. Method 1 corresponds to the method of log-space moments with the LP3 distribution assuming stationarity; method 3 is the same as method 1 but only uses the last 30 years of record. Method 4 adds a safety factor (+25% for case 1-2-5, and –25% for cases 3 and 4 for 25 year projection; and ±30% for the 50-year projection available elsewhere) to the real-space 100-year flood estimator calculated by stationarity method-of-moments.

Method 2 includes three different estimators: Trend_0, Trend_1, and Trend_2. Trend_0 assumes the trends β_1 and β_2 in the log-space mean and variance are known, and uses that information to update the distribution parameters with time. Trend_1 estimates the trend in the log-space mean β_1 from the historical records by linear regression and assumes there is no trend in the variance (β_2=0). Trend_2 estimates the trend in both the mean β_1 and variance β_2 from the historical records. All three trend estimators estimate the two constants, α_1 and α_2.

As described in Section 2, the regression equation to estimate the trend in the mean is:

$$X_t = \alpha_1 + \beta_1 t + \varepsilon_{1,t} \quad t = 0, 1, \ldots, T\text{-}1 \qquad (10)$$

where T is the total time elapses from the first year in the record to the year of concern. The natural estimator for the variance model would be Eqn. 5, but it gave very unstable results because extreme large values of the estimator of β_2 generated unreasonable large results as a result of the exponentiation. Here we employ a new model to estimate the trend in the scale parameter for the distribution:

$$\left|X_r - \hat{\mu}_r\right| = \sqrt{\alpha_2 \frac{r\exp\left(0.5\beta_2 r\right) + \exp\left(-0.5\beta_2 r\right)}{\exp\left(0.5\beta_2 r\right) + r\exp\left(-0.5\beta_2 r\right)}} + \delta_{2,r} \quad r = t + 1 = 1, 2, \ldots, T \quad (11)$$

In Eqn. 11 X_r is the logarithm of a recorded flow at the gauged station; $\delta_{2,r}$ is the random error; $\hat{\mu}_r$, equals to $\hat{\alpha}_1 + \hat{\beta}_1(t+1)$, is the estimated expected value of X in

year t+1 provided by Eqn. 10. Eqn. 11 does not exactly match the model used to generated the variance for each year, though the agreement is close for the values of β_2 actually adopted.

In Chapter 3, the trend in the scale parameter is estimated by the regression:

$$\left(X_r - \hat{\mu}_r\right)^2 = \alpha_2 \frac{r\exp\left(0.5\beta_2 r\right) + \exp\left(-0.5\beta_2 r\right)}{\exp\left(0.5\beta_2 r\right) + r\exp\left(-0.5\beta_2 r\right)} + \varepsilon_{2,r} \quad \text{r = t+1 = 1, 2, …, T} \quad (12)$$

The problem with Eqn. 12 is when X has a Pearson type 3 distribution with $|\gamma| > 0$, $(X_r - \hat{\mu}_r)^2$ can be a very highly skewed distribution with very thick tails. As a result, least squares regression for computing model parameters is not efficient and is poorly behaved. Note that in the case of no trend, the residual errors equal $(X_r - \hat{\mu}_r)^2$ minus a constant. When the distribution of X is normal so $\gamma = 0$, $(X_r - \hat{\mu}_r)^2$ is chi-squared with one degree of freedom (i.e. gamma distribution with $\alpha = 0.5$) and has a skewness coefficient of 2.8 with a lower bound of zero; it is not clear that least squares regression is an appropriate choice for $\gamma = 0$, the most benign value for the at-site skew γ. Least squares regression are the MLEs for models with normally distributed residuals, and may perform very poorly for non-normal residuals [White and MacDonald, 1980; Mîndrilă, 2010]. The other problem we experienced is that the simple equation for the variance in Eqn. 5, while always positive, can if extrapolated for a number of years with large $\beta > 0$, produce ridiculously large variances. Employing equation 11, the regression on $|X_r - \hat{\mu}_r|$ should yield an almost unbiased estimator of the absolute deviation. Thus, there is a need to convert the absolute deviation to estimate the standard deviation. A correction factor $\sqrt{\pi/2}$ for normal series is employed in the quantile estimation.

Both Eqn. 11 and 12 are basically method of moments estimators. One could attempt to employ maximum likelihood estimators, which are substantially more complicated and would require identification of all 5 parameters simultaneously, while

avoiding combination for which some observations are either impossible or have infinite likelihood. MLEs would be particularly attractive if one thought the X-variance changed substantially over the period of record. However, Bulletin 17B requires method of moments estimators and we investigate how such estimators might be extended to reflect trends.

6. Results

We denote the LMSE of $\hat{Q}_{100,t}$ in Eqn. 7 as $LMSE_Q$, and the LMSE of $A\hat{E}P_{0.01,t}$ in Eqn. 9 as $LMSE_P$. Figures 1-5 show the $LMSE_Q$ for $\hat{Q}_{100,t}$ with sample size of 100 for a 25-year projection. Each figure summarizes the results of one of the five cases; each panel in the figure displays the LMSEs for one configuration of the trend magnitudes and skews in the range [-1,+1]. To make the comparison clearer, the $LMSE_Q$ in Figures 3 and 4 are truncated at 2.5 and 3.

Comparing the six estimators whose performance is displayed in Fig. 1-5, Trend_0 always performs best as expected – it is always the correct model with the correct values of β_1 and β_2; α_1 and α_2 were estimated. Trend_1, which estimates a trend in only the mean and estimates β_1, performs well in most scenarios. Trend_2 which estimates the trends in both the mean and variance, β_1 and β_2, only beats Trend_1 when trends are large in both the mean and variance ($\beta_1 = \beta_2 = \pm 1\%$). The stationary estimator does poorly when the trend in the mean is ±0.25% or a more extreme value. The safety factor only works well when $\beta_1 = \pm 0.25\%$. The 30-year record estimator is generally a poor estimator – the limited sample size results in high variance even if it mitigates failure to model trends.

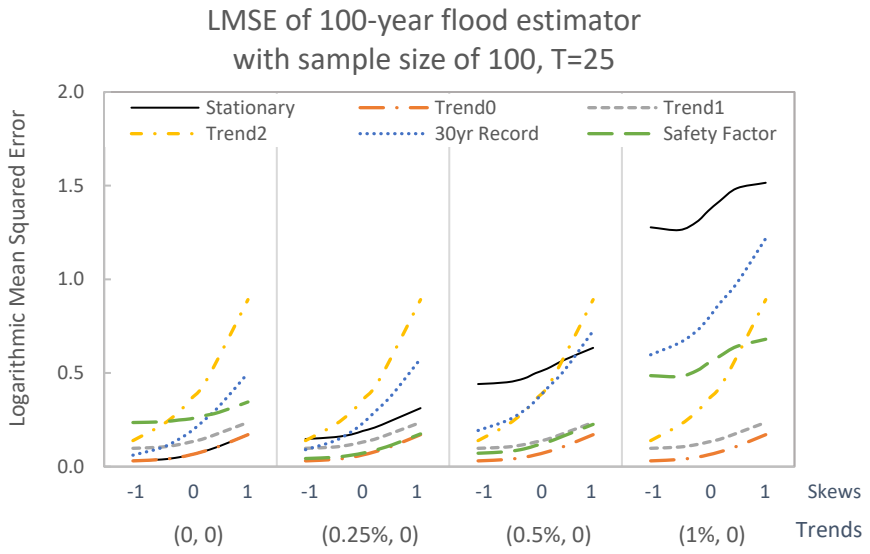


Figure 1. $LMSE_Q$ with different skews for T=25, sample size n = 100, Case 1

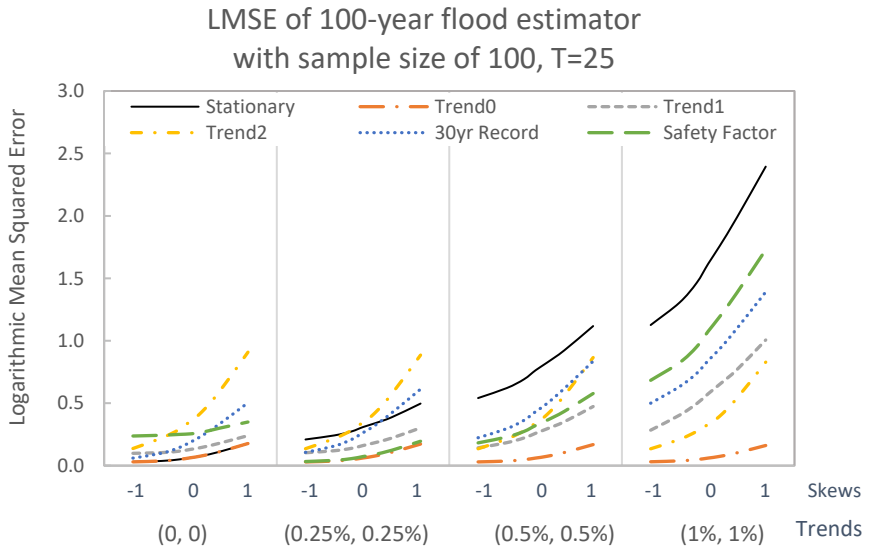


Figure 2. $LMSE_Q$ with different skews for T=25, sample size n = 100, Case 2

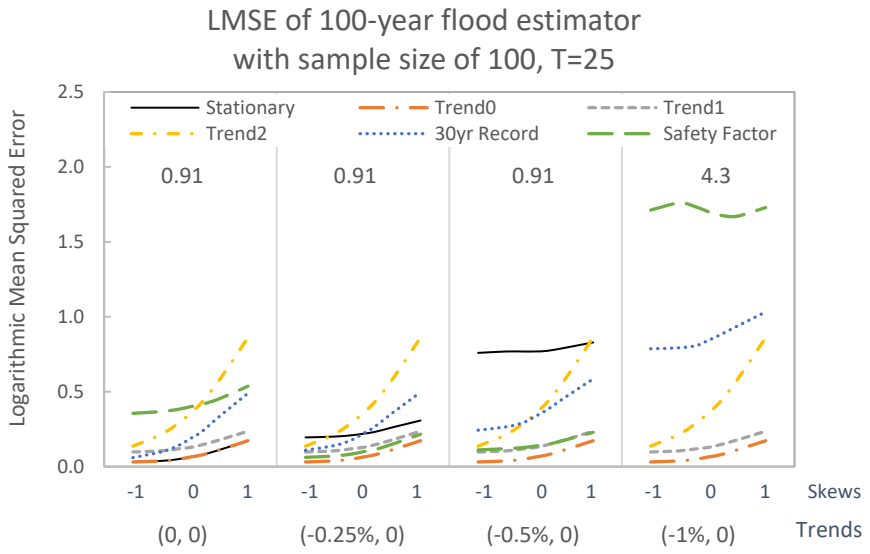


Figure 3. $LMSE_Q$ with different skews for T=25, sample size n = 100, Case 3
LMSEs are truncated at 2.5, "Stationary" on panel 4 has a LMSE larger than 3.4

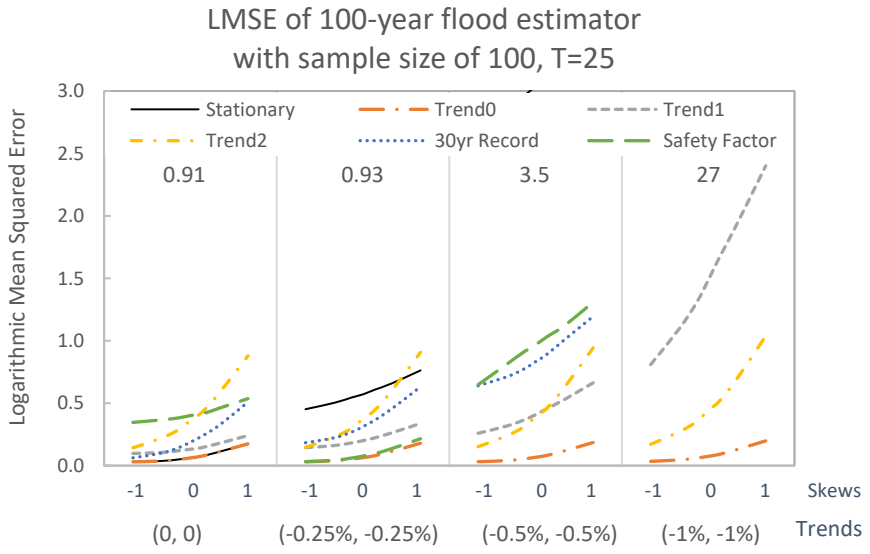


Figure 4. $LMSE_Q$ with different skews for T=25, sample size n = 100, Case 4
LMSEs are truncated at 3. On panel 3: "Stationary" > 2.4
On panel 4: "Stationary" > 20, "30yr Record" > 4, "Safety Factor" > 12

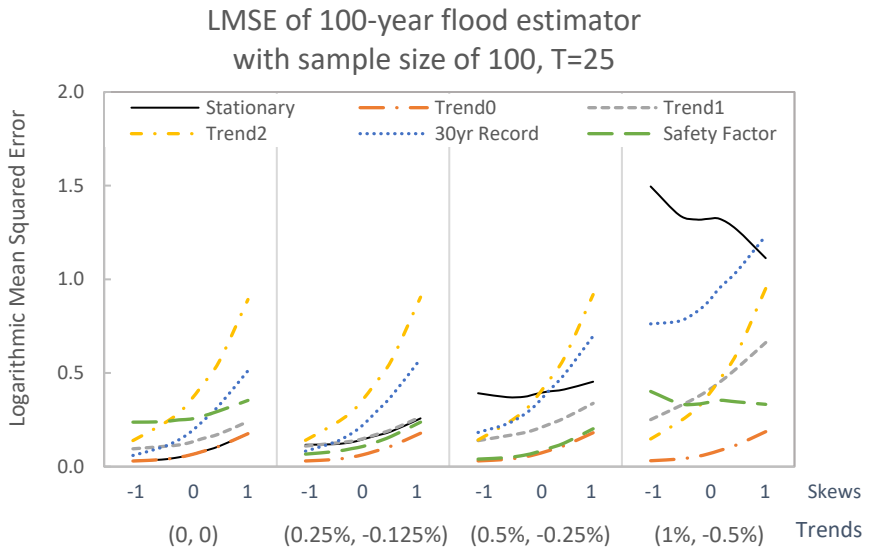


Figure 5. $LMSE_Q$ with different skews for T=25, sample size n = 100, Case 5

The coefficients of skewness within [-1,+1] generally have considerable impact on the relative values of $LMSE_Q$. The values of $LMSE_Q$ for a method and trend in each panel almost always increase with the skew. However, small non-monotonic anomalies occurred with Stationary and Safety Factor estimators in the 4th panels of Cases 3 and 5 (Figures 3 and 5) whose $LMSE_Q$ generally decrease with the skew. Because Safety Factor is just the Stationary estimator plus a constant safety factor, it is not a surprise seeing similar behavior. Stationary does particularly poorly in Case 3, and even worse in Case 4.

Figure 6 displays the $LMSE_P$ for Case 1; $LMSE_P$ for Cases 2-5 appear in Appendix B. Comparing the $LMSE_P$ for $\hat{AEP}_{0.01,t}$ yields conclusions similar to those for $LMSE_Q$, except Trend_2 now beats Trend_1 both when $\beta_1 = \beta_2 = 0.5\%$ as well as and 1% (in Case 2). Whereas $LMSE_Q$ generally increased with skew for a particular trend and method, the values of $LMSE_P$ always decrease with skew: see Case 1 in Figure 5 and Cases 2-5 in Appendix B.

It is interesting that $LMSE_Q$ generally increases with increasing skew, whereas $LMSE_P$ decreases. The skewness has a major impact on the asymmetry of the sample values X_t, and thus the asymmetry of the sample mean, variance, and skew coefficient. However, $LMSE_Q$ and $LMSE_P$ put very different weights on over- and under-estimation errors, resulting in positive trends with skew, γ, for $LMSE_Q$, and negative trends with skew for $LMSE_P$. But why? For example, with $\gamma > 0$, especially large X_t can be generated, resulting in quantile estimators that are also much too large, and the estimation error is squared in Eqn. 7. On the other hand, with $\gamma < 0$, samples would rarely have high outliers, and thus flood quantile estimators that were unusually large were rare. Looking the other way, one does not get unusually-small 100-year flood estimators that result in large-squared losses when $\gamma < 0$ or for any population skew – the n = 40 or 100 sample values effectively put a lower bound on $\hat{Q}_{100}$ [Lamontagne, et al., 2016]. Thus, $LMSE_Q$ increases with increasing skew, γ, because of the increasing likelihood of very large squared errors.

One sees the opposite effect with $LMSE_P$. For $\gamma > 0$, the skew in the observations results in positive skew in the fitted P3 distribution for X_t and a heavy upper tail. In this case, the $A\hat{E}P_{0.01}$ for the true 100-year flood can easily be overestimated, but not by that much, so the ratio of $A\hat{E}P_{0.01}$ to 1% should not be more than an order of magnitude. (It is very unlikely that Q_{100} would be mistaken for the 10-year event with n = 40 or 100.) However for $\gamma < 0$, the skew in the observations results in a very thin upper tail, and sometimes an upper bound that is less than the true 100-year flood value ($\hat{Q}_{100}$). Thus, $A\hat{E}P_{0.01}$ can be several orders of magnitude less than 1%; our $LMSE_P$ with a 10^{-4} base limited the effective error to 2 orders of magnitude. The point is that $A\hat{E}P_{0.01}$ estimators are sensitive to being orders of magnitude too small when $\gamma < 0$, and much less sensitive to large AEP errors when $\gamma > 0$; hence we

see a downward trend in $LMSE_P$ values with increasing γ because large squared losses become less likely.

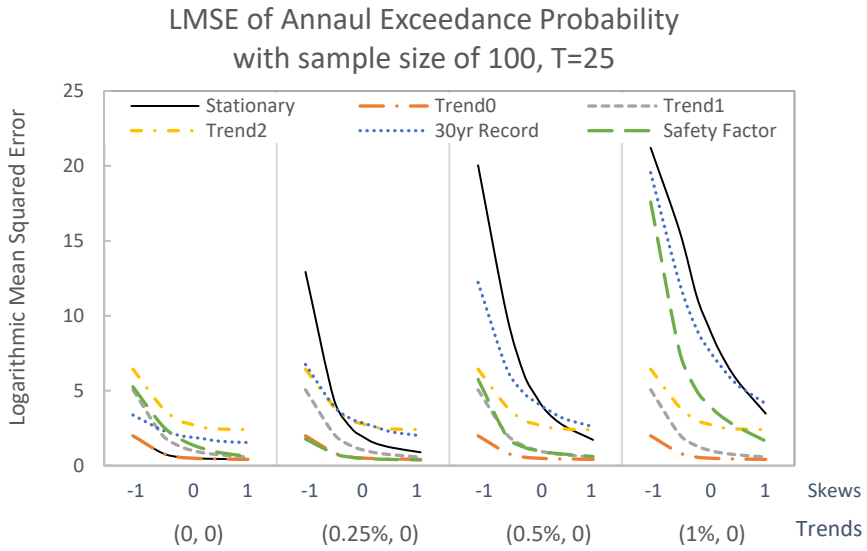


Figure 6. $LMSE_P$ with different skews for T=25, sample size n = 100, Case 1

Figures 7-11 and 12 display the $LMSE_Q$ and $LMSE_P$ with a sample size of only 40 for a 25-year projection. Here the 30-year-record method estimator is not included because the total length of the annual maximum series is only 40 years. The largest values of LMSE for Trend_2 in Figures 6-10 are truncated. Due to the shorter record length, for the quantile estimators, Trend_2 is almost always the worst. Among the estimators, with small trends (within ±0.5%), Trend_1 seems to lose its general superiority. The stationary estimator works well except for the most extreme scenario ($\beta_1 = \pm 1\%$ or $\beta_1 = \beta_2 = \pm 1\%$). The safety factor works surprisingly well with moderate trends (±0.25% and ±0.5%) in all four cases – the correction was about the right value. However, the analysis assumes we know it is supposed to make an upward adjustment in cases 1 and 2, and a downward adjustment in cases 3 and 5.

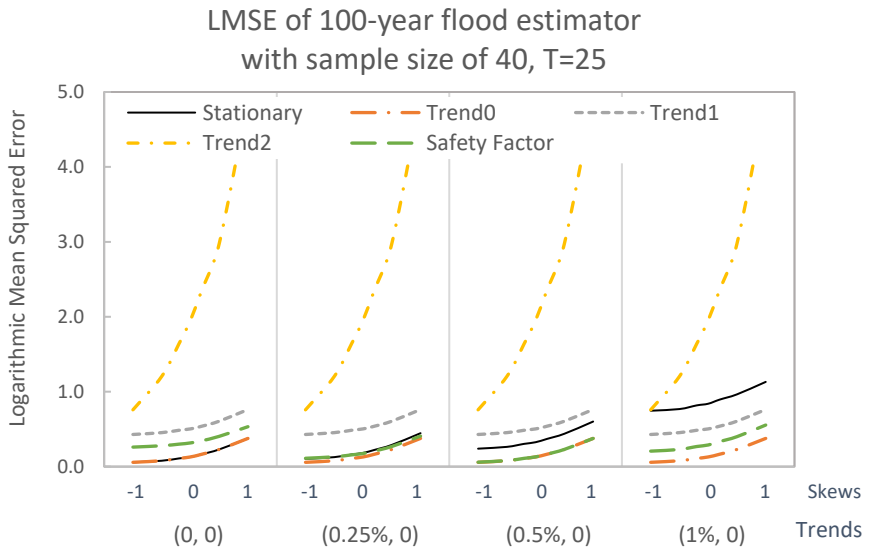


Figure 7. $LMSE_Q$ with different skews for T=25, sample size n = 40, Case 1

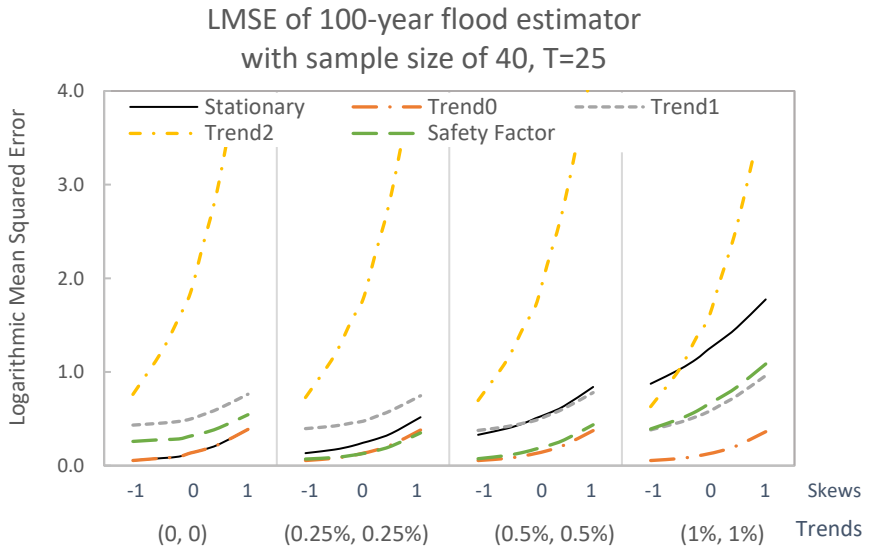


Figure 8. $LMSE_Q$ with different skews for T=25, sample size n = 40, Case 2

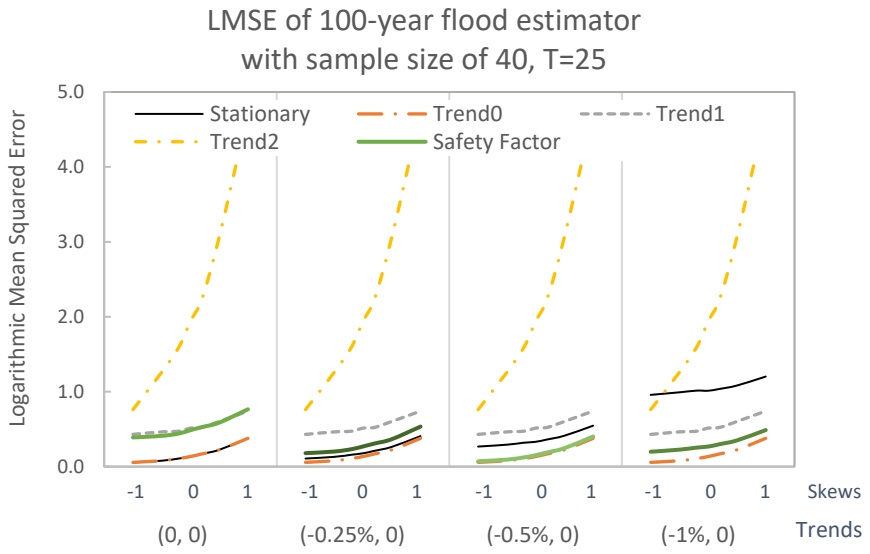


Figure 9. $LMSE_Q$ with different skews for T=25, sample size n = 40, Case 3

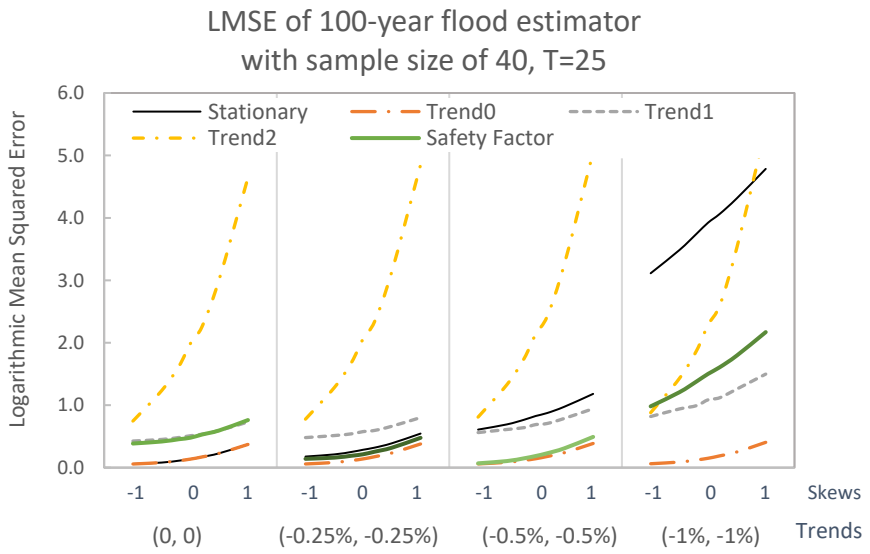


Figure 10. $LMSE_Q$ with different skews for T=25, sample size n = 40, Case 4

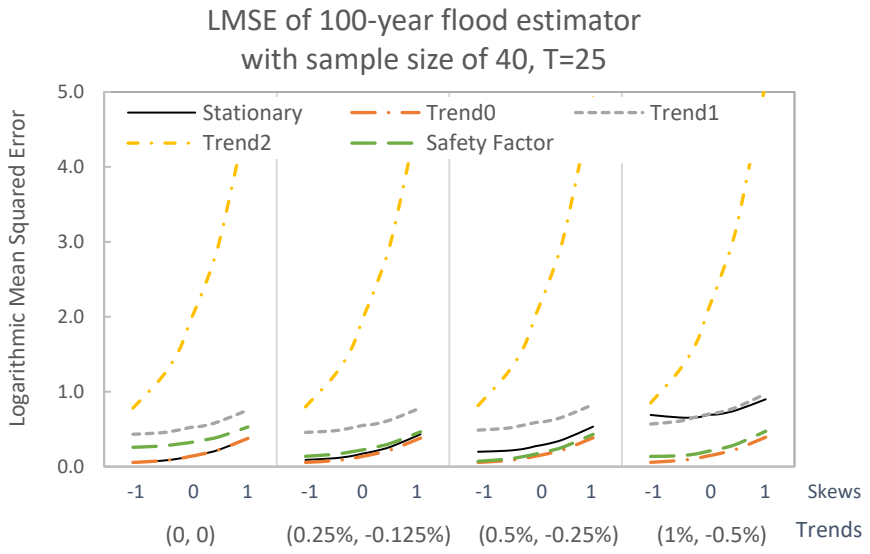


Figure 11. $LMSE_Q$ with different skews for T=25, sample size n = 40, Case 5

Comparing the $LMSE_P$ values in Figure 11 (Case 1), and in Figures 4.B-5 – 4.B-8 in Appendix B (Case 2-5), Trend_0 is always the best. The Stationary and Safety factor estimators tie Trend_0 in several scenarios. Trend_2 is not a good choice except $\beta_1 = \beta_2 = 1\%$. Trend_1 would be an alternative for other extreme scenarios (±1% trends) in all four cases.

With the record length of 40, the coefficients of skewness have relatively consistent impacts on $LMSE_Q$ and $LMSE_P$. For a specific method and trend, the values of $LMSE_Q$ increase with the skew, while, the values of $LMSE_P$ decrease with the skew. However, the clear lesson is that for n = 40, Trend 2 which estimates a trend in the variance is not attractive with the estimator considered.

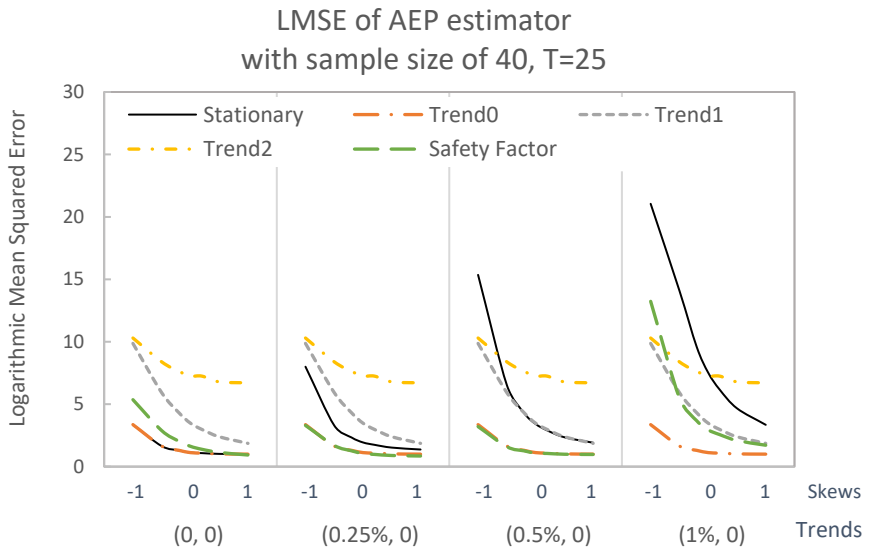


Figure 12. $LMSE_P$ with different skews for T=25, sample size n = 40, Case 1

7. Conclusion

The situation is rather complex and the data do not justify a general sweeping statement as to which method is best. Sample size is very important, as is the magnitude of the trend. And the two interact: because with n = 100, the magnitude of the difference in the true 100-year event and in the X_t variance between the first and last year of the record is over twice that for the n = 40 cases.

Though the trend with skew is different, generally the LMSE for $\hat{Q}_{100,t}$ and $A\hat{E}P_{0.01,t}$ provided the same ranking among methods across every case and panel. Thus we included $LMSE_P$ for $A\hat{E}P_{0.01,t}$, only for Case 1. The results for Cases 2-5 are included in Appendix B.

For n = 40 and small trends (within ±0.25%), the stationarity estimator is about the best among realistic methods (excludes Trend_0 and maybe Safety Factor) across all cases regardless of the skews; for larger trends one should use Trend_1.

For n = 100, surprisingly Trend_1 generally does about as well as can be done – it does nearly as well as stationarity when trends are zero. With 100 years of record

the loss of accuracy with Trend_1 is not very large for the no-trend panel ($\beta_1 = \beta_2 = 0$), whereas in other situations, stationarity which does not estimate a trend can have relatively large errors. On the other hand, only for ±1% or more-extreme-values, Trend_2 is the best realistic choice. Clearly, with the larger sample size, Trend_1 is the robust choice.

A critical issue is that with n = 40, seldom does a standard trend test with our data sets identify a statistically significant trend in the mean when the trend is ±0.5% or less: the type II error is 88% or more. The situation is better for n = 100, but the type II error is still in excess of 66% with ±0.25% per year and less than 30% for cases with trends of ±0.5% per year (see Appendix 4.C). Thus, we are in a situation where we cannot depend on seeing a statistically significant trend to determine whether one should use the Stationary model (no trend) or include the possible trend in the mean (i.e. the Trend_1 estimator). Clearly, flood frequency analysis in an uncertain world will be a challenge.

References

Bobée, B. (1975), The Log Pearson Type 3 Distribution and Its Application in Hydrology, *Water Resources Research,* 11(5), 681-689.

Bobée, B., and R. Robitaille (1977), The Use of the Pearson Type 3 and Log Pearson Type 3 Distributions Revisited, *Water Resources Research,* 13(2), 427-443.

Bobée, B., and Ashkar, F. (1991). *The gamma family and derived distributions applied in hydrology.* Water Resources Pubns. ISBN 0918334683.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values.* Springer Series in Statistics. Springer-Verlag, London.

Chowdhury, J. U., and J. R. Stedinger (1991), Confidence Interval for Design Floods with Estimated Skew Coefficient, J. Hydraul. Eng., 117(7): 811-831

El Adlouni S., Ouarda T.B.M.J., Zhang X., Roy R. et Bobée B. (2007). Generalized maximum likelihood estimators for the nonstationary generalized extreme

value model. *Water Resour. Res.*, 43 (3): W03410. doi: 10.1029/2005WR004545

Greene, William H. (2012), Econometric Analysis, seventh edition, Pearson Education, Inc., Upper Saddle River, N.J. ISBN 0-13-139538-6

Griffis, V. W. (2003). Evaluation of log-Pearson type 3 flood frequency analysis methods addressing regional skew and low outliers. MS thesis, Cornell Univ., Ithaca, N.Y.

Griffis, V. W., J. R. Stedinger, and T. A. Cohn (2004), Log Pearson type 3 quantile estimators with regional skew information and low outlier adjustments, *Water Resour. Res.*, 40, W07503, doi:10.1029/2003WR002697.

Griffis, V. W., and J. R. Stedinger (2007a), The LP3 distribution and its application in flood frequency analysis, 1. Distribution Characteristics, *J. Hydrol. Eng.*, 12(5), 482-491.

Griffis, V. W., and J. R. Stedinger (2007b), The LP3 distribution and its application in flood frequency analysis, 2. Parameter estimation methods, *J. Hydrol. Eng.*, 12(5), 492-500.

Griffis, V. W., and J. R. Stedinger (2009), The log-Pearson type 3 distribution and its application in flood frequency analysis, 3. Sample skew and weighted skew estimators, *J. Hydrol. Eng.*, 14(2), 209-212.

Gruber, A. M., and Stedinger, J. R. (2008). Models of LP3 regional skew, data selection, and Bayesian GLS regression. *Proc., World Environmental & Water Resources Conf. 2008*, R. Babcock and R. Walton, eds., ASCE, Reston, VA, Paper No. 596.

Hirsch, R. M. and Ryberg, K. R. (2012). Has the magnitude of floods across the USA changed with global CO_2 levels? *Hydrological Sciences Journal*, 57.1:1-9.

Hirsch, R. M. and S. A. Archfield (2016). Not higher but more often. *Natural Climate Change*, Vol 5, 198-199.

IACWD (1982). Bulletin 17b of the hydrology subcommittee: Guidelines for determining flood flow frequency. Technical report, U.S. Geological Survey.

IPCC (2012). *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK, and New York NY, USA.

Jain, S. and Lall, U. (2001). Floods in a changing climate: does the past represent the future? *Water Resources Research*, 37.12:3193-3205.

Koutsoyiannis, D. and Montanari, A. (2007). Statistical analysis of hydroclimatic time series: Uncertainty and insights. *Water Resour. Res.,* 43.5: W05429.

Lamontagne, J.R., J.R. Stedinger, Xin Yu, C.A. Whealton, and Ziyao Xu (2016). Robust Flood Frequency Analysis: Performance of EMA with Multiple Grubbs-Beck Outlier Tests, *Water Resour. Res.,* 52(4), 3068–3084 doi:10.1002/2015WR018093

Landwehr, J. M., and N. C. Matalas (1978), Some Comparisons of Flood Statistics in Real and Log Space, *Water Resour. Res.*, 14.5:902-920.

Lins, H. F. and Slack, J. R. (1999). Streamflow trends in the United States. Geophysical Research Letters, 26.2:227-230.

Madsen, H., Lawrence, D., Lang, M., Martinkova, M., and Kjeldsen, T. (2013). *A Review of Applied Methods in Europe for Flood-Frequency Analysis in a Changing Environment*. FLOODFREQ COST Action ES0901. Center for Ecology & Hydrology.

Mentaschi, L. et al. (2016). The transformed-stationary approach: a generic and simplified methodology for non-stationary extreme value analysis. *Hydrol. Earth Syst. Sci.*, 20, 3527–3547

Mîndrilă, D. (2010), Maximum Likelihood (ML) and Diagonally Weighted Least Squares (DWLS) Estimation Procedures: A Comparison of Estimation Bias with Ordinal and Multivariate Non-Normal Data, *International Journal of Digital Society (IJDS),* 1(1): 60-66

Parrett, C., Veilleux, A., Stedinger, J. R., Barth, N. A., Knifong, D. L., and Ferris, J. C. (2011), *Regional skew for California, and flood frequency for selected sites in the Sacramento–San Joaquin River Basin, based on data through water year 2006,* U.S. Geological Survey Scientific Investigations Report 2010–5260, 94 p.

Stedinger, J. R. and Griffis, V. W. (2011). Getting from here to where? Flood frequency analysis and climate. *JAWRA Journal of the American Water Resources Association,* 47.3:506-513.

Tasker, G. D., and Stedinger, J. R. (1986). Estimating generalized skew with weighted least squares regression. *J. Water Resour. Plann. Manage.,* 112(2): 225–237.

Veilleux, A. G., J. R. Stedinger, and J. R. Lamontagne (2011), Bayesian WLS/GLS Regression for Regional Skewness Analysis for Regions with Large Cross-Correlations among Flood Flows, *World Environmental and Water Resources Congress*: 3103-3112

Vogel, R. M., W. O. Thomas Jr., and T. A. McMahon (1993). Flood-flow frequency model selection in southwestern united states, *Journal of Water Resources Planning and Management,* 119(3), 353-366

Weaver, J.C., Feaster, T.D., and Gotvald, A.J., (2009), *Magnitude and frequency of rural floods in the Southeastern United States, through 2006—Volume 2, North Carolina:* U.S. Geological Survey Scientific Investigations Report 2009–5158, 110 p.

White, H., and G. M. MacDonald (1980), Some Large-Sample Tests for Nonnormality in the Linear Regression Model. Journal of the American Statistical Associates, 75(369): 16-28.

Appendices for Chapter 4

Appendix 4.A Approximation of MSE of at-site skew estimator

The mean squared error of the at-site skew estimator $\hat{\gamma}$ ($MSE_{\hat{\gamma}}$ in Eqn. 6) is calculated using Eqn. A1-A4, which were derived by Griffis and Stedinger [2009]:

$$MSE\left[\hat{\gamma}\right] \approx \left(\frac{6}{N}+a\right)\left[1+\left(\frac{9}{6}+b\right)\gamma^2+\left(\frac{15}{6*8}+c\right)\gamma^4\right] \qquad \text{(A1)}$$

where a, b, and c are correction factors for small samples:

$$a = -\frac{17.75}{N^2}+\frac{50.06}{N^3} \qquad \text{(A2)}$$

$$b = \frac{3.93}{N^{0.3}}-\frac{30.97}{N^{0.6}}+\frac{37.1}{N^{0.9}} \qquad \text{(A3)}$$

$$c = -\frac{6.16}{N^{0.56}}+\frac{36.83}{N^{1.12}}-\frac{66.9}{N^{1.68}} \qquad \text{(A4)}$$

valid for $|\gamma| \leq 1.4$ and sample sizes N from 10 to 150 [Griffis, 2003]. Here we use the regional skew G as the estimator of the at-site skew (γ) whose value is needed in (A1).

Appendix 4.B Results of LMSE for 1% AEP (Cases 2–5)

Figures B1-B8 plot the logarithm mean squared errors of the 1% annual exceedance probability ($LMSE_P$) for Case 2, 3, 4 and 5; see Eqn. 9 in text. The results for Case 1 are presented in Chapter 4, Section 6 (Figures 6 and 12).

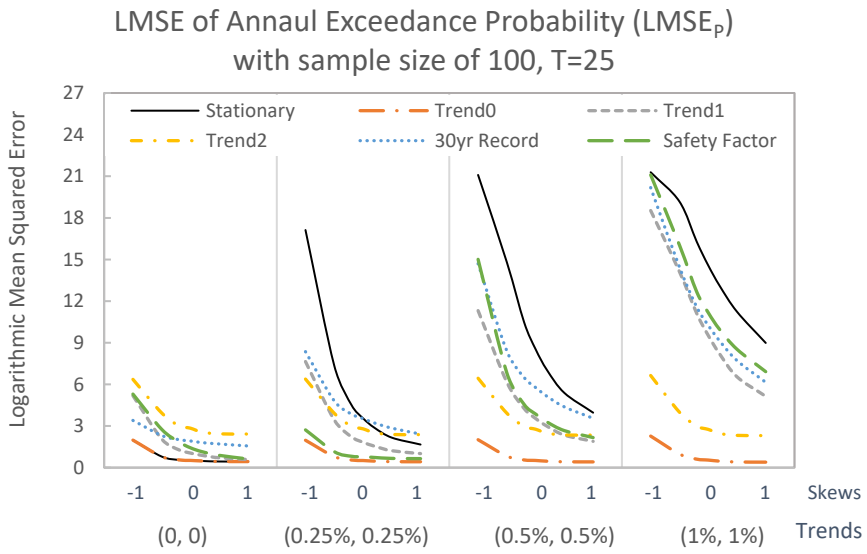


Figure 4.B-1. $LMSE_P$ with different skews for T=25, n = 100, Case 2

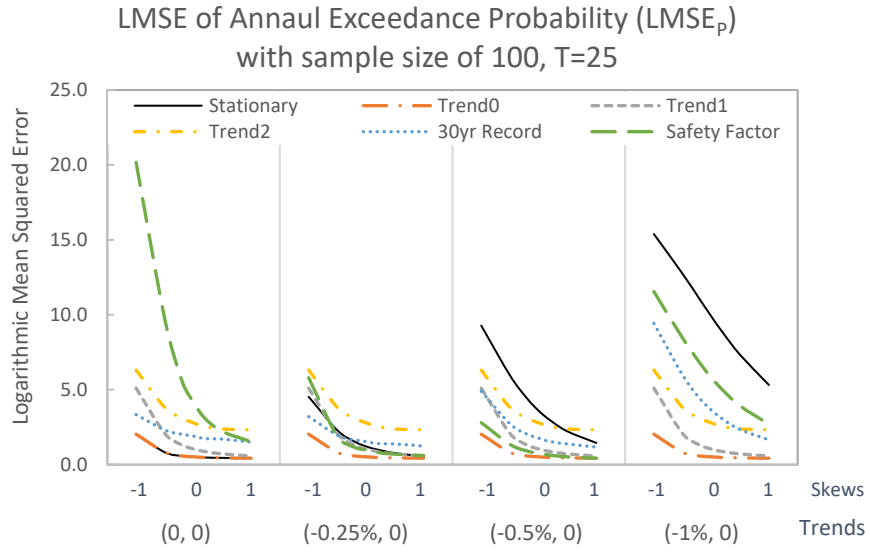


Figure 4.B-2. $LMSE_P$ with different skews for T=25, n = 100, Case 3

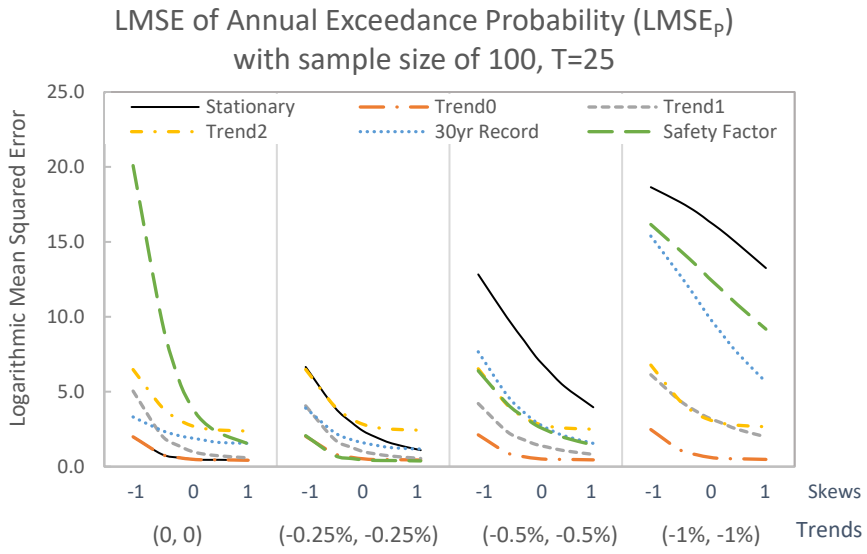


Figure 4.B-3. LMSE$_P$ with different skews for T=25, size n = 100, Case 4

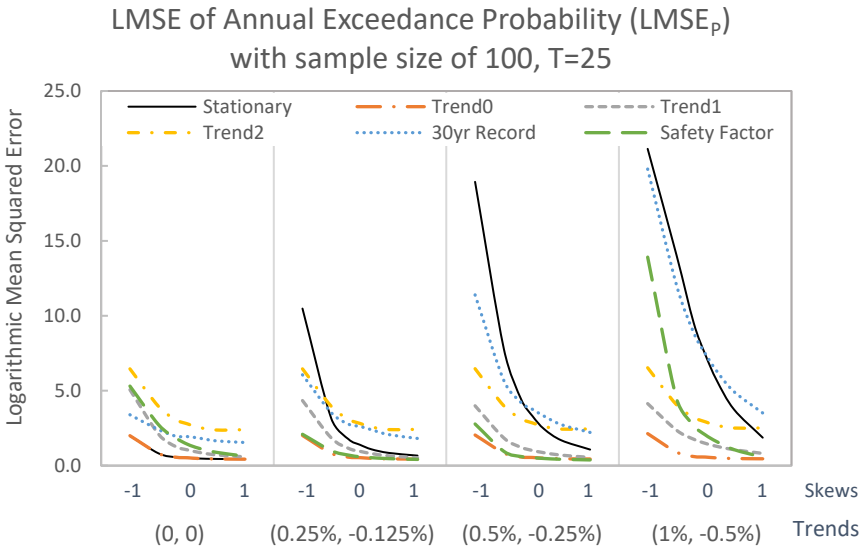


Figure 4.B-4. LMSE$_P$ with different skews for T=25, size n = 100, Case 5

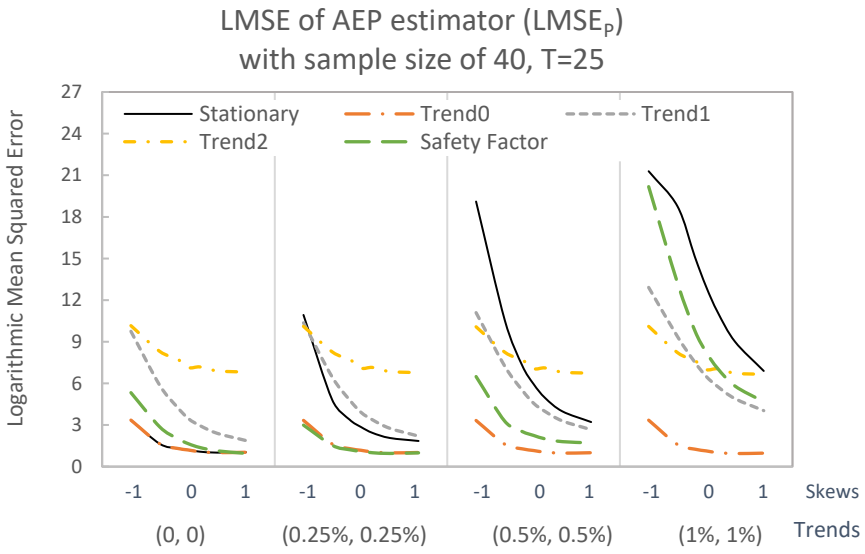


Figure 4.B-5. LMSE$_P$ with different skews for T=25, n = 40, Case 2

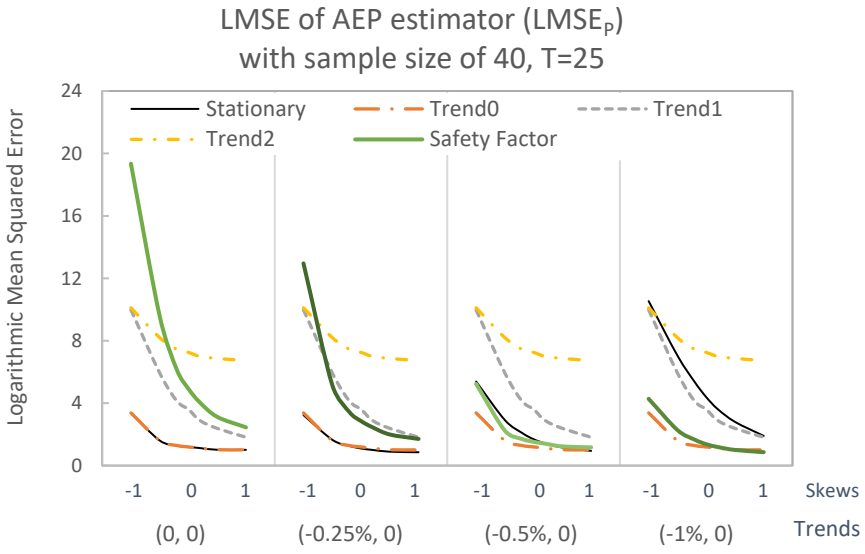


Figure 4.B-6. LMSE$_P$ with different skews for T=25, n = 40, Case 3

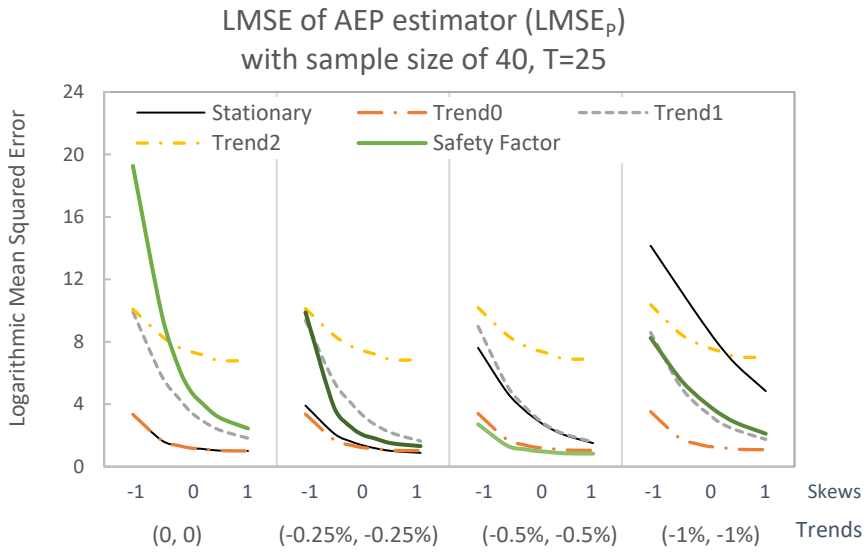


Figure 4.B-7. LMSE$_P$ with different skews for T=25, n = 40, Case 4

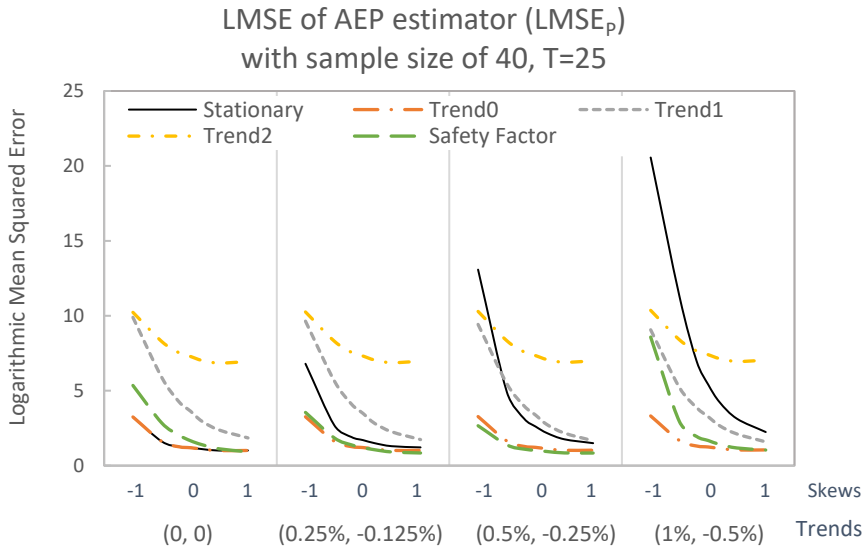


Figure 4.B-8. LMSE$_P$ with different skews for T=25, n = 40, Case 5

Appendix 4.C Type II Error

This appendix reports results on a type II error analysis for the five cases considered in Chapter 4. The type II error for a trend test is the probability the test concludes (based on the systematic record) that there is no trend in the mean when a trend in the mean did exist. The type II error is important because one would like to use a trend test to determine if the stationary estimator is appropriate, or one should use Trend_1.

The analysis considers applying the trend test to the annual maximum series and assumes the test statistic $\hat{\beta}_1$ follows a normal distribution. Based on a Monte Carlo simulation with 100,000 replicate samples, the standard error of that estimator $\hat{\beta}_1$ is approximately 0.002 for sample size of 100, and 0.007 for sample size of 40. Table 4.C-1 and 4.C-2 list the estimated type II errors for a z-test (with $\alpha = 0.05$) for selected cases. In all cases the regional skewness was zero. The regional skew effected the values generated for the at-site skew which determined the distribution of the generated *X*-record representing possible flood records.

When n=40 (Table 4.C-1), the type II error is 88% or more for all the cases included, corresponding to trends within ±0.5% per year. That indicates that with a 40 year record the analysis would generally not identify a trend in the mean. When n=100 (Table 4.C-2), the type II errors are still in excess of 66% for trends of ±0.25% per year or less. However, the 100-year flood will be significantly under-/over-estimated if such trends are neglected. For n = 100, the test does much better for trends of ±0.5% per year, but that is a large trend. Clearly trend tests should be employed, but they are not sufficiently sensitive to identify all of the situations in which it would be appropriate to estimate a trend in the mean.

Table 4.C-1. Type II errors of trend detection for log-mean flood $\widehat{\beta}_1$ (n=40)

	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.
Trend (per year)	**±0.25%**	**0**	**0.25%**	**0.25%**	**-0.25%**	**-0.25%**	**0.25%**	**-0.125%**
Type II error	93.4%		93.5%		93.4%		93.4%	
Trend (per year)	**±0.5%**	**0**	**0.5%**	**0.5%**	**-0.5%**	**-0.5%**	**0.5%**	**-0.25%**
Type II error	88.6%		89.2%		88.0%		88.3%	

Table 4.C-2. Type II errors of trend detection for log-mean flood $\widehat{\beta}_1$ (n=100)

	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.
Trend (per year)	**±0.25%**	**0**	**0.25%**	**0.25%**	**-0.25%**	**-0.25%**	**0.25%**	**-0.125%**
Type II error	69.6%		72.6%		66.3%		68.0%	
Trend (per year)	**±0.5%**	**0**	**0.5%**	**0.5%**	**-0.5%**	**-0.5%**	**0.5%**	**-0.25%**
Type II error	17.5%		27.9%		9.68%		13.2%	

CHAPTER 5
CONCLUSIONS

This thesis addresses two challenges in flood frequency analysis: efficient parameter estimation with mixed populations, and addressing the possible impact of climate change on flood risk over time. Chapter 2 discussed the estimators that can be used to describe the risk of large floods when the annual maximum series consists of events that are generated from more than two distinct physical processes (e.g. rainfall and snowmelt). Chapters 3 and 4 explored the methods that incorporate climate change impacts into the quantile and exceedance probability estimation.

Conclusions from Chapter 2

The analysis evaluated three risk estimation models – Mixture model, Joint model, and annual maximum model – and different parameter estimators that go with each model (see Table 2 in Chapter 2).

Ideally, fitting the joint model would be the best approach for describing the mixed population when both the rainfall and snowmelt flood series records are complete, whether or not they are independent, because that method provides the most complete model and employs the largest and most complete data set. However, for all the estimators, the log-space cross correlation (ρ) between the rainfall and snowmelt events has a modest effect on the mean squared errors of 10- and 100-year flood estimators for a concurrent snowmelt-rainfall correlation $\rho \leq 0.5$. Thus, the simple mixture estimator that neglects any cross-correlation between the two series provides an accurate approximation of the annual maximum distribution for quantile estimation when ρ of 0.5 or less; this is the case for most rivers in the western U.S. [Elliott, et al., 1982]. If the rainfall events determine the flood risk, the two-parameter lognormal estimator of rainfall-only model should be adequate for describing the flood risk.

When available flood records include only the annual maximum flood series, it is more difficult to develop models of the two series. Two such estimators were considered. The expected moments algorithm (EMA) uses a censored sampling paradigm to develop models of each series by describing annual maximum snowmelt and rainfall that are not recorded as less than the larger values of the annual maximum in the other series. Kirby on the other hand suggested using the conditional distributions for snowmelt and for rainfall given they are the annual maximum for the year considering both series. In realistic cases considered, EMS's flood risk estimator (EMA) performed better than the Kirby estimator for 100-year flood estimation when the concurrent snowmelt-rainfall correlation ρ was 0.5 or less; for 10-year flood estimation, the Kirby and EMA estimators' performance is about the same, though EMA is often a little better.

In situation where one source of floods provides the largest annual maxima, and a second source provides most of the small floods, it is reasonable to split the annual maximum series into distinct physical processes and model them separately. When both rainfall and snowmelt floods records are available, the simple mixture estimator is recommended, because correlation between rainfall and snowmelt floods has modest effects on the distribution of the annual maximum series. Among the three estimators that only use the annual maximum series, the EMA is the best choice for 100-year flood estimation when the annual maximum series results from two distinct flood sources with different simple distributions. The Kirby and single LN3 estimators are less accurate. Still the single LN3 estimator is simpler than EMA, and is certainly appropriate in application where mixtures are not a critical issue because one source of flooding dominates the flood risk, or the two sources have similar distributions.

Conclusions from Chapter 3

This chapter considers several simple approaches to flood frequency analysis when records may have trends, and evaluates their performance via a Monte Carlo re-sampling study. Unfortunately, the probability of failing to detect a modest trend in annual maximum series is high (see Table 4 in Chapter 3); however, the 100-year flood estimator could in expectation appreciably underestimate/overestimate the true value if such trends are neglected.

When the trend magnitudes are known and employed to update the parameters of LP3 distribution year by year, the 100-year flood estimators (Trend_0) have the smallest log-space MSE. In practice, it is difficult to know the exact magnitude of trends. The alternative is to estimate trends based on the record. Trend_1 estimates the trend in the mean of the annual maximum log-series; it performs well for most cases except the extreme scenario ($\beta_1 = \beta_2 = \pm 1\%$). Trend_2 estimates the trend in both the mean and the variance of the log-space peak-flow, resulting in a large variance in the 100-year flood estimator and the largest MSE when trends were modest (within ±0.5% per year). The procedure to develop the model for Trend_2 is more complicated than a linear regression that people usually suggest, thus, estimation of a trend in the variance should only be done with caution.

Use of a shorten record (30 years in our case) is not a good choice because of the shorter record and the past observations don't accurately reflect the trends in the future without an adjustment. Using a safety factor does well when a safety factor happens to have about the right value; it is not recommended.

In the context of climate change, the selection of a best estimator is essentially a trade-off between the variance and the bias of the estimators.

Conclusions from Chapter 4

This chapter expanded the work in Chapter 3, which applied the log-Pearson type 3 distribution to samples obtained by resampling a real record, and considered the impacts of both the skewness coefficient and the record length. The situation is even more complex and the Monte Carly study results do not justify a general statement as to which method is best.

Sample size is very important, as is the magnitude of the trend. And the two interact: with n = 100, the magnitude of the difference in the true 100-year event and in the variance between the first and last year of the record is over twice that for the n = 40 cases.

Generally, the LMSE for the 100-year flood estimator and the 1% annual exceedance probability estimator provided the same ranking among methods across every case and configuration of different trends.

For n = 40 and small trends (within ±0.25%), the stationarity estimator is about the best among realistic methods (excludes Trend_0 and maybe Safety factor) across all cases regardless of the skews; for larger trends one should use Trend_1.

For n = 100, surprisingly Trend_1 generally does about as well as can be done – it does nearly as well as stationarity when trends are zero. With 100 years of record the loss of accuracy with Trend_1 is not very large for the no-trend scenario ($\beta_1 = \beta_2 = 0$), whereas in other situations, stationarity which does not estimate a trend can have relatively large errors. Only for ±1% or more-extreme-values is Trend_2 the best realistic choice. Clearly, with the larger sample size, Trend_1 is the robust choice.

Future Work and Unanswered Questions

Flood frequency analysis with mixtures has not received much attention in the literature. A critical issue is if the distribution of the annual maximums is poorly

behaved with a large skew making it difficult to fit. Chapter 2 provides a basis for making reasonable recommendations. The expected moments algorithm (EMA) proposed for the first time here had advantages when dealing with mixed population when all that was available was the annual maximum series. If one has both the rainfall and snowmelt series, which can be the situation in hydrologic practice, then the simple mixture is very attractive. Comparing the mean squared errors among the six estimators in Chapter 2, EMA performs generally well when the correlation is within the reasonable range. On the other hand, EMA generates the complete models of the annual maxima for both the rainfall and snowmelt events. which can be augmented with a regional skew, historical information, or information on land use or climate change to improve the efficiency of flood risk estimation. Future work should look more carefully at the properties of snow and rainfall generated floods in different regions so as to determine how much might be gained by use of mixture models, rather than just modelling the annual maximum series.

Results in Chapter 3 and 4 show that the estimator with the best performance is Trend_0, which unrealistically assumes the true trends are known. It indicates that obtaining accurate estimates of the magnitudes of any trend for the future annual maximum flood series is critical. Besides relying on the data from flood records, we naturally think of a regional estimator or getting information from atmospheric models that represent the change in climate state more closely. With the longer records (n = 100), Trend_1 that estimated a trend in the mean did relatively well, even when the trend parameters were zero. However, Trend_2 that estimated trends in both the mean and the variance seldom did well. Fortunately, a constant log-space variance, corresponding to a constant coefficient of variation in real space, is a reasonable physical assumption. Trends in the real or log-space mean are a true concern.

As mentioned in some studies [Bloschl, 2006; Mallakpour and Villarini, 2015; Hirsh and Archfield, 2016], incorporating precipitation records or long-term climatic indexes (such as temperature) into flood frequency analysis remains an opportunity for extreme quantile estimation whose promise has yet to be captured. Alternatively, reliable downscaling methods that are needed to employ GCMs results for a specific region could be applied [Benestad et al., 2007; Gutmann, et al., 2014]. Clearly, flood risk management in an uncertain world will be a challenge.

References

Benestad, R. E., I. Hanssen-Bauer, and E. J. Forland (2007). An evaluation of statistical models for downscaling precipitation and their ability to capture long-term trends. *Int. J. Climatol*. 27: 649-665

Bloschl, G. (2006). Statistical Upscaling and Downscaling in Hydrology. *Encyclopedia of Hydrological Sciences.1:9*. Published Online.

Elliott, J. G., R. D. Jarrett, and J. L. Ebling (1982), *Annual snowmelt and rainfall peak-flow data on selected foothills region streams, South Platte River, Arkansas River, and Colorado River basins, Colorado,* U.S. Geological Survey Open File Report 82-426.

Gutmann, E., T. Pruitt, M. P. Clark, L. Brekke, J. R. Arnold, D. A. Raff, and R. M. Rasmussen (2014). An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.*, 50, doi: 10.1002/2014WR015559

Hirsch, R. M. and S. A. Archfield (2016). Not higher but more often. *Natural Climate Change,* Vol 5, 198-199.

Mallakpour, I. and G. Villarini (2015), The changing nature of flooding across the central United States, *Nature Climate Change*, 5, 250–254.