Reservoir Re-operation, Risk, and Levee Failure Analysis: Mokelumne River Case

By

PATRICK JI
B.S. (Tsinghua University, China) 1994
M.S. (University of Houston) 2000

DISSERTATION
Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY
in

Civil and Environmental Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Jay R. Lund, Chair

Bassam A. Younis

Alexander Aue

Committee in Charge

2011

ABSTRACT

Reservoir operation for flood control requires accurate inflow frequency analysis which involves the multivariate characteristics of flood peaks, volumes and duration. A complete understanding of flood events involves the joint probabilistic behaviors of these correlated variables. The inflow raw data are commonly transformed for flood frequency analysis, often to a form of lognormal distribution. The multivariate distribution is important for analyzing a flood episode. Flood hydrograph design is a key component flood control rule design in reservoir operation. Numerous methods have been developed to represent flood hydrograph magnitude, duration, volume and shape. Using probability density functions (PDFs) to fit the shapes of flood hydrographs has drawn more attention recently due to improvements in statistical techniques, including algorithms for fitting. Reservoirs transform unregulated flow to regulated flow with different operation rules. Regulated versus unregulated flow curves represent aspects of a reservoir flood control system. However, an accurate curve is difficult due to the complicated physical setting and uncertainties from operations. Lastly, levee failure has drawn attention due to rapid urbanization behind levees and climate change increasing hydrologic extremes. Levee failure can have several mechanisms. In California, levee failure mechanisms mainly are overtopping and erosion.

To address these issues, this dissertation presents a vertical process from inflow analysis through reservoir re-operation to levee failure analysis. First, it presents a procedure for using the bivariate normal distribution to describe the joint distributions of correlated flood peaks and volumes, and correlated flood volumes and durations. Joint distributions, conditional distributions, and the associated return periods of these random variables can be readily derived from their marginal distributions. The theoretical distributions show a good fit to observed ones. The return periods will be used for risk analysis of flood storage space changes.

After inflow multivariate analysis, this dissertation presents three steps to design flood hydrographs for reservoir reoperation: 1) Flood hydrographs separation and modification: Typical flood hydrographs were separated, selected and converted to dimensionless ones; 2) PDF fitting and selection: Beta, Gamma. Lognormal and Weibull distributions were selected and compared to be scaled to fit modified hydrographs based on goodness of fit criteria including RMSE and coefficients of determinations. 3) Development of design flood hydrographs: Design shape variables were estimated from frequency analysis and finally, design flood hydrographs including 10-, 20-, 50-, 100- and 200-year return periods were derived from the combinations of hydrographs shape, flood volume and durations.

To estimate regulated flow frequency for a reservoir's flood storage allocation, this dissertation presents three main steps including: unregulated flow frequency analysis, unregulated/regulated flow transformation and regulated flow frequency estimation. The main contributions include separating flood pulses from daily inflow time series by base flow criteria, modification of unregulated flow calculations, and fitting unregulated and regulated flow to appropriate probability distributions. Unregulated versus regulated flow curves are found using USACE's ResSim software.

Lastly, this dissertation introduces a framework to assess levee failure probability from both overtopping and erosion incorporating uncertainties from hydrologic, hydraulic and geotechnical factors. Two main contributions include overall risk estimation and load-resistance interference risk analysis. Overtopping and erosion failure probability analysis are performed separately in water resources and geotechnical engineering. This chapter presents a more comprehensive risk combining these two failure mechanisms. Load-resistance interference risk analysis is introduced to consider overtopping between flood magnitude and levee capacity and erosion failure between velocity and soil strength. Both analyses are performed by Monte Carlo simulation to estimate overall levee failure probability. Failure probability can be very sensitive to geotechnical variables and less sensitive to reservoir operation.

Camanche and Pardee reservoirs and Lower Mokelumne River levee system in Northern California are used as example applications.

ACKNOWLEDGEMENTS

I would like to thank Professor Jay R. Lund, my major advisor and committee chairman, for his help, support, encouragement, and guidance throughout the past three years. His kindness, enthusiasm and insight make working with him full of joy. I would also like to thank Professor Alexander Aue, committee member, for his valuable advice on my research. I am grateful to Professor Bassam Younis for severing on my committee. I would like to thank the former and current research group members, in particular, Tingju Zhu, for his suggestions and Nate Burley, David Rheiheimer, Kent Ke, Dan Nover and John Hickey, for their many helps. Finally, I would like to thank my family. My parents, Dingquan Ji and Junqing Li, my parent-in-law, Qiusheng Lian and Lingpeng Zeng were always supportive and understanding. I particularly want to thank my wife, Quan Zeng, for her love, patience and understanding.

TABLE OF CONTENTS

ABSTRACT ..... **ii**

ACKNOWLEDGEMENTS ..... **iii**

TABLE OF CONTENTS ..... **iv**

TABLE OF FIGURES ..... ***vi***

LIST OF TABLES ..... ***viii***

CHAPTER 1 INTRODUCTION **1**

1.1 MOTIVATION ***1***

1.2 ORGANIZATION ***2***

CHAPTER 2 MULTIVARIATE FLOOD EPISODE ANALYSIS OF RESERVOIR INFLOW **4**

SUMMARY ***4***

2.1 INTRODUCTION ***4***

2.2 CONCEPTS DEVELOPMENT ***5***

2.3 METHODS ***8***

2.4 APPLICATION TO CAMANCHE/PARDEE RESERVOIR ***9***

2.5 CONCLUSIONS ***16***

CHAPTER 3 DEVELOPMENT OF DESIGN FLOOD HYDROGRAPHS USING PDF SHAPES **22**

SUMMARY ***22***

3.1 INTRODUCTION ***22***

3.2 CONCEPT DEVELOPMENT ***23***

3.3 METHOD ***28***

3.4 APPLICATION ON LOWER MOKELUMNE RIVER LEVEE SYSTEM ***31***

3.5 CONCLUSIONS AND LIMITATIONS ***38***

CHAPTER 4 REGULATED FLOOD FLOWS FREQUENCY CALCULATION THROUGH RESERVOIR RE-OPERATION **42**

SUMMARY ***42***

4.1 INTRODUCTION ***42***

4.2 LITERATURE REVIEW AND CONCEPTS DEVELOPMENT ***44***

4.3 METHODS ***51***

4.4 APPLICATION IN CAMANCHE/PARDEE RESERVOIRS ***53***

4.5 COMPARISON WITH USACE APPROACH – CHARACTERISTIC TIME (CT) TO SEPARATE FLOODS *65*

4.6 CONCLUSION AND FUTURE RESEARCH SUGGESTIONS *70*

CHAPTER 5 FLOOD LEVEE FAILURE ANALYSIS WITH HYDROLOGIC, HYDRAULIC AND GEOTECHNICAL UNCERTAINTIES **73**

SUMMARY *73*

5.1 INTRODUCTION *73*

5.2 LITERATURE AND CONCEPTS DEVELOPMENT *75*

5.3 METHODS *85*

5.4 APPLICATION ON LOWER MOKELUMNE RIVER LEVEE SYSTEM *88*

5.5 CONCLUSIONS AND LIMITATIONS *91*

CHAPTER 6 SUMMARY AND CONCLUSIONS **96**

TABLE OF FIGURES

FIGURE 1.1 FLOW CHART OF RESERVOIR RE-OPERATION, RISK AND LEVEE FAILURE ANALYSIS 2
FIGURE 2.1: CHARACTERISTICS VALUES OF A FLOOD EVENT 6
FIGURE 2.2 SCHEMATIC ILLUSTRATION OF SEPARATION OF SURFACE RUNOFF FROM BASEFLOW 6
FIGURE 2.3 FLOW CHART OF FREQUENCY ANALYSIS OF OBSERVED FLOOD HYDROGRAPH 9
FIGURE 2.4 DAILY UNREGULATED INFLOW TO PARDEE RESERVOIR (1921- 2009, USGS AND DWR, LATITUDE 38.1061°N, LONGITUDE 121.5711°W.) 11
FIG 2-5(A): TRANSFORMED FLOOD PEAK(Q) ON NORMAL PAPER 13
FIG 2-5(B): TRANSFORMED FLOOD DURATION (D) ON NORMAL PAPER 13
FIG 2-5 (C): TRANSFORMED FLOOD VOLUME (V) ON NORMAL PAPER 14
FIGURE 2-6 COMPARISON OF OBSERVED AND THEORETICAL PROBABILITIES OF FLOOD PEAKS AND VOLUMES 15
FIGURE 2-7(A-C): JOINTLY PDF OF FLOOD PEAK AND VOLUME WITH DIFFERENT VIEW DIRECTIONS 17
FIGURE 2-8 (A): JOINTLY CDF OF FLOOD PEAK AND VOLUME 18
FIGURE 2-8 (B): JOINTLY RETURN PERIOD OF FLOOD PEAK AND VOLUME 19
FIGURE 2-9 (A): CONDITIONAL PROBABILITY OF FLOOD VOLUME GIVEN FLOOD PEAKS 19
FIGURE 2-9 (B): CONDITIONAL RETURN PERIOD OF FLOOD VOLUME GIVEN FLOOD PEAKS 20
FIGURE 3.1 SCHEMATIC ILLUSTRATION OF DIFFERENT FLOOD HYDROGRAPH SHAPE: A) PRIOR-PEAK SHAPE; B) MIDPEAK SHAPE; C) POSTERIOR-PEAK SHAPE 22
FIGURE 3.2 BETA, GAMMA, LOG NORMAL AND WEIBULL PROBABILITY DISTRIBUTIONS SHAPES 26
FIGURE 3.3 SCHEMATIC ILLUSTRATION OF CALCULATION OF CENTROID OF FLOOD HYDROGRAPH 27
FIGURE 3.4 FLOWCHART OF DESIGN FLOOD HYDROGRAPH PROCEDURE 29
FIGURE 3.5 SCHEMATIC ILLUSTRATION OF SEPARATION OF SURFACE RUNOFF FROM BASE FLOW 30
FIGURE 3.6 SCHEMATIC ILLUSTRATION OF A) A FLOOD HYDROGRAPH; B) MODIFIED FLOOD-HYDROGRAPH 30
FIGURE 3.7 INFLOW TIME SERIES OF CAMANCHE/PARDEE RESERVOIR AND 32
TYPICAL FLOOD HYDROGRAPH (DWR, 2008) 32
FIGURE 3.8 (A-B) ORIGINAL AND MODIFIED HYDROGRAPHS OF FLOOD EVENTS IN MOKELUMNE RIVER 33
TABLE 3.2 PERFORMANCE INDICES OBTAINED FROM FITTING THE PDF FUNCTIONS WITH OBSERVED HYDROGRAPHS 34
FIGURE 3.9 (A-K) COMPARISON OF OBSERVED AND FITTED HYDROGRAPH FOR DIFFERENT FLOOD EVENTS 34
FIGURE 3.9 (A-K) COMPARISON OF OBSERVED AND FITTED HYDROGRAPH FOR DIFFERENT FLOOD EVENTS 35
FIGURE 3.9 (A-K) COMPARISON OF OBSERVED AND FITTED HYDROGRAPH FOR DIFFERENT FLOOD EVENTS 36
FIGURE 3.10 (A-B) DISTRIBUTIONS OF SHAPE MEAN AND STANDARD DEVIATION ON NORMAL PAPER 37
FIGURE 3.11 (A-B) COMPARISON OF DESIGNED FLOOD HYDROGRAPHS OF 10, 20, 50, 100 AND 200 YEAR 38
FIGURE 4.1 TRADITIONAL METHOD FOR ESTIMATING REGULATED FREQUENCY CURVE (USACE, 1996) 43
FIGURE 4.2 MODIFIED METHOD FOR ESTIMATING REGULATED FREQUENCY CURVE 44
FIGURE 4.3 SEPARATING FLOOD PULSES BY BASE FLOW FROM FLOW TIME SERIES 46
[MAIDMENT, ET AL. 1993] 46
FIGURE 4.4 A RESERVOIR STORAGE ALLOCATION [FORD, D. 1990] 49
FIGURE 4.5 TYPICAL RESERVOIR OPERATION FOR INFLOW/OUTFLOW RELATIONSHIPS [USACE, 1993] 52
FIGURE 4.6 REGULATED FLOOD FLOW FREQUENCY ANALYSIS FLOW CHART 53
FIGURE 4.7 CAMANCHE/PARDEE RESERVOIR OPERATION RULE FOR FLOOD CONTROL [USACE, 2006] 54
FIGURE 4.8 RESERVOIR REALLOCATION (BASED ON USACE 2006) 55
FIGURE 4.9A FLOOD PULSE VOLUME FREQUENCY UNDER DIFFERENT BASE FLOWS (A: Q_B=2,000 CFS; B: Q_B=3,000 CFS; C: Q_B=4,000 CFS AND D: Q_B=5,000 CFS) 56

Figure 4.9b Flood pulse duration frequency under different base flows (A: Q_B=2,000 CFS; B: Q_B=3,000 CFS; C: Q_B=4,000 CFS and D: Q_B=5,000 CFS)....57
Figure 4.9c Flood pulse flow (V/D) frequency under different base flows (A: Q_B=2,000 CFS; B: Q_B=3,000 CFS; C: Q_B=4,000 CFS and D: Q_B=5,000 CFS)....58
Figure 4.11 Unregulated flood flow histograms with base flow 2,000 CFS....59
Figure 4.12 Illustration of unregulated-regulated flow transform curves at Camanche/Pardee Reservoirs for various storages....62
Figure 4.13 Regulated flow density plots with/without uncertainty comparison....63
Figure 4.14 Regulated flow histograms at three storage size options....64
Figure 4.15a Flood pulse volume (V) frequency under different duration (A: D = 3 days; B: D = 5 days; C: D = 10 days and D: D = 20 days)....65
Figure 4.15a Flood pulse volume (V) frequency under different duration (A: D = 3 days; B: D = 5 days; C: D = 10 days and D: D = 20 days)....66
Figure 4.15b Flood pulse inflow (V/D) frequency under different duration (A: D = 3 days; B: D = 5 days; C: D = 10 days and D: D = 20 days)....66
Figure 4.15b Flood pulse inflow (V/D) frequency under different duration (A: D = 3 days; B: D = 5 days; C: D = 10 days and D: D = 20 days)....66
Figure 4.15b Flood pulse inflow (V/D) frequency under different duration (A: D = 3 days; B: D = 5 days; C: D = 10 days and D: D = 20 days)....67
Figure 4.16 Unregulated flood flow histograms with CT 10 days....67
Figure 4.17 Regulated flow histograms at three storage size options with USACE approach....68
Figure 4.18 Regulated flow histograms at three storage size options with two approaches....69
Figure 5.1 Common levee failure mechanisms [Wood, 1977]....74
Figure 5.2 Levee failure mechanisms....76
Figure 5.2 Levee erosion process sketch (W_E is effective levee width, $Slope_1$ is levee waterside slope, $Slope_2$ is levee foundation slope, WSE is water surface elevation)....83
Figure 5.3 Flow chart of computation of overall levee failure....86
Figure 5.4 A simple levee/channel system sketch (Not to scale)....86
Figure 5.6 Monte Carlo results with 100, 1,000 and 10,000 realization....89
Figure 5.7 Lower Mokelumne River and levee system (from USGS and Robinson and Bryon, 2006)....89
Figure 5.8 One typical levee cross section in Lower Mokelumne River (Robinson and Bryon, 2006)....90
Figure 5.9 Failure probabilities for flood storage capacity in Camanche/Pardee Reservoirs....91

LIST OF TABLES

TABLE 2.1: FLOOD PEAK Q, VOLUME V AND DURATION D (INFLOW OF PARDEE RESERVOIR) 12
TABLE 2.2: STATISTICS OF TRANSFORMED FLOOD DATA 12
TABLE 2.3: CORRELATIONS OF TRANSFORMED FLOOD DATA 14
TABLE 3.1: FLOOD DURATION D, VOLUME V AND PEAKS Q IN MOKELUMNE RIVER 32
TABLE 4.1 STATISTICS OF CANDIDATE DISTRIBUTIONS FOR MODIFIED UNREGULATED FLOW WITH BASE FLOW APPROACH 59
TABLE 4.2 THE DETERMINISTIC RELATIONSHIP BETWEEN INFLOW AND REGULATED FLOWS 60
TABLE 4.3 SUMMARY OF VARIANCE COEFFICIENTS (ERROR PART OF E) 60
TABLE 4.5 STATISTICS OF CANDIDATE DISTRIBUTIONS FOR UNREGULATED FLOW WITH USACE APPROACH 67
TABLE 4.6 REGULATED FLOWS STATISTICS WITH USACE APPROACH 68
TABLE 5.1 STRENGTH OF LEVEE SOIL CHARACTERISTICS [BRIAUD ET AL., 2001A, B; HANSON AND SIMON 2003] 84
TABLE 5.2 LEVEE GEOMETRIC VARIABLE AND DISTRIBUTION 87
TABLE 5.3 CRITICAL SOIL STRENGTH AND DISTRIBUTION 87
TABLE 5.4 SUMMARY OF HYDRAULIC VARIABLES FOR EXAMPLE 87
TABLE 5.5 LEVEE GEOMETRIC VARIABLE AND DISTRIBUTION IN LOWER MOKELUMNE RIVER LEVEE 90
TABLE 5.6 LOWER MOKELUMNE RIVER LEVEE FAILURE ANALYSIS SUMMARY 91

CHAPTER 1 INTRODUCTION

Among the five disasters in a nation, flood is the most dangerous one. However, as long as you conquer it, you can rule the whole nation.

- Guanzi, Chinese Philosopher, 400 BC

1.1 MOTIVATION

Reservoir systems have long been built for flood protection, water conservation, hydropower and recreation [Wurbs, 1993]. The development of environmental consciousness has increased concerns for environmental flows downstream of dams in recent decades [Richter, 2006]. Also with climate change related uncertainties, reservoir re-operation to balance hydropower, flood pool, water storage and environmental purposes is now commonly discussed. A common re-operation practice is to reduce existing flood storage to gain hydropower, water supply and ecological benefits. However, this reservoir re-operation will increase downstream flood and levee failure frequency.

Reservoir operation and re-operation requires accurate inflow frequency analysis which involves multivariate characteristics of flood peaks, volumes and duration. A complete understanding of flood events involves the joint probabilistic behaviors of these correlated variables. Flood hydrograph design is a key component for designing flood control reservoir operation. Numerous methods exist to develop flood hydrographs. Using probability density functions (PDFs) to fit the shapes of flood hydrographs has drawn attention recently due to improvements in statistical techniques, including algorithms for fitting.

Reservoir re-operation can change regulated/unregulated flood flows dramatically. Regulated versus unregulated flow relationships are a key aspect of reservoir flood control system. Reservoir re-operation with less flood storage will change not only the parameters of regulated flow distributions but also the distributions themselves. However, due to the complexity of reservoir systems and uncertainties in operation, few articles discuss this issue.

Urbanization behind levees and climate change increasing hydrological extremes have increased attention to levee failure. Levee failure can have several mechanisms. Among them, overtopping and erosion failure analysis are performed separately in water resources and geotechnical engineering. This is due to the different approaches to uncertainty in hydrology and geotechnical engineering. Few papers discuss levee failure with both hydrologic and geotechnical uncertainties.

Statistics deals with methods for drawing inferences about the properties of a population based on the properties of a sample from that population. Statistics has long been widely used in water resources for decades [Hirsch, R., et al, 1993]. With efficient computation algorithms and developed statistical packages, statistical techniques can be applied more widely to water resources engineering and can help us understand inflow and develop design hydrographs.

This dissertation presents a vertical framework to address these issues. Figure 1.1 illustrates this framework. It includes inflow analysis, regulated/unregulated flow relation curves and levee failure analysis through reservoir re-operation.

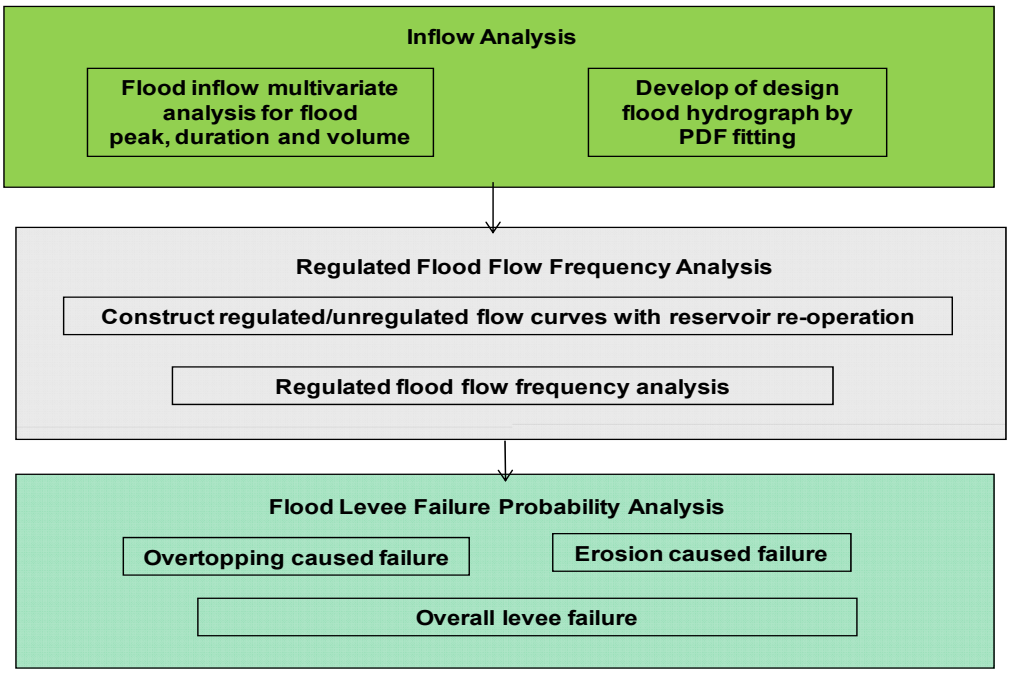


FIGURE 1.1 FLOW CHART OF RESERVOIR RE-OPERATION, RISK AND LEVEE FAILURE ANALYSIS

1.2 ORGANIZATION

This chapter reviews the organization of the dissertation. Chapter 2 presents a procedure for using the bivariate normal distribution to describe the joint distributions of correlated flood peaks and volumes, and correlated flood volumes and durations. Joint distributions, conditional distributions, and the associated return periods of these random variables can be readily derived from their marginal distributions. The theoretical distributions show a good fit to observed ones. The return periods will be used for risk analysis of flood storage options.

Chapter 3 introduces a framework to develop design flood hydrographs or reservoir reoperation. Three steps are presented. 1) Flood hydrograph separation and modification: Typical flood hydrographs were separated, selected and converted to dimensionless ones; 2) PDF fitting and selection: Beta, Gamma. Lognormal and Weibull distribution forms were compared to fit modified hydrographs based on goodness of fit criteria including RMSE and coefficients of determinations. 3) Development of design flood hydrographs: The design shape variables were determined from frequency analysis and finally, the design flood hydrographs including 10-, 20-, 50-, 100- and 200-year return periods were derived from the combinations of hydrograph shape, flood volume and durations. Gamma PDF form seems most suitable to fit the application site's flood events.

Chapter 4 introduces a framework to estimate regulated flow frequency for a reservoir's flood storage re-allocation. Three main steps including unregulated flow frequency analysis, unregulated/regulated flow transformation and regulated flow frequency estimation are presented. The main contributions include separating flood pulses from daily inflow time series by base flow criteria, modification of unregulated flow calculations, and fitting unregulated and regulated flow to appropriate probability distributions. Unregulated versus regulated flow curves are found using USACE's ResSim software.

Chapter 5 introduces a framework to assess levee failure probability from overtopping and erosion incorporating with uncertainties in hydrologic, hydraulic and geotechnical factors. Overtopping and erosion failure probability analysis are usually performed separately in water resources and geotechnical engineering. This chapter presents a more comprehensive calculation combining these two failure mechanisms. Also load-resistance interference failure probability analysis is introduced to consider overtopping between flood magnitude and levee capacity and erosion failure between velocity and soil strength. Both analyses are performed by Monte Carlo simulation to estimate overall levee failure

probability. Failure probability can be very sensitive to geotechnical variables and less sensitive to reservoirs reoperation.

The Lower Mokelumne River levee system below Camanche/Pardee reservoirs in Northern California is the application site. Chapter 6 summaries and concludes the findings throughout the dissertation.

REFERENCES

Wurbs, R.,(1993) “Reservoir-system Simulation and Optimization Models”. J. Water Resources Planning and Management.Vol.119, No.4, July/August, 1993.

Richter, B. D., Warner, A. Meyer, J. and Lutz, K.(2006), "A collaborative and adaptive process for developing environmental flow recommendations, River Research and Application, 22: 297-318.

Hirsch, RM., Helsel, D.R., Cohn, T.A. and Gilroy, E.J. (1993). “Statistical Analysis of Hydrologic Data” In Handbook of Hydrology, Maidment, D. (ed) from 17-1 to 17-55. McGraw Hill: New York.

Venables, W., Smith, D. and the R Development Core Team (2010) “An Introduction to R: Notes on R: A Programming Environmental for Data Analysis and Graphics Version 2.12.0”, http://cran.r-project.org/doc/manuals/R-intro.pdf

CHAPTER 2 MULTIVARIATE FLOOD EPISODE ANALYSIS OF RESERVOIR INFLOW

After Gun, his father was killed due to failure to control flood, Da Yu regarded flood as his teacher. To learn from floods, he spent 13 years to observe rivers and have enough knowledge of flood characteristics.

- Huai Nan Zi, Han Dynasty Philosopher,135 BC

SUMMARY

Reservoir operation for flood control requires accurate inflow frequency analysis which involves multivariate characteristics of flood peaks, volumes and duration. A complete understanding of flood events involves the joint probabilistic behaviors of these correlated variables. Raw flow data are commonly transformed for flood frequency analysis, often to a form of lognormal distribution. The multivariate distribution is important for analyzing floods. This chapter presents a procedure for using the bivariate normal distribution to describe the joint distributions of correlated flood peaks and volumes, and correlated flood volumes and durations. Joint distributions, conditional distributions, and the associated return periods of these random variables can be readily derived from their marginal distributions. The theoretical procedure is applied to modified unregulated inflow for the Mokelumne River for reservoir re-operation. The theoretical distributions show a good fit to observations. The return periods will be used for risk analysis of flood storage space operation.

2.1 INTRODUCTION

Flood severity is often a function not only of the flood flow peak, but also volume and duration characteristics of the flood. In reservoir operation for floods, inflow peaks, volumes and duration are all important. In the past, however, flood-frequency analysis has often concentrated only on flow peaks volumes. Reviews of single-variable flood frequency analysis are available [Cunnane,1987; Bobee and Rasmussen, 1994]. However, a flood event is multivariate, characterized by its peak, volume and duration, which may be mutually correlated. Flood-peak frequency analysis provides a limited assessment of a flood episode, whereas reservoir operation requires more information concerning the flood event (flood peak, flood volume and flood duration).

Some meaningful attempts have been made to address this topic in past decades [Ashkar, 1980; Correia, 1987; Sackl and Bergmann, 1987; Krstanovic and Singh, 1987; Kelly and Krzysztofowicz, 1997; Goel et al 1998; Yue, 2000]. Ashkar [1980] considered a flood as a multivariate event and derived relationships between flood peak, duration and volume. Correia [1987] deduced the joint distribution of flood peaks and durations using the partial duration series method (PDS) on the basis of assumptions that (i) both flood peaks and durations are exponentially distributed; and (ii) the conditional distribution of flood peaks for given flood durations is normal. Krstanovic and Singh (1987) derived multivariate Gaussian and exponential distributions using the principle of maximum entropy (POME) and used these distributions to describe the joint distribution of flood peaks and volumes. Kelly and Krzysztofowicz proposed a bivariate meta-Gaussian model for hydrological frequency analysis [Kelly and Krzysztofowicz, 1997]. The joint distribution of flood peaks and volume have been represented by the bivariate normal distribution [Sackl and Bergmann, 1987, Goel et al., 1998 and Yue,1999]. In practice, extreme events such as flood peak and flood volume maybe represented by the Gumbel distribution (EV1 distribution) [Gumbel 1958; Todorovic 1978; Castillo 1988; Watt et al. 1989]. However, these models have mainly remained their theoretical developments and seldom succeeded in resolving practical problems in the field of hydrological frequency analysis. Yue developed Gumbel logistic model to represent the joint distribution of flood peak and volume in the province of Quebec, Canada. The results show the correlation coefficient between two random variables must be in the range: $0 < \rho < 2/3$. However, a large number of hydrological extreme events may be closely correlated and the correlation between them may be greater than 213. In such cases, the Gumbel mixed model is no longer valid [Yue 1999].

Usually, hydrologic events such as flood peak and flood volume are positively skewed and may follow the lognormal distribution [Chow, 1954; Sangal and Biswas,1970; Watt et al., 1989; Stedinger et al, 1993]. It will be of great interest to hydrological engineers to use the bivariate normal distribution after transforming raw data to analyse the joint probability distribution of two correlated random variables with marginals that are normally distributed. This chapter presents a procedure for using the bivariate normal distribution to represent joint distributions of flood peaks and flood volumes as well as flood volumes and durations. On the basis of the marginal distributions of these random variables, the joint distributions, the conditional distributions and the associated return periods could be derived.

The organization of this chapter is as followed. The next section provides basic concepts including flood event characteristics, the Box-Cox data transformation, and bivariate normal distribution formulas. In the next section the above procedure is verified using Mokelumne River modified unregulated inflow flood data in northern California, U.S.A. The final section summarizes results and application.

2.2 CONCEPTS DEVELOPMENT

This section presents the flood events characteristics, common data transforming methods and the bivariate normal distribution.

2.1.1 Characteristics of flood events

Flood Hydrograph

The most significant flood characteristics are the flood peak (Q), flood volume (V) and flood Duration (D) as illustrated in Figure 2.1. Determination of flood duration involves establishing times of the start and end of flood runoff. Generally time boundaries of a flood are marked by a rise in stage and discharge from base flow (start of flood runoff) and a return to base flow (end of flood runoff). In this study, start of surface runoff is usually marked by the abrupt rise of the hydrograph. The end of flood runoff can be identified by the flattening of the hydrograph recession limb. As the characteristics of surface runoff recession differ from the base flow, there is a significant change in the slope of the hydrograph as the transition occurs from surface runoff to base flow. On the basis of these criteria, the estimation of the start date (SD_i) and end date (ED_i) of flood runoff for the ith year can be carried out, and the flood duration series (D_i) is constructed as

$$D_i = (ED_i - SD_i) \qquad (2.1)$$

The flood volume series can be constructed using the following formula

$$V_i = \sum_{j=SD_i}^{ED_i} (q_{ij} - 1/2(q_{is} + q_{ie})) \qquad (2.2)$$

where q_{ij} is the j th day observed daily stream flow value for the i th year; q_{is} and q_{ie} are observed daily stream flow values on the start date and end date of flood runoff for the i th year, respectively. The flood peak Q_i is the maximum surface flow and is given by

$$Q_i = \max(q_i - q_b) \qquad (2.3)$$

where q_b is base flow value and q_i is the daily flow.

Annual Maximum Series (AMS) and Partial Duration Series (PDS)

A storm time series can be constructed using the annual maximum series (AMS) approach, or the partial duration series (PDS) (or peak over threshold (POT)) approach [Stedinger et al., 1993]. An AMS is constructed by selecting the annual maximum value of each year, i.e., only one event per year is retained. This leads to events that are generally independently and identically distributed. The PDS contains of all values that exceed a specified threshold. The main advantage of the PDS approach is that it is not confined to only one event per year, allowing additional large events to be considered. However, the key unresolved problem of the PDS is selecting appropriate thresholds [Cunnane, 1987; Valadares Tavares and Evaristo Da Silva, Wang, and Rasmussen et al.]. Wang finds AMS and PDS similar for a long-term time series [Langbein, 1949]. In this chapter, the AMS approach is employed to provide a storm peak and the corresponding storm volumes and durations.

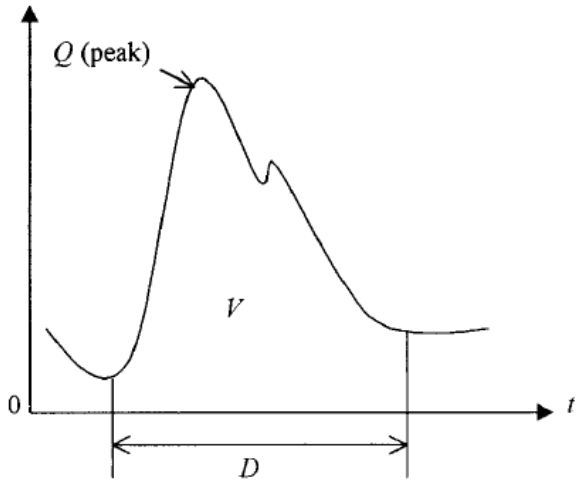


FIGURE 2.1: CHARACTERISTICS VALUES OF A FLOOD EVENT

Separation of flood hydrograph

There are two common ways to separate flood hydrographs from time series: the specific duration method and base flow method. The specific duration method separates the hydrograph based on specified duration such as 3 days, 7 days or 30 days. All separated hydrographs have the same duration. The base flow method sets a specific flow and compares all hydrographs exceeding this threshold. All the separated hydrographs have same starting flow value and ending flow values. In this study, the base flow method was selected to extract the direct stream flow hydrographs. Many techniques are practiced to separate the base flow from the stream flow hydrographs. Three commonly used graphical techniques, are constant discharge, constant slope and concave methods. These techniques aim mainly to separate quick flow from slow flow for flood analysis and prediction [Pramanik et al, 2010]. Figure 2.4 illustrates of separation of surface runoff from base flow. For simplicity, the start and end of a flood hydrograph are connected by a straight line, and this straight line is considered as the shape of base flow.

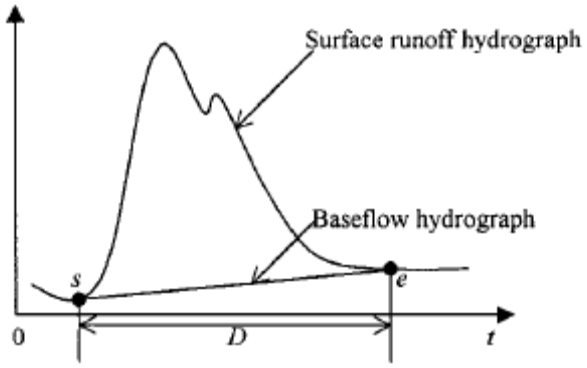


FIGURE 2.2 SCHEMATIC ILLUSTRATION OF SEPARATION OF SURFACE RUNOFF FROM BASEFLOW

2.1.2 Box-Cox Transformation

The Box-Cox transformation or the power transformation [Box and Cox, 1964] is applied to normalize sample data. It has been widely used in hydrology and is defined as

$$y_{\lambda i} = \frac{x_i^{\lambda}}{\lambda} \quad (\lambda \neq 0) \qquad (2.4)$$

$$y_{\lambda i} = \log x_i \quad (\lambda = 0) \qquad (2.5)$$

Where x_i are original sample data, $y_{\lambda i}$ are the transformed sample data, and λ is the transformation parameter.

If the transformed sample $y_{\lambda i}$ follows the normal distribution $N(\mu,\sigma^2)$, λ is set by maximum likelihood function methods. The original variable x_i can be readily obtained using the following back transformations:

$$x_i = (y_{\lambda i} + 1)^{1/\lambda} \quad (\lambda \neq 0) \qquad (2.6)$$

$$x_i = \exp(y_{\lambda i}) \quad (\lambda = 0) \qquad (2.7)$$

The transformation parameter is estimated using the maximum likelihood method. The log likelihood function is given as

$$\log(\lambda) = -\frac{n}{2}\log(2\pi) - \frac{n}{2} - \log\left[\sum_{i=1}^{n} \frac{(z_{\lambda i} - z_m)^2}{n}\right] \qquad (2.8)$$

Where

$$z_{\lambda i} = \frac{y_{\lambda i}}{x_{gm}^{\lambda-1}} \qquad (2.8a)$$

$$x_{gm} = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \qquad (2.8b)$$

$$z_m = \frac{1}{n}\sum_{i=1}^{n} z_{\lambda i} \qquad (2.8c)$$

The maximum likelihood estimation problem reduces to the minimization with respect to λ of

$$t_\lambda^2 = \sum_{i=1}^{n} \frac{(z_{\lambda i} - z_m)^2}{n} \qquad (2.9)$$

The parameter λ will be readily reached by trial-and-error.

2.1.3 Bivariate normal distribution

If two correlated continuous random variables X and Y are normally distributed with different parameters (mean and standard deviation) as follows: [DeGroot, M., Schervish, M., 2002]

$$f(x) = \frac{1}{\sigma_x\sqrt{2\pi}}\exp\left[-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right] \quad (-\infty < x < +\infty) \qquad (2.10)$$

$$f(y) = \frac{1}{\sigma_y\sqrt{2\pi}}\exp\left[-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right] \quad (-\infty < y < +\infty) \qquad (2.11)$$

then the joint distribution of these two variables can be represented by the bivariate normal distribution. The bivariate probability density function (pdf) is:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp\left(-\frac{q}{2}\right) \quad (-1 < \rho < +1) \qquad (2.12)$$

$$q = \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{x-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right] \qquad (2.13)$$

where μ_x, σ_x, μ_y and σ_y, are respectively the mean and standard deviation of X and Y, and are estimated using the method of moments (MOM); ρ is the product-moment correlation coefficient of X and Y, and is computed by:

$$\rho = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sigma_x\sigma_y} \qquad (2.14)$$

The conditional pdf of X given Y = y and pdf of Y given X=x are given as follows [Hogg and Craig, 1978]:

$$f(x|Y=y) = \frac{1}{\sigma_{x|y}\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_{x|y})^2}{2\sigma_{x|y}^2}\right] \quad (-\infty < x < +\infty) \qquad (2.15)$$

$$f(y|X=x) = \frac{1}{\sigma_{y|x}\sqrt{2\pi}} \exp\left[-\frac{(y-\mu_{y|x})^2}{2\sigma_{y|x}^2}\right] \quad (-\infty < y < +\infty) \qquad (2.16)$$

$$\mu_{x|y} = \mu_x + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_y) \qquad (2.17)$$

$$\sigma_{x|y} = \sigma_x\sqrt{1-\rho^2} \qquad (2.18)$$

Thus, the conditional distributions of X and Y are also normally distributed with different means and standard deviations.

$$\mu_{y|x} = \mu_y + \rho\frac{\sigma_y}{\sigma_x}(y-\mu_x) \qquad (2.19)$$

$$\sigma_{y|x} = \sigma_y\sqrt{1-\rho^2} \qquad (2.20)$$

In this study, bivariate normal distribution is used due to it's simplicity.

2.1.4 Frequency analysis for return periods determination

The return periods exceeding some values of the variables X and Y are presented as follows:

$$T_x = \frac{1}{1-F_x} \quad (F(x) = P_r[X \le x]) \qquad (2.21)$$

$$T_y = \frac{1}{1-F_y} \quad (F(y) = P_r[Y \le y]) \qquad (2.22)$$

Similarly, the joint return period Txy, of two variables x and y, the conditional return period Tx/y of X given y, and the conditional return period Ty\x of Y given X are respectively given as follows:

$$T_{xy} = \frac{1}{1-F(x,y)} \quad (F(x,y) = P_r[X \le x, Y \le y]) \qquad (2.23)$$

$$T_{x|y} = \frac{1}{1-F(x|y_0)} \quad (F(x,y) = P_r[X \le x \mid Y = y_0]) \qquad (2.24)$$

$$T_{y|x} = \frac{1}{1-F(y|x_0)} \quad (F(x,y) = P_r[Y \le y \mid X = x_0]) \qquad (2.25)$$

where F(*) is the cumulative distribution function (cdf). As the cdf of the normal distribution is not attainable, it was computed by numerically integrating the corresponding pdf f(*).

2.3 METHODS

In this section, the process of computing the frequency of flood peak, volume and duration for all observed hydrographs is presented. Figure 2.3 shows this procedure's flowchart. Three sections including flood hydrograph separation and variables calculation, multivariate analysis and goodness of fit statistics and development of design flood episodes are presented.

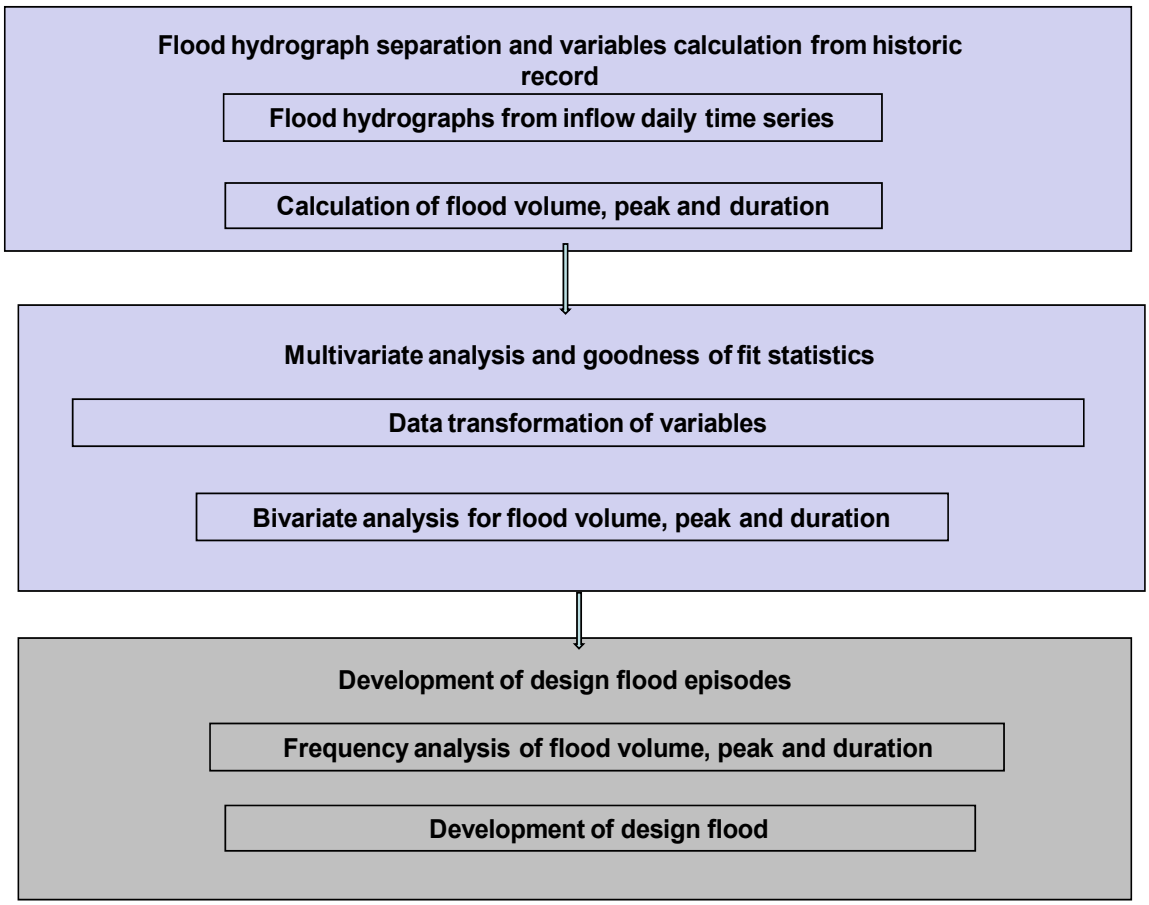


FIGURE 2.3 FLOW CHART OF FREQUENCY ANALYSIS OF OBSERVED FLOOD HYDROGRAPH

2.4 APPLICATION TO CAMANCHE/PARDEE RESERVOIR

To test the above methods, the Mokelumne River's historical flow data were used. The application results are presented below. In this basin, as with most basins larger than 500 km^2 in northern California, a mix of snow melt and rainfall usually causes the annual maximum flood, both in flow peak and volume. Using the annual maximum series approach, the joint probability distributions of flood peaks and volumes as well as flood volumes and durations are analyzed for reservoir re-operation flood risk analysis in Mokelumne River.

2.1.5 Site and inflow data description

Two reservoirs, Pardee and Camanche, are on the Mokelumne River, a major tributary of San Joaquin River in California. Ten major flood events have occurred on this river in the past 50 years, four in the past 20 years. These four events have accounted for an average flood damage value of $4 million per event [DWR, 2006]. The reservoirs and watershed are described below.

Camanche/Pardee Reservoirs

Camanche reservoir is on the Mokelumne River near Jackson, California, approximately 10 miles downstream from Pardee reservoir. Camanche reservoir provides storage for flood control, irrigation, power, recreation, and downstream fishery needs. The combination of Camanche and Pardee reservoirs provide maximum flood control space reservation of 200,000 acre-feet during the winter. Camanche reservoir was completed in 1963 and Pardee was constructed in 1927.

Mokelumne River Watershed Description

The Mokelumne River watershed covers about 920 square miles of mountainous to valley floor terrain. Elevations range from a peak of above 8,800 feet msl to slightly below sea level in the vicinity of the Delta. The Mokelumne River is highly regulated by reservoirs for waters supply and power generation, with Camanche Reservoir providing flood control capacity for the lower watershed.

The upper watershed supports approximately 272,860 acres of conifer forest, or 46% of the land area. The uppermost watershed is in the alpine region of the Sierra Nevada, above the main forested areas of the watershed. The alpine region is characterized by granite peaks, lake basins, and other rock structures carved by past glaciation [Storer and Usinger 1963]. Below the alpine region and down to the foothill region, the watershed is dominated by the forested canyons of the north, middle, and south forks of the Mokelumne River. The middle watershed, generally from Highway 49 to Camanche Dam, supports about 30,000 acres of oak woodlands, which make up approximately 5% of the watershed. This portion of the watershed narrows to the width of the main stem river canyon, generally less than 2 miles wide as it enters Pardee and Camanche reservoirs and flows to the valley floor. Both reservoirs are owned and operated by EBMUD. The lowest part of the watershed, below Camanche Dam, includes over 70,800 acres of cropland and nearly 60,300 acres of orchards and vineyards. This area also includes the communities of Clements, Lockeford, Lodi, and Woodbridge [Robinson and Bryon, 2006].

Inflow Description

Modified unregulated daily inflow to Pardee and Camanche reservoirs are used. The inflow gauge is just above Pardee Reservoir (USGS 11319500). Several reservoirs in the drainage basin lie above Pardee Reservoir for irrigation, power, and/or water supply. The capacities are quite small except Lower Bear Reservoir and Salt Spring Reservoir with capacities of 52,000 ac-ft and 142,000 ac-ft respectively. The unregulated modified inflow to Pardee Reservoir was based on the gauge data with the consideration of the Salt Spring and Lower Bear Reservoirs. The final modified daily inflow is shown in Figure 2.4.

A normal distribution is assumed as the appropriate distribution for all three transformed variables, i.e, log(peak), log(volume) and log(duration). To test the goodness of fit of the normal distribution, R package's QQNORM/QQLINE code is applied [R Core Team, 2010]. Figure 2-5(a) to 2-5(c) show the comparisons of transformed inflow data with normal distribution and theoretical normal distribution. In this study, equation 2.5 from Box-Cox methods is used. The points on the plots are the transformed peaks, duration and volume data with normal distribution. A 45-degree reference line representing theoretical normal distribution is also plotted. If the transformed data follows the chosen distribution, i.e. normal distribution, the points should fall approximately along this reference line. The greater the departure from this reference line, the greater the evidence that the data have come from a non-normal distribution. From the plots, the transformed peaks and volumes points are closer to the reference line than the transformed duration points. Therefore, the transformed peaks and volumes follow a normal distribution.

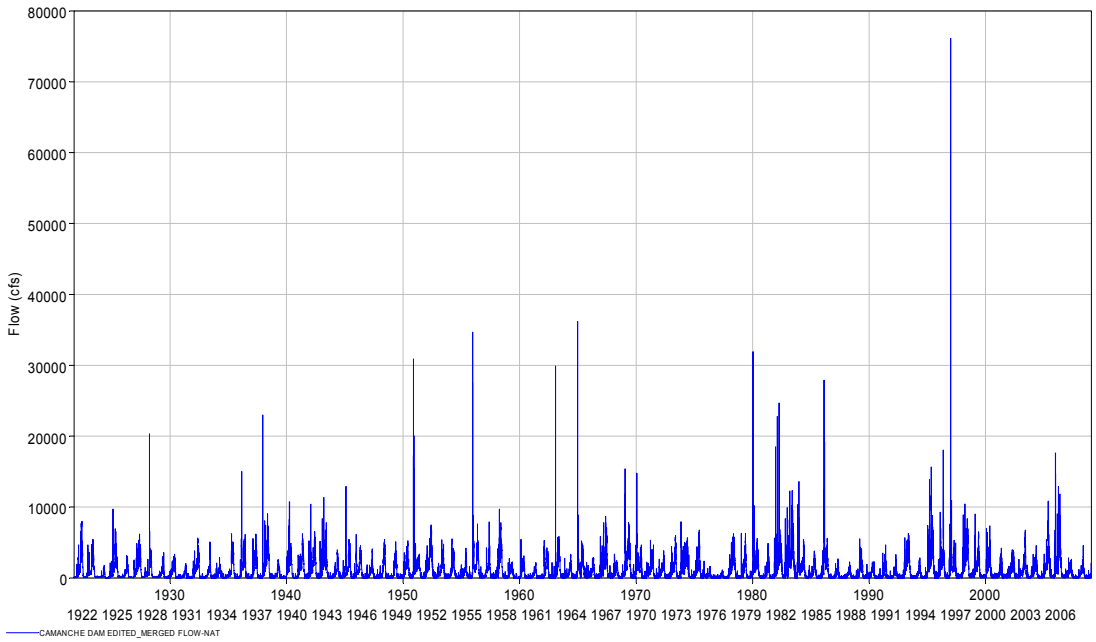


FIGURE 2.4 DAILY UNREGULATED INFLOW TO PARDEE RESERVOIR (1921- 2009, USGS AND DWR, LATITUDE 38.1061°N, LONGITUDE 121.5711°W.)

2.1.6 Flood hydrograph separation and peak, volume and duration calculation

AMS method was selected to analyze flood hydrograph using a base flow separation technique. Base flow was selected at 5,000 CFS because it is the capacity of the downstream channel [USACE, 1983]. Table 2.1 shows the results of peak, volume and duration of floods from year 1924 to 2009.

2.1.7 Box-Cox transformation and normalization of flood data

From the Box-Cox transformation theory, λ=0 in Equation (1) is selected to transform the raw data for peaks, volume and duration. Therefore, the transformed data y_i will replace raw data x_i for analysis.

$$y_i = \log(x_i) \qquad (2.26)$$

Table 2.2 shows the statistics of log transformed data y_i including mean (μ), standard deviation (σ), maximum (max), minimum (min) and coefficient of skew (C_s). From the table, flood peaks have the lowest standard deviation while flood volumes have the highest deviation. For skew coefficient, flood peaks have positive skew while flood duration and volume have negative skew.

Table 2.1: Flood peak Q, volume V and duration D (Inflow of Pardee Reservoir)

Year	Q (CFS)	V (taf)	D (days)	Year	Q (CFS)	V (taf)	D (days)
1924	5,430	236	51	1967	2,845	100	34
1925	1,770	4	1	1968	8,651	628	82
1926	9,700	52	10	1969	3,352	47	12
1927	3,100	27	9	1970	15,415	174	18
1928	6,160	233	58	1971	14,756	243	30
1929	20,300	237	30	1972	5,335	45	9
1930	3,530	21	5	1973	3,818	43	10
1931	3,319	69	18	1974	5,987	378	77
1932	2,022	15	6	1975	7,905	47	9
1933	5,616	191	31	1976	6,733	399	74
1934	5,105	195	34	1977	2,355	10	4
1935	2,858	17	5	1978	1,122	2	1
1936	6,214	519	82	1979	6,211	244	56
1937	15,034	91	11	1980	6,292	326	70
1938	6,194	330	1	1981	31,924	206	16
1939	22,970	97	8	1982	4,873	58	13
1940	2,570	44	13	1983	24,642	706	83
1941	10,740	252	43	1984	12,304	907	131
1942	6,214	436	74	1985	13,559	137	15
1943	10,397	78	14	1986	3,766	82	27
1944	11,380	164	27	1987	27,878	596	61
1945	3,964	110	27	1988	2,689	12	4
1946	12,891	77	11	1989	2,255	9	4
1947	6,133	71	14	1990	5,492	58	11
1948	4,046	65	15	1991	2,273	9	4
1949	5,415	237	42	1992	4,654	20	5
1950	5,108	246	62	1993	3,298	10	4
1951	5,227	237	51	1994	6,262	482	88
1952	30,862	254	14	1995	2,790	31	9
1953	7,447	873	163	1996	15,637	865	100
1954	5,342	116	27	1997	18,015	366	73
1955	5,491	33	7	1998	76,137	370	28
1956	4,191	102	22	1999	10,417	135	23
1957	34,657	258	20	2000	9,021	59	11
1958	7,874	187	31	2001	7,319	114	32
1959	9,679	82	16	2002	4,190	72	17
1960	2,719	19	6	2003	3,952	42	10
1961	5,426	27	6	2004	6,745	256	49
1962	2,143	20	7	2005	4,532	67	22
1963	5,241	22	5	2006	10,819	503	82
1964	29,861	158	14	2007	17,633	143	18
1965	3,321	100	24	2008	2,791	12	7
1966	36,173	291	17	2009	4,581	97	29

Table 2.2: Statistics of transformed flood data

	μ	σ	C_s	max	min
Q (log(CFS))	8.8002	0.806	0.5986	11.2403	7.0223
D (log(days))	2.6535	1.031	-0.1776	4.6052	0
V (log(TAF))	4.5596	1.3158	-0.5175	6.8102	0.7982

The conclusion above can be examined by Shapiro-Wilk test. Shapiro-Wilk test is one of the most powerful normality tests, especially for small samples. Normality is tested by matching two alternative variance estimates: a non-parametric estimator got by a linear combination of ordered sample values and the usual parametric estimator [DeGroot, M., Schervish, M., 2002]. The statement performing Shapiro-Wilk test is shapiro.test() and it supplies the p value. The p-value is higher than significance levels usually used to test statistical hypotheses, we accept the null hypothesis that the sample data is from a normal

distribution. From the R package, the p-values for transformed peak, duration and volume are 0.03372, 0.3234, and 0.05825, respectively. The transformed peaks and volumes follow the normal distribution at the 95% confidence level.

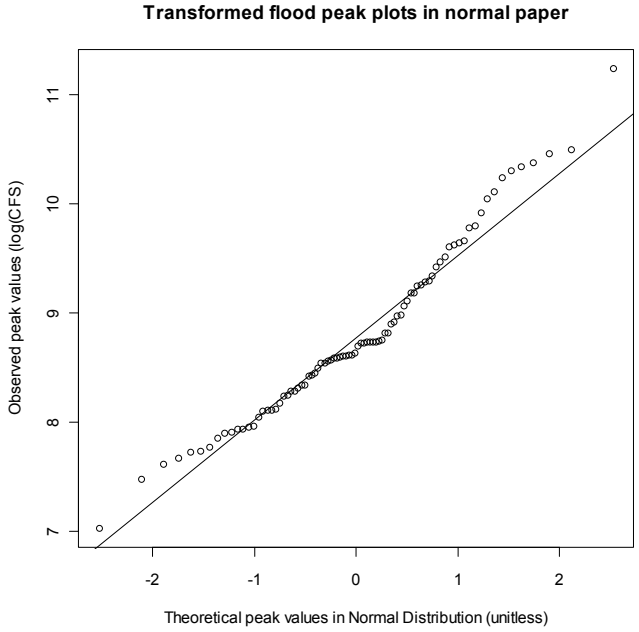


FIG 2-5(A): TRANSFORMED FLOOD PEAK(Q) ON NORMAL PAPER

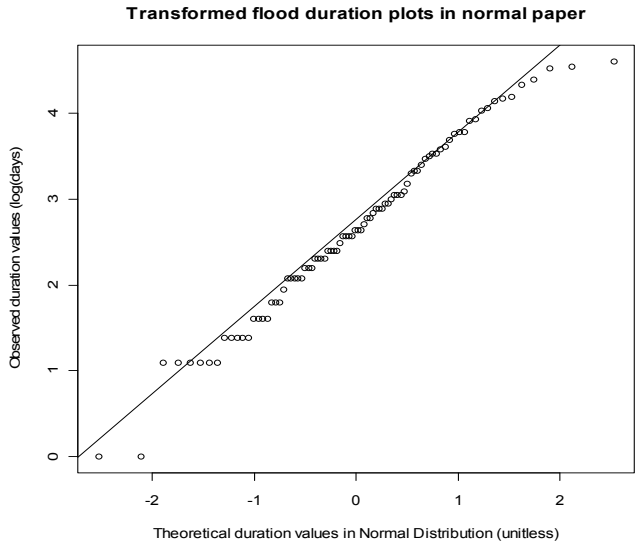


FIG 2-5(B): TRANSFORMED FLOOD DURATION (D) ON NORMAL PAPER

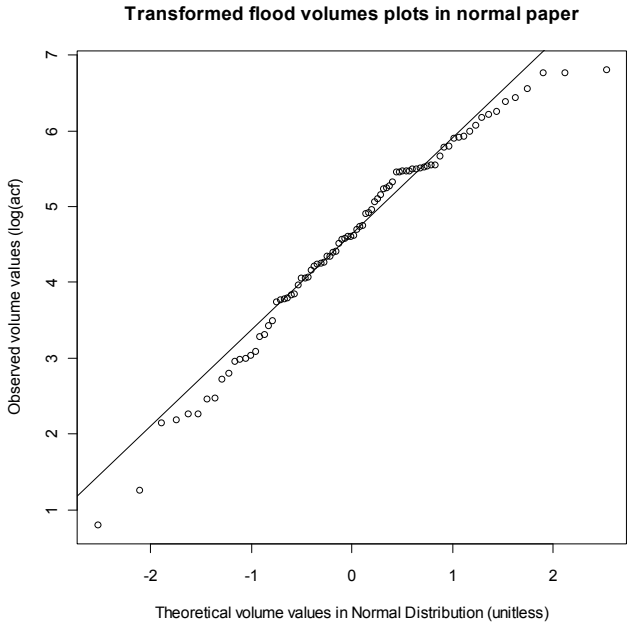


FIG 2-5 (C): TRANSFORMED FLOOD VOLUME (V) ON NORMAL PAPER

2.1.8 Association between flood peaks (Q), flood duration (D) and flood volume (V)

From the physical point of view, flood peak and duration are the least correlated of the three variables and we assume they are mutually independent. Thus we can explore the bivariate normal model to represent the joint behaviors of a flood episode, i.e. to analyse the different two-way combinations of the flood event: the joint distribution of flood peaks and volumes and the joint distribution of flood volumes and durations. The correlation coefficients between flood peaks and volumes, and between flood durations and volumes are estimated using Equation (5) and presented in Table 2.3. From the table, the volume and duration is closely correlated. The correlation between the volume and peak is greater 2/3. Thus, the bivariate normal distribution is appropriate.

Table 2.3: Correlations of transformed flood data

	Peak	Duration	Volume
Peak	1	0.4679	0.67798
Duration	0.5679	1	0.9577
Volume	0.67798	0.9577	1

2.1.9 Statistics of the joint distribution of flood peaks (Q) and volumes (V)

Validity of the bivariate normal model.

Observed joint probabilities are computed based on the same principle as in the case of a single variable. A two dimensional table is first constructed in which the variables Q and V are arranged in ascending order. The element in row i and column j of the table is defined as the joint frequency function of the two random variables and is estimated by

$$f(q_i, v_j) = P(Q = q, V = v) = \frac{n}{N+1} \quad (2.27)$$

Where N is the total number of observations (N = 86) and n is the number of occurrences of the combinations of q_i and v_j. The joint cumulative frequency (non exceedance joint probability) is then

$$f(q, v) = P(Q \le q, V \le v) = \sum_{m=1}^{i} \sum_{l=1}^{j} f(q_m, v_l) \quad (2.28)$$

Theoretical joint probabilities of the real occurrence combinations of q_i and v_i are estimated by numerically integrating Equation 13. The observed and theoretical joint probabilities are presented in Figure 2-6. In Figure 2-6 the solid-line represents the theoretical joint probabilities of flood peaks and volumes, which are arranged in ascending order, and the corresponding observed joint probabilities are expressed by the plus sign. The x axis is the corresponding order number of a combination of q_i and v_j. The theoretical joint probabilities fit the observed ones well.

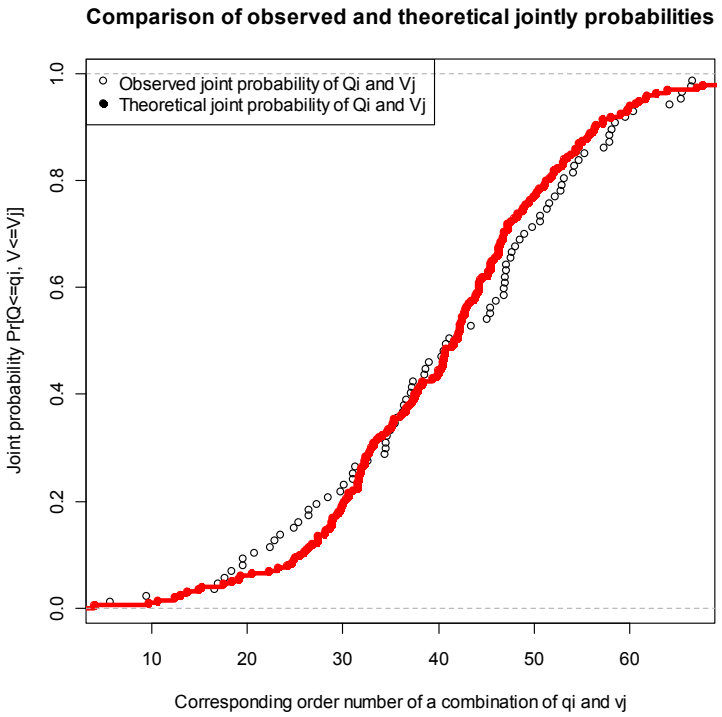


FIGURE 2-6 COMPARISON OF OBSERVED AND THEORETICAL PROBABILITIES OF FLOOD PEAKS AND VOLUMES

Joint PDF, CDF and return period of Q and V

The joint PDF, CDF and return period of flood peaks and volumes corresponding to given flood peak and volume values are computed. The joint PDFs are displayed in Figure 2-7 (a-c) with different view directions. R package is employed to perform the calculating and plotting. For ease of understanding, the joint CDF and return period given specific flood peak values is also plotted and illustrated in Figure 2-8 (a-b).

Conditional return periods

The conditional return period $T_{v|q}$ of flood volume V given flood peaks Q is computed using Equation (7) and is shown in Figure 2-9 (a). Similarly, the conditional return period $T_{v|q}$ of V given Q is represented in Figure 2-9 (b). Figures 2-9(a) and (b) indicate that the proposed method can help solve several problems in hydrological engineering design and management for which single variable flood frequency analysis cannot provide answers. For example, given a flood event return period, it is possible to obtain various occurrence combinations of flood peaks and volumes, and vice versa. These various scenarios can be

useful in the analysis and assessment of the probability of hydrological problems such as spillway design and flood control. The method proposed also allows one to obtain information concerning the occurrence probabilities of flood volumes under the condition that a given flood peak occurs, and vice versa.

Statistics of the joint distribution of flood duration (D) and volumes (V)

Similarly, observed and theoretical joint probabilities of flood durations and volumes are computed. There are no major differences between observed and theoretical probabilities. Therefore, the model appears suitable for representing the joint distribution of flood durations and volumes. The joint PDF, CDF and return period of flood durations (D) and volumes (V) are computed as outlined in the previous section. These results confirm the usefulness of the proposed model. Aside from flood peak, the duration and the volume of a given flood are the two main flood characteristics which affect flood damages. Results from the proposed model can then be used to calculate flood damages for pre-flood or post-flood studies.

This study selects bivariate normal distribution to represent the multivariate flood events including volume/ peak and volume/duration due to values of correlation coefficients. The correlation coefficient between volume and peak of transformed flow in Mokelume River is 0.677 and between volume and duration is 0.95. Both of the values are greater than 2/3. According to Yue [1999], the bivariate normal distribution is appropriate when the correlations are greater than 2/3. Otherwise, EV1 distribution would be considered [Yue, 1999].

2.5 CONCLUSIONS

This study presents the procedure of applying the bivariate normal distribution model with normal marginal to analyse multivariate flood events. The model is used to develop the joint distributions of combinations of flood characteristics, namely flood peaks and volumes, and then flood volumes and durations. Based on this model, if the marginal distributions of two random variables can be represented by the lognormal distribution, one can readily obtain the joint probability distributions, the conditional distributions and the associated return periods of these variables. The parameters of the model can be estimated easily from the sample data based on the single variable normal distribution.

The method is tested using observed flood data from the Mokelumne River basin in Northern California. A good agreement is observed between the theoretical and observed distributions. The results point out that the proposed method provides additional information that cannot be obtained by single variable flood frequency analysis, such as the joint return periods of the combinations of variables of interest (flood peak and volume, or flood volume and duration), and the conditional return periods of these variables. These results also indicate that the model can contribute in solving several problems of hydrologic engineering design and management. For example, given a flood-event return period, it is possible to obtain various occurrence combinations of flood peaks and volumes, and vice versa. These results can be useful in the analysis and assessment of the frequency of several hydrologic problems, such as spillway design and flood control.

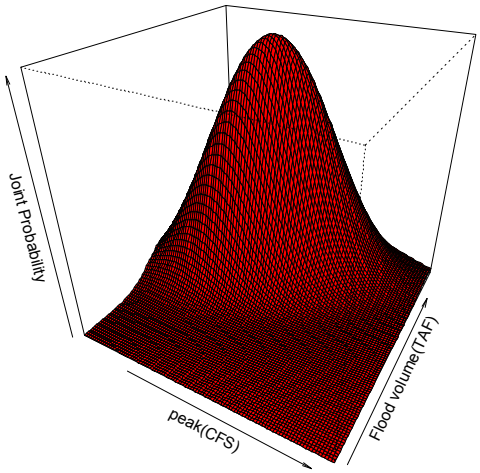


FIGURE 2-7(A-C): JOINTLY PDF OF FLOOD PEAK AND VOLUME WITH DIFFERENT VIEW DIRECTIONS

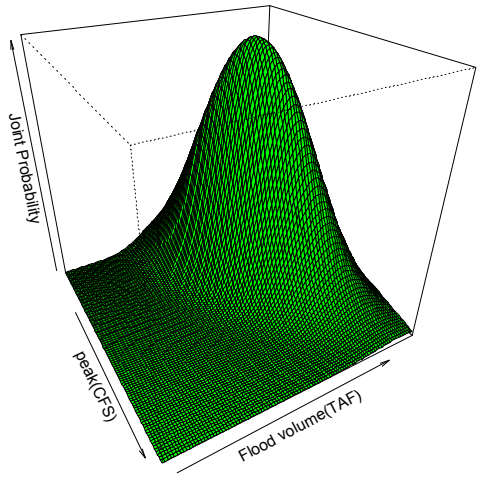


Figure 2-7(a-c): Jointly PDF of flood peak and volume with different view directions

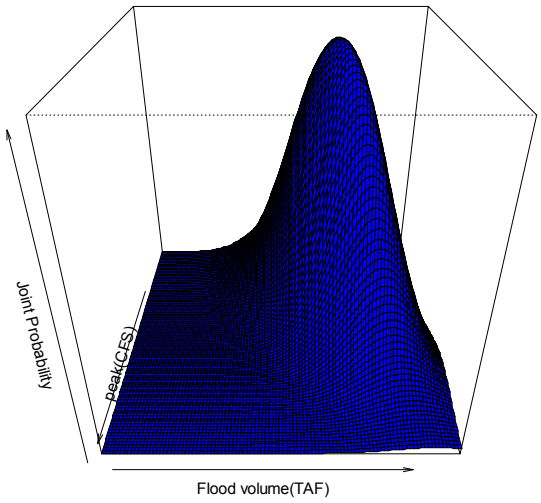


Figure 2-7(a-c): Jointly PDF of flood peak and volume with different view directions

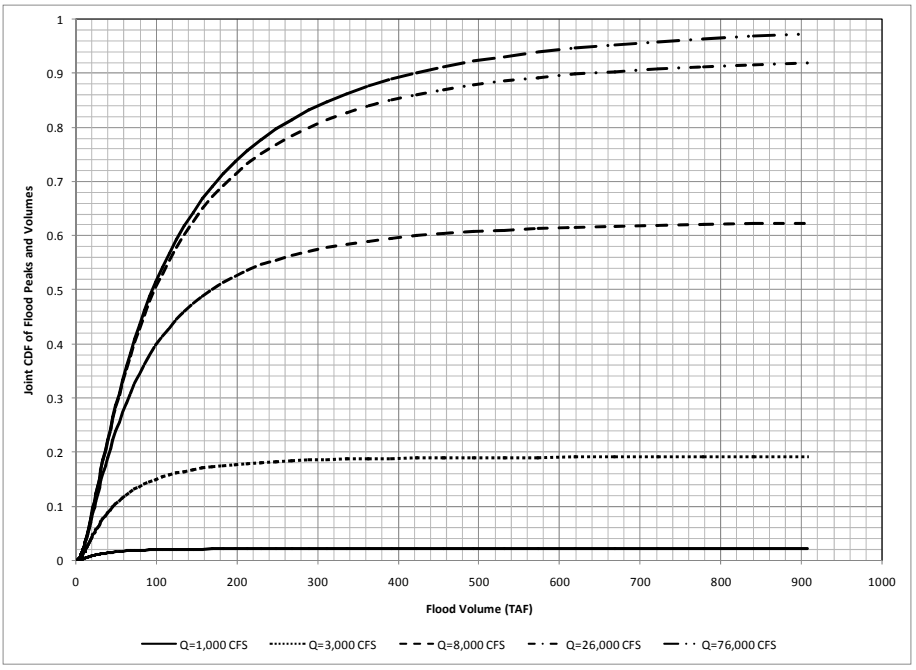


FIGURE 2-8 (A): JOINTLY CDF OF FLOOD PEAK AND VOLUME

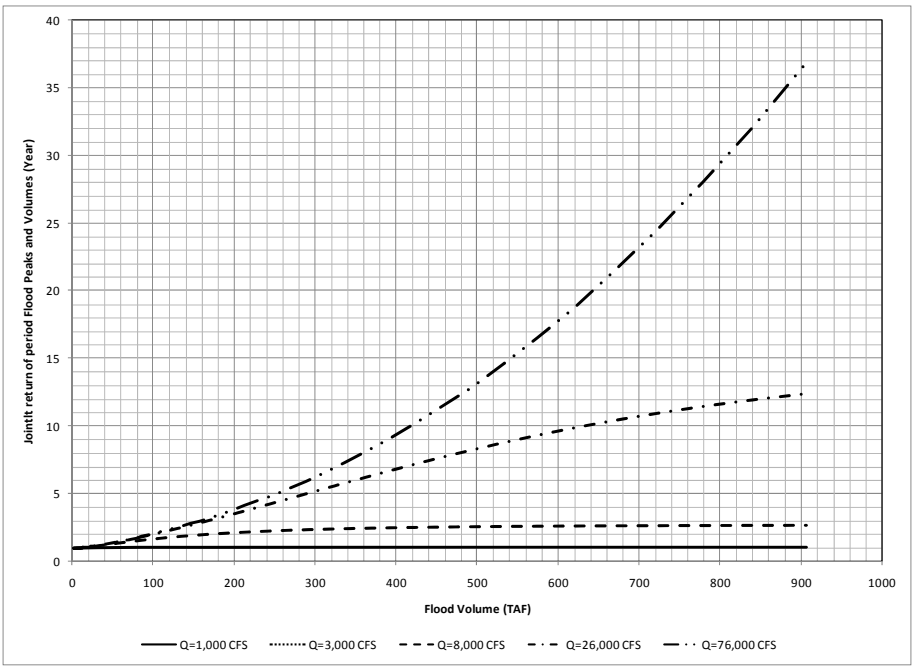


FIGURE 2-8 (B): JOINTLY RETURN PERIOD OF FLOOD PEAK AND VOLUME

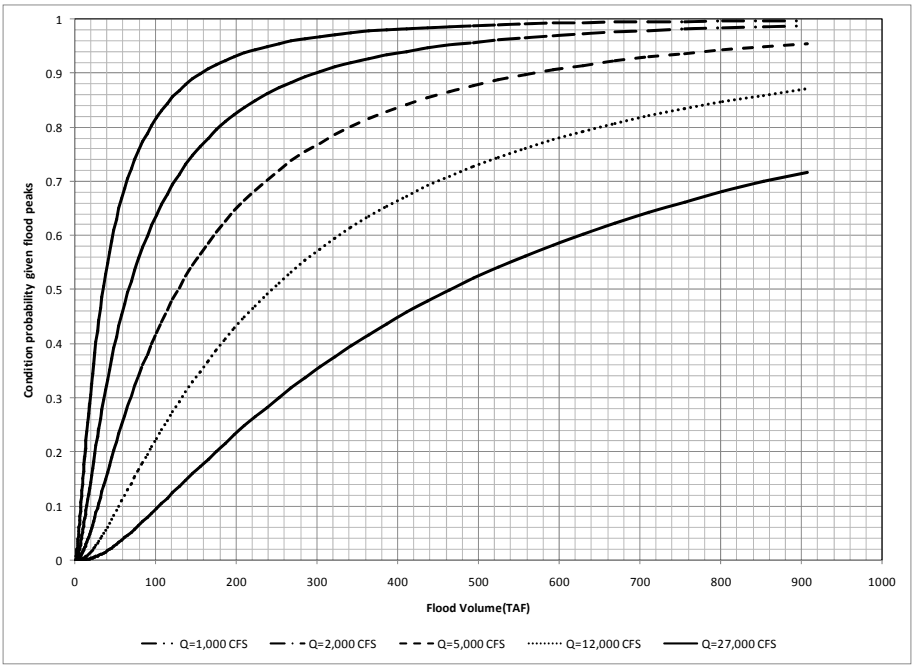


FIGURE 2-9 (A): CONDITIONAL PROBABILITY OF FLOOD VOLUME GIVEN FLOOD PEAKS

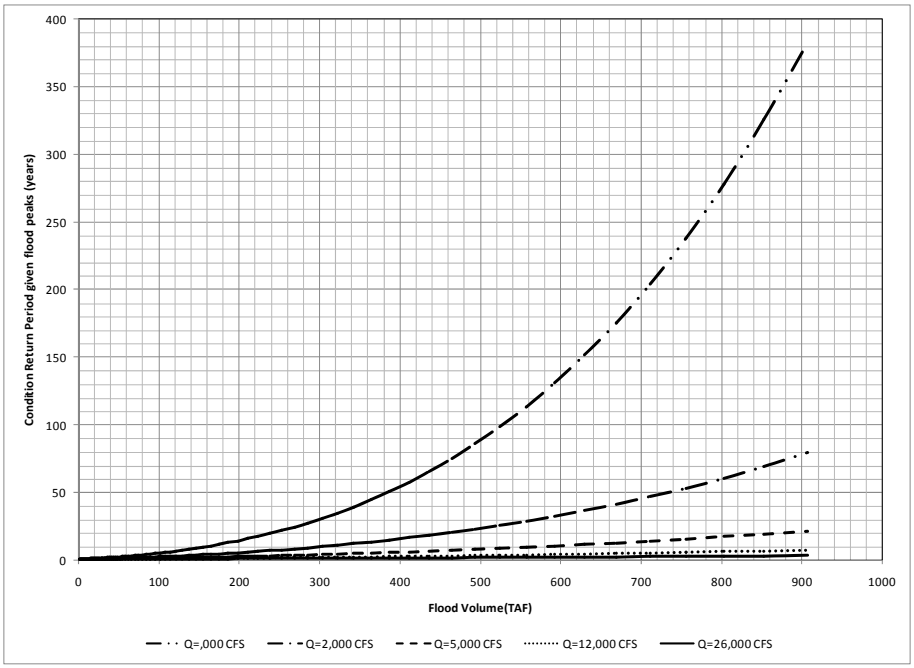


FIGURE 2-9 (B): CONDITIONAL RETURN PERIOD OF FLOOD VOLUME GIVEN FLOOD PEAKS

REFERENCES

Ashkar F. (1980). “Partial duration series models for flood analysis”. Ph.D Thesis, Ecole Polytechnique of Montreal.

Bobee B, Ashkar F. (1991). “The Gamma Family and Derived Distribution Applied in Hydrology." Water Resources Publications. U.S.A.

Castillo, E. (1988) “Extreme Value Theory in Engineering”, Academic Press Inc.

Chow VT.(1953). “Frequency analysis of hydrological data with special application to rainfall intensities." University of Illinois, Engineer in Experimental Station Bulletin 414.

Chow VT. (1954). “The log-probability law and its engineering applications” Proceedings, American Society of Civil Engineers 80 (Separate No.536):25pp

Correia FN. (1987) “Multivariate partial duration series in flood risk analysis” In Hydrologic Frequency Modeling, Singh VP (ed) Reidel: Dordrecht: 541 – 554.

Cunnane C. (1987) “Review of statistical models for flood frequency estimation” In Hydrologic Frequency Modelling, Singh VP (ed) Reidel: Dordrecht: 49-95.

DeGroot, M., Schervish, M.(2002) “Probability and Statistics”, Addison-Wesley: New York.

Goel NK, Seth SM, Chandra S. (1998) “ Multivariate modeling of flood flows” American Society of Civil Engineers, Journal of Hydraulics Engineers 124(2): 146-155.

Gumbel, E. J. (1958) “Statistics of Extremes”, Columbia University Press, New York.

Lambein, W.(1949) “Annual floods and the partial-duration flood series”, Transactions, American Geophysical Union 30(6):879-881.

Kelly KS, Krzysztofowicz R.(1987) “A bivariate meta-Gaussian density for use in hydrology”, Stochastic Hydrology and Hydraulics 11:17-31.

Krstanovic PF, Singh VP. (1987) “A multivariate stochastic flood analysis using entropy” In Hydrologic Frequency Modelling, Singh VP (ed) Reidel: Dordrecht: 515-539.

R Core Team. (2010) “An Introduction to R”. Website: http://cran.r-project.org/doc/manuals/R-intro.html

Robinson-Bryon Inc. (2006) “Floodplain resources characterization report”, in “Cosumnes & Mokelumne Rivers Floodplain Integrated Resources Management Plan”, California Bay-Delta Authority.

Sackl B., Bergmann H. (1987) “A bivariate flood model and its application” In Hydrologic Frequency Modelling, Singh VP (ed) Reidel: Dordrecht: 571-582.

Sangal B., Biswas AK (1970) “The 3-parameter lognormal distribution and its applications in hydrology” Water Resources Research 6(2): 505-515.

Stedinger JR., Vogel RM., Foufoula-Georgiou E. (1993) “Frequency analysis of extreme events” In Handbook of Hydrology, Maidment DR. (ed) from 18-1 to 19-65. McGraw Hill: New York.

Todorovic, P. (1978) “Stochastic models of floods”, Water Resources Research, Vol. 14 (2), pp. 345-356.

Watt WE., Lathem KW., Neill CR., Richard TL., Rousselle J. (1989) “ Hydrology of Floods in Canada] a Guide to Planning and Design” National Research Council of Canada, Canada.

Weibull W. (1939) “A Statistical Theory of the Strength of Materials” Ingeniors Vetenskaps Akademiens-Handlingar, 151, The Royal Swedish Institute for Engineering Research Proceedings: Stockholm.

Yue, S. (1999) “A bivariate extreme value distribution applied to flood frequency analysis” Nordic Hydrology, 32(1), 2001, 49-64.

Yue S. (1999) “Applying the bivariate normal distribution to flood frequency analysis” Water International 24(3): 248-252.

Yue S., Ouarda TBMJ., Bobee B., Legendre P., Bruneau P. (1999) “The Gumbel mixed model for flood frequency analysis” Journal of Hydrology 226: 88-100.

Explanation of Variables and simplified terms

PDF – probability of density function

CDF – cumulative density function

AMS – Annual Maximum Series

DWR – Department of Water Resources

USACE – U.S. Army of Corps of Engineer

USGS – United States of Geological Services

CHAPTER 3 DEVELOPMENT OF DESIGN FLOOD HYDROGRAPHS USING PDF SHAPES

Da Yu worked so hard and he spent almost all of his life in fields. During 13 years, he had passed through his home three times but never step in.

- Huai Nan Zi, Han Dynasty Philosopher,135 BC

SUMMARY

Flood hydrograph design is a key aspect flood control studies for reservoir operation. Numerous methods have been developed to summarize flood hydrograph characteristics. Using probability density functions (PDFs) to fit the shapes of flood hydrographs has drawn more attention recently due to improvements in statistical techniques, including algorithms for fitting. This chapter introduces a framework to develop design flood hydrographs for reservoir reoperation. Three steps are presented. 1) Flood hydrograph separation and modification: Typical flood hydrographs were separated, selected and converted to dimensionless ones; 2) PDF form fitting and selection: Beta, Gamma. Lognormal and Weibull distributions were selected and compared to fit modified hydrographs based on goodness of fit criteria including RMSE and coefficients of determinations. 3) Development of design flood hydrographs: The design shape parameter values were found from frequency analysis and finally, the design flood hydrographs for 10-, 20-, 50-, 100- and 200-year return periods were derived from the combinations of hydrograph shape, flood volume and durations. The applicability of the proposed method is demonstrated using observed data from flood hydrographs for Mokelumne River in Northern California. The gamma PDF is suitable for this application site's floods.

3.1 INTRODUCTION

Three important parameters of a hydrograph are peak discharge (Q_p), duration (D) and volume (V), which can be computed for effective flood control planning, design and management. Using these three parameters, hydrographs could be developed synthetically for partially gauged watersheds using empirical relationships among them. Generally, flood peak discharge and hydrograph volume are considered for designing flood reservoir storage and other hydraulic structures and flood duration is considered for reservoir operation [Cruise, 1996].

In addition to the three parameters above, the shapes of flood hydrographs are important for flood management. Based on the relationship between flood peak-occurrence time (t_p) and the time interval between the centroid of a flood hydrograph from the origin (t_c), flood hydrographs are categorized into three different shapes, namely, prior-peak or positively skewed shape ($t_p < t_c$), midpeak or symmetrical shape ($t_p = t_c$) and posterior-peak or negatively skewed shape ($t_p > t_c$). Different flood hydrograph shapes may cause significant differences in cost and flood policies of water resources management projects. Figure 3.1 shows the schematic illustration of three flood hydrograph shapes. For example, for two floods with the same peak and volume, one has the prior-peak shape and the other has the posterior shape. To get the same degree of protection, i.e., to cut the same amount of the flood peak by reservoir routing, the flood with the posterior shape requires more reservoir storage volume than that with the prior shape if a channel capacity operating rule is used [Chow et al. 1988; Ergish, 2010].

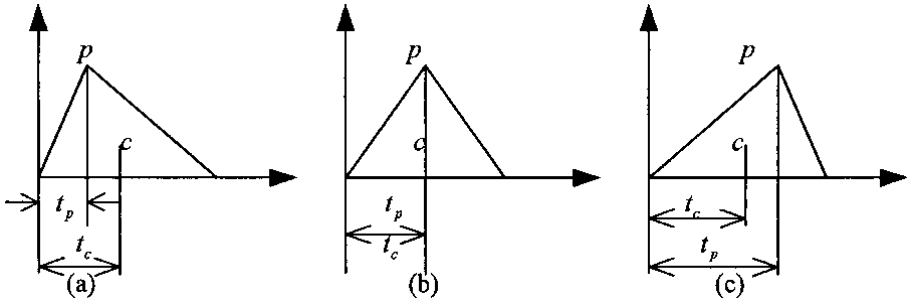


FIGURE 3.1 SCHEMATIC ILLUSTRATION OF DIFFERENT FLOOD HYDROGRAPH SHAPE: A) PRIOR-PEAK SHAPE; B) MIDPEAK SHAPE; C) POSTERIOR-PEAK SHAPE

The shapes of flood hydrographs are highly varied and stochastic. Many factors govern the shape of a flood hydrograph, such as precipitation intensity, precipitation pattern, precipitation amount, storm center and movement, snow-depth spatial distribution, temperature, basin geometry and others are random phenomena. In practice, a river may have various shapes of flood hydrographs, as different storm or snowmelt events produce different floods. To catch hydrograph shape is essential for a design flood hydrograph (DFH). In the past, various methods have been proposed to construct a design flood hydrograph (DFH). These methods may be grouped into the three classes: 1) the unit-hydrograph (UH) methods including traditional and synthetic hydrographs [Sherman 1932; Dooge 1959; Chow 1964; Chow et al. 1988; Pilgrim and Cordery 1993; Yue and Hashino 2000; Snyder1938; US-SCS 1985]; 2) the typical hydrograph (TH) method [Nezhikhovsky 1971; Sokolov et al. 1976]; and 3) the statistical method (SM) mainly using PDFs to provide hydrograph shapes [Gray 1961;Sokolov et al. 1976; Ciepielowski 1987; Haktanir and Sezen 1990]. Each DFH construction approach has pros and cons. In this study, SM method which employs PDFs to represent the shape of a DFH was selected due to mathematical convenience and versatility.

Four PDFs, namely, two parameter Beta, Weibull, Gamma and Lognormal have been widely used to develop design flood hydrographs [Pramanik, et al., 2010; Koutsoyiannis and Xanthopoulos, 1989; Bhunya et al., 2004; Nadarajah, 2007; Rai et al., 2008]. The major criterion for selecting of the PDFs is the resemblance of the shape of the PDF to the shape of observed flood hydrographs. In this study, annual maximum series (AMS) observed flood hydrographs were used and modified to fit the four PDFs. The most suitable PDF fitted hydrographs was selected to represent the dimensionless design flood hydrograph. Further, the best-fitted PDF parameters, flood duration, volume of the hydrographs of different return periods were estimated and used to develop design flood hydrographs using the best fitted PDF. The design flood hydrographs will be used as input for reservoir operation and reoperation studies.

The organization of this chapter is as follows. Section 2 describes the concept development including review of methods to generate hydrographs and four widely used PDFs to fit hydrographs. Section 3 presents the methods flow chart used here, including three steps from observed hydrographs separation and modification, PDF fitting methods, and development of design flood hydrographs. Section 4 gives an application to Camanche/Pardee reservoirs inflow in Northern California. Section 5 summarizes conclusions and limitations.

3.2 CONCEPT DEVELOPMENT

This section presents the common methods to construct flood hydrographs and four widely used PDFs, Beta, Gamma, Lognormal and Weibull, to fit flood hydrographs.

3.2.1 Generation of hydrographs

To construct flood hydrographs synthetically, various empirical models and statistical techniques have been proposed. These methods may be grouped into the three classes including Unit Hydrograph (UH), Traditional Hydrograph (TH), and Statistical Methods (SM).

Unit Hydrographs (UH) are most widely used method to develop flood hydrographs. The traditional UH method [Chow et al. 1988; Pilgrim and Cordery 1993] is based on assumptions of spatial homogeneity and a linear response of runoff to net rainfall. These assumptions are not always valid. Also, derivation of a flood hydrograph requires precipitation data to represent the areal rainfall of a given basin. Basins seldom have data to meet this requirement. The synthetic UH method proposed by Snyder [1938] is also frequently used. This method is based on assumption of a triangular shaped hydrograph in which the watershed lag time, the time to attain the peak and the magnitude of peak discharge are estimated empirically. Practically, few hydrographs have triangular shapes.

Traditional Hydrographs (TH) concepts were developed by Nezhikhovsky [1971] and Sokolov et al. [1976] to obtain the shape of the design flood hydrographs. It selects the highest peak hydrograph from a series of historical hydrographs as the TH. The ratio between the peak discharge of a given return period and the peak discharge of the TH is estimated and used as an amplifier, which is multiplied with the ordinates

of TH to develop the design flood hydrograph. The amplifier also can be the ratio of design volume with a given return period to the TH volume. Hereafter the former is referred to as a peak-amplitude method (PAM) and the latter is a volume-amplitude method (VAM). In reservoir flood control, VAM has been widely used [Hickey, et al., 2002]. Though these methods are commonly used to derive design flood hydrographs, it fails to get the volume of the design flood hydrographs the same as the volume obtained from frequency analysis when PAM was applied. Also it fails to get the peak of the design flood hydrographs the same as the peak obtained from frequency analysis when VAM was applied.

Many studies have been conducted to develop UHs using Statistical Methods (SM) [Gray, 1961; Sokolov et al.,1976; Ciepielowski, 1987; Haktanir and Sezen, 1990; Bhunya et al., 2004, 2007]. Mostly, Gamma and Beta PDFs have been used to represent the shape of UH [Koutsoyiannis and Xanthopoulos, 1989; Haktanir and Sezen, 1990; Haan et al., 1994; Bhunya et al., 2003, 2004; Rai et al., 2008]. Researchers have also discussed the flexibility of PDFs to produce different shapes by changing their parameter values [Bhunya et al., 2008]. Rai et al. (2009) evaluated nine probability density models in 18 different watersheds to develop UHs and ranked their suitabilities according to the size of the watersheds. However, these approaches do not present a systematic view to catch the randomness of flood hydrographs.

To describe the statistics of flood hydrograph shapes, Yue et al. attempted to model the shape of hydrographs using two parameter Beta PDFs by introducing shape mean and shape variance of the hydrographs. The shape mean and shape variance have been used analogous to mean and standard deviation of a pool of random events. With the frequency analysis of shape variables, design flood hydrographs will be derived with corresponding flood volume, peaks and duration [Yue, et al., 2002].

3.2.2 Probability density functions

A brief description of four widely used PDFs including Beta, Gamma, Weibull and Lognormal distributions and hydrographs shape variables is presented.

Beta probability density function

The two parameter Beta PDF was developed by Johnson and Kotz [1970]. A two parameter Beta PDF has two shape parameters, which regulate the shape of the Beta PDF curve. The mathematical expression for the hydrograph using Beta PDF is presented as followed:

$$q(t, a, b) = \frac{1}{B(a,b)} t^{a-1}(1-t)^{b-1}; \qquad B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,dt \qquad (3.1)$$

The function B(a,b) is called the Beta function. The notations a and b are the shape parameters and are always positive numbers. The values of t must lie on the interval [0 1]. Many studies have used two and three parameter Beta PDF for synthesis of UHs [Koutsoyiannis and Xanthopoulos, 1989; Haktanir and Sezen, 1990; Bhunya et al., 2004; Rai et al., 2008; Rai et al., 2009]. The shape of Beta PDFs can be regulated by their shape parameters. The condition of a = b produces a symmetric hydrograph, whereas the positively skewed hydrographs are produced when a < b and negatively hydrographs are obtained for the condition of a > b. For any values of a and b more than unity, the shape of the PDF curve becomes concave and resembles shape of the hydrograph. With simultaneous increase in the values of a and b beyond one, the peak of the PDF increases. Therefore, to simulate a prior-peak hydrograph, b exceeds a, while to simulate posterior-peak hydrograph, a exceeds b. When the hydrograph shape is symmetric, a is equal to b. Figure 3-2 (a) shows several Beta PDF's shapes with different a and b values.

The parameters of the Beta pdf can be estimated using the method of moments [DeGroot,, Schervish, 2002]:

$$\mu = \frac{a}{a+b}; \qquad \sigma^2 = \frac{ab}{(a+b)^2(a+b+1)} \qquad (3.2)$$

where μ and σ^2=population mean and variance of a variable, respectively. In practice, the sample mean and variance are used to replace the population mean and variance, which can be estimated from sample data.

Gamma probability density function

The two parameter Gamma PDF is widely used to model the hydrograph shapes [Haktanir and Sezen, 1990; Haan et al., 1994; Bhunya et al., 2003; Nadarajah, 2007]. The mathematical expression of Gamma PDF is:

$$q(t,\tau,\theta) = t^{\tau-1}\frac{\exp(-\frac{t}{\theta})}{\theta^{\tau}\Gamma(\tau)} \qquad (3.3)$$

The notation τ is the shape parameter and θ is the scale parameter. Γ(τ) is called the Gamma function and is expressed Γ(τ)= (τ-1)! and is valid for integer values of τ. Bhunya et al. [2003] used the Marquardt algorithm to optimize the weights of non-linear equations for estimating the PDF parameters. Eventually, they developed a series of approximate equations to estimate the ordinates of the dimensionless hydrographs. Therefore, most of Gamma PDFs are negative skewed and used to represent the prior-peak hydrographs. When the shape parameters increase, the PDF shapes are more symmetric. Figure 3-2 (b) shows several Beta PDF's shapes with τ and θ values.

The parameters of the Gamma pdf can be estimated using the method of moments [DeGroot & Schervish 2002]:

$$\mu = \tau\theta\ ; \qquad \sigma^2 = \tau\theta^2 \qquad (3.4)$$

where μ and σ^2=population mean and variance of a variable, respectively. In practice, the sample mean and variance are used to replace the population mean and variance, estimated from sample data.

Lognormal probability density function

The lognormal PDF is a single-tailed probability density function of any random variable whose logarithm is normally distributed. The mathematical expression for the hydrograph using this PDF is:

$$q(t,\mu,\sigma) = \frac{1}{t\sigma}\exp\left(-\ln(t)-\mu\right)^2/2\sigma^2 \qquad (3.5)$$

The notations μ and σ are the mean and standard deviation respectively. Values of t must be always exceed 0. The value of μ can range between -∞ to +∞, and σ is always exceed 0. Few papers use Lognormal PDF to develop flood hydrographs [Nadarajah, 2007; Rai et al., 2009]. Most Lognormal PDFs are negative skewed which represent the prior-peak hydrographs. Figure 3-2 (c) shows a few Lognormal PDF shapes with different μ and σ values.

The parameters of the Lognormal pdf can be estimated using the method of moments [DeGroot & Schervish, 2002]:

$$\mu = \exp(\mu+\sigma^2); \qquad \sigma^2 = \exp(\sigma^{-2}-1)\exp(2\pi+\sigma^{-2}) \qquad (3.6)$$

where μ and σ^2=population mean and variance of a variable, respectively. In practice, the sample mean and variance are used to replace the population mean and variance, estimated from sample data.

Weibull probability density function

The two parameter Weibull PDF was first introduced by Rosin and Rammler [1933]. For some values of its parameters, the function mimics normal and exponential distribution functions. The mathematical expression for this PDF is:

$$q(t,k,\lambda) = \frac{k}{\lambda}\left(\frac{t-k}{\lambda}\right)^{k-1}\exp-\left(\frac{t-k}{\lambda}\right)^k \qquad (3.7)$$

Where, k and λ are positive numbers and are treated as shape and scale parameters respectively. The above expression is valid for all t > 0. The Weibull functions return 0 for non-zero t. The scale parameter is responsible for shrinking and widening the hydrograph, whereas the skewness of the hydrographs is obtained with the certain combinations of k and λ. Considerably lower values of λ with higher k produce positively skewed hydrographs. Most of Weibull PDFs are negative skewed and are used to represent the prio-peak hydrographs. Figure 3-2 (d) shows a few Weibull PDF's shapes with different k and λ values.

The parameters of the Weibull pdf can be estimated using the method of moments [DeGroot & Schervish, 2002]:

$$\mu = \lambda\Gamma\left(1+\frac{1}{k}\right); \qquad \sigma^2 = \lambda^2\Gamma\left(1+\frac{2}{k}\right) - \mu^2 \tag{3.8}$$

where μ and σ^2=population mean and variance of a variable, respectively. In practice, the sample mean and variance are used to replace the population mean and variance, which can be estimated from sample data.

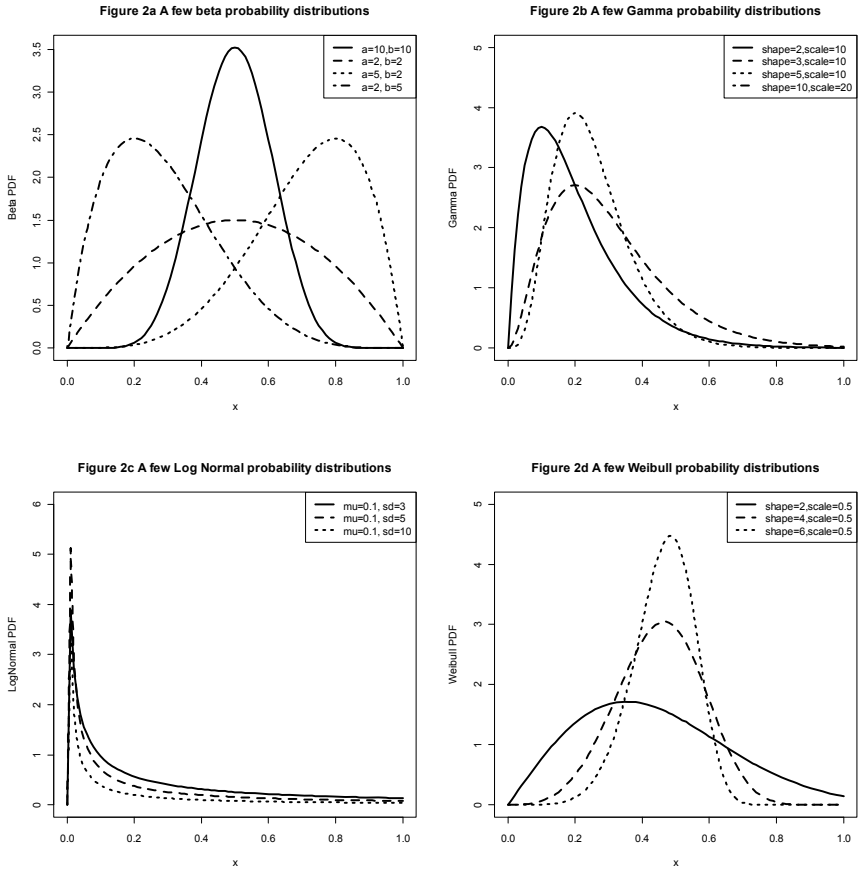


FIGURE 3.2 BETA, GAMMA, LOG NORMAL AND WEIBULL PROBABILITY DISTRIBUTIONS SHAPES

Shape variables of flood hydrographs

The shape of a flood hydrograph is a random event. For representing the statistical properties of the shape of a flood hydrograph, two shape variables - shape mean (S_m) and shape variance (S_v) - are defined as follows.

1). Shape mean S_m is the centroid of a flood hydrograph that represents the central tendency of a flood hydrograph. It is estimated using (Fig. 3.2):

$$S_m = \frac{1}{A}\sum_{i=1}^{n} x_i A_i = \frac{1}{A}\sum_{i=1}^{n} x_i f(x_i)dx_i \tag{3.9}$$

where A = volume of the flood hydrograph; x_i = horizontal distance from the origin to the center of subarea Ai ; and f (x_i) = ordinate (flow magnitude) corresponding to x_i .

2). Shape variance S_v is the variance of a flood hydrograph that represents the spreadness of a flood hydrograph. It is defined as

$$S_V = \sum_{i=1}^{n}(x_i - S_m)^2 A_i = \sum_{i=1}^{n}(x_i - S_m)^2 x_i \cdot f(x_i)dx_i \quad (3.10)$$

These two equations indicate that the principle used for numerically calculating the shape variables is the same as that for deriving the mean m and variance σ^2 of a random variable X, in which A=1, as expressed

$$\mu = \int_{-\infty}^{+\infty} xf(x)\,dx; \qquad \sigma^2 = \int_{-\infty}^{+\infty}(x-\mu)^2 f(x)dx \quad (3.11)$$

where f(x)=probability density function (PDF) of a distribution that the random variable X follows. From equations, the shape of a flood hydrograph can be interpreted as a ‘‘PDF’’ [f (x_i)=f (x)]. Thus, the PDF of a probability distribution can be used to represent the shape of a flood hydrograph. The shape variables—shape mean and shape variance derived from a hydrograph—are also random variables. Distributions of these shape variables can be selected using the goodness-of-fit test such as the Kolmogorov-Smirnov (KS) test. Therefore, given the aforementioned shape variables S_m=μ and S_v=σ^2, the parameters of the PDFs can be derived and the shape of a hydrograph can be represented by the PDFs. Given some quantiles of shape variables corresponding to a specific return period (T), the corresponding shape of a flood hydrograph can be obtained using the PDF.

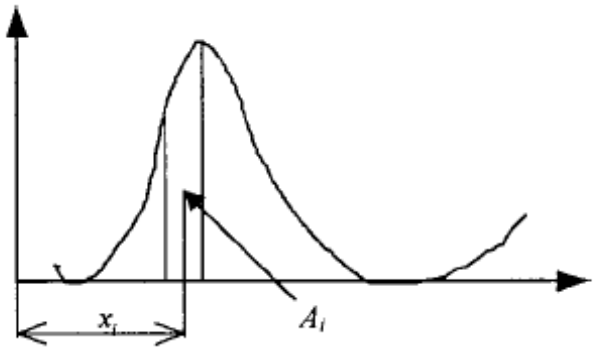


FIGURE 3.3 SCHEMATIC ILLUSTRATION OF CALCULATION OF CENTROID OF FLOOD HYDROGRAPH

Distribution of shape variables

The nonexceedance probability is estimated using the Weibull plotting position formula [Weibull 1939; Chow 1953].

$$P_k = \frac{k}{N+1} \quad (3.12)$$

where P_k is cumulative frequency, the probability that a given value is less than the kth smallest observation in the data set of N observations.

To select the distributions from which these random variables are drawn, the Kolmogorov-Smirnov (KS) test [Degroot, et al., 2002] was used as a goodness of fit statistical test. Several methods for goodness of fit test including chi-square, Kolmogorov-Smirnov and Anderson-Darling (AD) tests are used in hydrology and water resources. The Chi-square test is the oldest goodness of fit test and maybe thought of as a formal comparison of a histogram with the fittest density. The Chi-square test can be applied either to discrete distributions or continuous ones. One disadvantage of this test is it requires a sufficient sample size for the Chi square approximation to be valid. The AD test is another test available for a few specific distributions. The KS test is often used in continuous distributions. It is based on a comparison between the empirical distribution function (ECDF) and the theoretical one. KS is more powerful than Chi-square test when sample size is not large.

KS is based on a comparison between the empirical distribution function (ECDF) and the theoretical one defined as $F(x) = \int_{\alpha}^{x} f(y,\theta)dy$, where f(y,q) is the PDF. Given n ordered data points X1, X2, ..., Xn, the ECDF is defined as:

$$F_n(X_i) = \frac{N(i)}{n} \qquad (3.13)$$

where N(i) is the number of points less than X_i (X_i are ordered from smallest to largest value). This is a step function that increases by 1/n at the value of each ordered data point. The test statistic used is:

$$D_n = \sup_{1 \le i \le n} |F(x_i) - F_n(x_i)| \qquad (3.14)$$

that is the upper extreme among absolute value differencies between ECDF and theoretical CDF. The hypothesis regarding the distributional form is rejected if the test statistic, D_n, exceeds the critical value obtained from a table, or, which is the same, if the p-value is less than the significance level. Therefore, the KS test is defined for the hypothesis:

H_0: the data follow a specified distribution; H_A: the data do not follow the specified distribution

We accept null hypothesis that the data follow a specific distribution because the p-value is enough higher than significance levels found from the statistical literature. Normally, p-value is compared with the value of 0.05 [DeGroot & Schervish, 2002].

To be convenient, lognormal distributions are assumed for shape mean and shape variance. The test indicates that both the shape mean and shape variance can be represented by the lognormal distribution at the significance level of 0.05. They are independent each other. The PDF of the lognormal distribution is represented by

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma_Y^2}} \exp\left[-\frac{(\ln(x)-\mu_Y)^2}{2\sigma_Y^2}\right] \qquad (3.15)$$

$$Y = \ln(X) \qquad (3.16)$$

where μ_Y and σ_Y mean and standard deviation of Y, respectively. They can be derived using the following formulas [Stedinger et al. 1993]:

$$\mu_y = [\ln(1 + \frac{\sigma_X^2}{\mu_X^2})]^{1/2}; \qquad \mu_x = \ln(\mu_X) - \frac{\sigma_Y^2}{2} \qquad (3.17)$$

and where μ_X and σ_X = mean and standard deviation of the random variable X. The cumulative distribution function (CDF) and the associated return period T of the variable X can be computed through the normal distribution as follows:

$$F(x) = Pr(X \le) = Pr[Y \le \ln(x)] = \Phi[\frac{\ln(X)-\mu_Y}{\mu_Y}] \qquad (3.18)$$

$$T_X = \frac{1}{1-F(X)} \qquad (3.19)$$

in which Φ = CDF of the standard normal distribution. As there is no analytical form of the CDF, it can be calculated by integrating the corresponding PDF.

3.3 METHOD

In the present study, the four PDFs described above were employed to fit the observed inflow time series for flood hydrographs. Continuous stream flow hydrographs at a river are the delayed response of precipitation or snowmelt, whose shapes are affected by the duration and intensity of precipitation or snowmelt as well as human interference and presence of structures within the catchments. In this section, the entire process of computation of best-fitted PDF curves for all observed hydrographs as well as the procedures followed to develop design flood hydrographs of 10-, 20-, 50-, 100- and 200-year return

periods are presented. Figure 3.4 shows this procedure's flowchart. Three sections including flood hydrograph modification, PDF fitting and development of design flood hydrographs are presented.

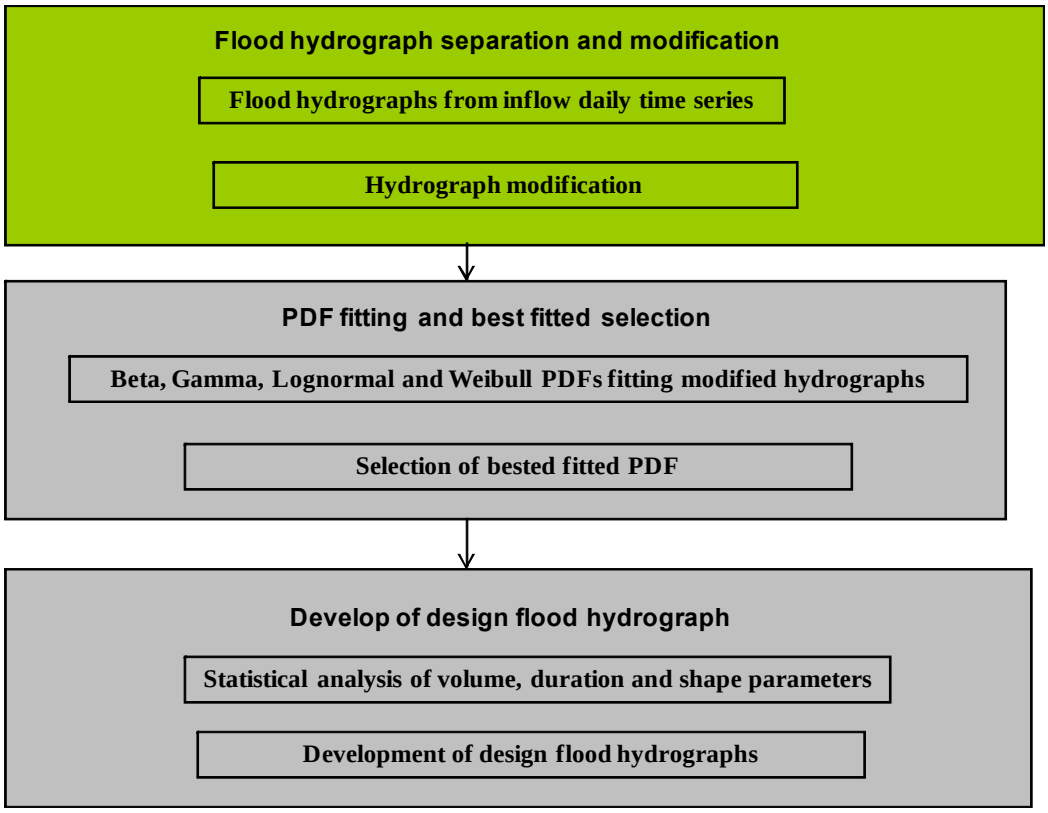


FIGURE 3.4 FLOWCHART OF DESIGN FLOOD HYDROGRAPH PROCEDURE

3.3.1 Flood hydrograph separation and modification

The inflow flood hydrographs having the highest peak flow were selected from each year's stream flow record and were used for further analysis. The starting points of the rising limb and the end point on the recession limb while the flow reduces to base flow of the hydrographs were determined. The base flow separations of the hydrographs were performed to extract the direct stream flow hydrographs. Many techniques are practiced to separate base flow from stream flow hydrographs. Three common graphical techniques, namely, constant discharge, constant slope and concave methods are practiced for this purpose. These techniques aim mainly to separate quick flow from slow flow for the purposes of flood analysis and prediction [Pramanik et al, 2010]. Figure 4 shows the schematic illustration of separation of surface runoff from baseflow. For simplicity, the start and end of a flood hydrograph are connected by a straight line, and this straight line is considered as the shape of base flow. The duration (D) of a flood is given by

$$D = e - s \qquad (3.20)$$

The flood volume V can be computed by

$$V = \left[\sum_{i=s}^{e} q_i - \frac{q_s + q_e}{2}\right] - \frac{D(q_s + q_e)}{2} \qquad (3.21)$$

where q_i is the ith day observed daily streamflow value; and q_s and qe are observed daily streamflow values on the start date and end date of a flood, respectively. The flood peak Q is the maximum surface flow and is given by

$$Q = \max(q_i - qb_i) \quad (i = s, s+1, \ldots e) \qquad (3.22)$$

$$qb_i = q_s + \frac{(i-s)}{D}(q_e - q_s) \qquad (3.23)$$

where qb_i = i th day base flow value. On the basis of the above procedure, the 12 flood-hydrograph series and the corresponding annual flood peak, volume, and duration series were obtained in this study. The results will be summarized in the next section.

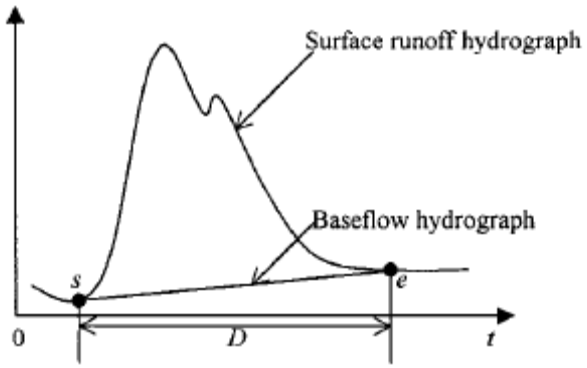


FIGURE 3.5 SCHEMATIC ILLUSTRATION OF SEPARATION OF SURFACE RUNOFF FROM BASE FLOW

The derived direct stream flow hydrographs (DSFH) were transformed to dimensionless hydrographs before they were used for curve fitting using the PDF equations. The volumes (V) under the direct flood hydrographs were computed using the trapezoidal method and the ratio of duration of each hydrograph (D) to the volume was computed. The computed fraction (D/V) was multiplied with ordinates of each DSFH and the values in the abscissa of each DSFH were divided by D to convert DSFH to dimensionless form. The objective of this transformation of DSFH to dimensionless hydrographs was to constrain the volume and time base as unity. The unit hydrographs are comparable to the PDFs in terms of describing the shapes and scales. Both of PDFs and dimensionless form of the DSFHs have characteristics with the unit area of 1 by integrating functions within the limit of 0 to 1. Figure 3.5 shows the schematic illustration of a flood hydrograph and modified flood-hydrograph.

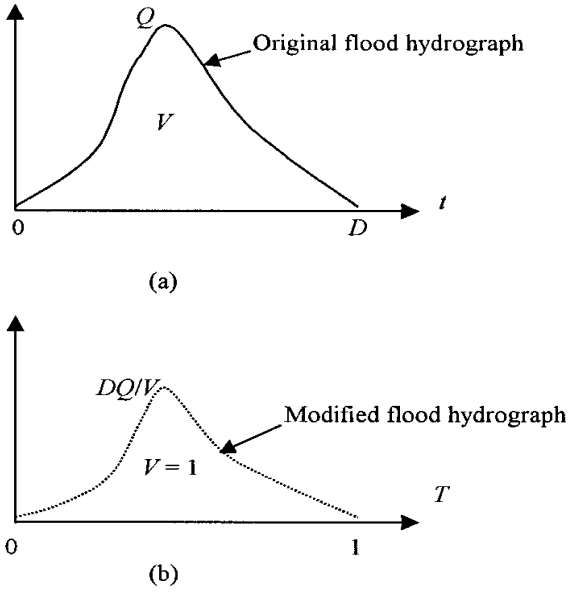


FIGURE 3.6 SCHEMATIC ILLUSTRATION OF A) A FLOOD HYDROGRAPH; B) MODIFIED FLOOD-HYDROGRAPH

3.2.2 PDF fitting and selection of best fitted PDF

The observed modified dimensionless flood hydrographs were fitted with the above-mentioned four PDFs. The shapes generated by the PDFs were then compared with the shape of the observed dimensionless hydrographs. Nonlinear least square method with Marquardt optimization algorithm was used to compute the best-fitted parameters for all selected distributions [DeGroot & Schervish, 2002]. Goodness of fit criteria, namely, root mean square error (RMSE) and coefficient of determination (R^2) were used to evaluate the fitting. The expressions of RMSE and R^2 were presented as followed, respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(q_{(o)i} - q_{(pdf)i})^2}{n}} \qquad (3.24)$$

$$R^2 = \left\{\frac{\sum_{i=1}^{n}(O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2 \sum_{i=1}^{n}(P_i - \bar{P})^2}}\right\}^2 \qquad (3.25)$$

where, $q_{(o)i}$ = observed discharge of i th point, $q_{(pdf)i}$ = PDF computed discharge of i th point, $q_{(o)}$= average observed discharge of all points.

The best fitted PDF function was selected based on the results of RMSE and R^2. After the best fitted PDF function selected, the best PDF parameters of each dimensionless hydrographs were obtained for frequency analysis. In this study, Gamma PDF has lowest RMSE and highest R^2 and was selected for further frequency analysis.

3.3.3 Development of design flood hydrograph

After the best fitted PDF was selected, the shape variables of different return periods of this distribution were determined by frequency analysis. In this study, Lognormal was selected as the distribution model for PDF's shape variables. The Kolmogorov–Smirnov (K–S) test was used to test the best fit distribution. Details about the parameter estimation methods were described by Rao and Hamed [2000] and Haan [2002]. For Gamma PDF, which is best fitted in this study, both shape parameters including mean and standard deviation of shape and scale parameters were subjected to frequency analysis and μ_T and σ_T were estimated for different return periods. The assumption was made that the two shape parameters are independent, although they probably covary somewhat.

Other important parameters of the stream flow hydrographs like flood duration (D) and volume (V) of the hydrographs were used for frequency analysis and the quartiles corresponding to 10-, 20-, 50-, 100- and 200-year return periods were estimated. The ratio V_T/D_T was multiplied with the ordinates of the dimensionless design hydrograph to convert the dimensionless design hydrographs to real valued design flood hydrographs. Finally, the Base flow values were added with the ordinates of the corresponding design flood hydrographs to develop the complete flood hydrographs for different return periods. In this study, flood hydrographs for 10-, 20-, 50-, 100- and 200-year return periods were developed.

3.4 APPLICATION ON LOWER MOKELUMNE RIVER LEVEE SYSTEM

This section presents the site description, results and discussions.

3.2.1 Study area

The Camanche/Pardee reservoirs on the Mokelumne River, a major tributary of San Joaquin River in California, are the application site. Ten major flood events have occurred in this river in the past 50 years, with four occurring in the past 20 years. These four events have accounted for an average flood damage value of $4 millions per each event across the four events [DWR, 2006]. The Mokelumne River watershed covers approximately 920 square miles of mountainous to valley floor terrain. Elevations range from a peak of above 8,800 feet MSL to slightly below sea level in the vicinity of the Delta. The floodplain land use mainly is for agriculture. The winter season lasts about four months, and precipitation is both rainfall and snowfall. Floods are caused mainly by rainfall. Spring also has high flows due to snowmelt, and the spring flood is part of the maximum annual flood both in flood magnitude and volume. As hydro-climatic conditions such as temperature, cumulated snow amount, precipitation, and others dramatically differ

from year to year, observed flood hydrographs clearly show randomness. Figure 3.7 shows the inflow time series of Camanche/Pardee Reservoir and one typical flood hydrograph [DWR, 2008]. In this study, R software is used for the statistical analysis.

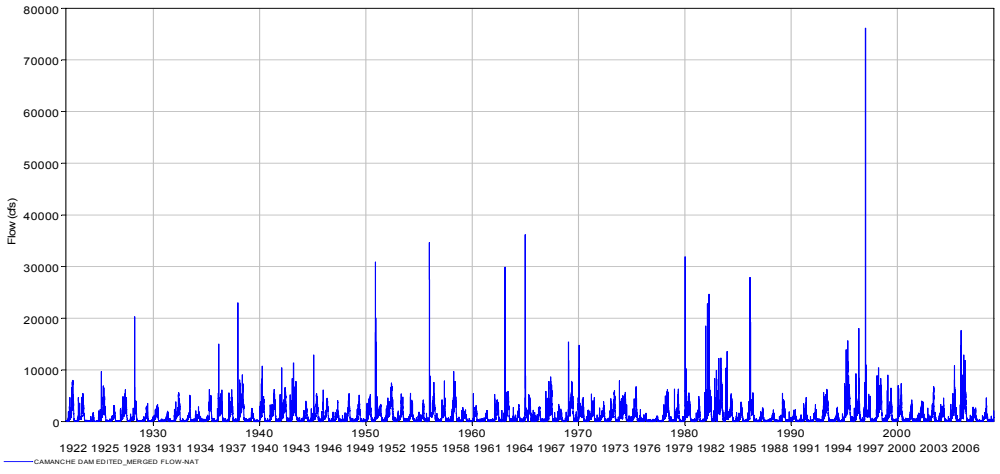


FIGURE 3.7 INFLOW TIME SERIES OF CAMANCHE/PARDEE RESERVOIR AND TYPICAL FLOOD HYDROGRAPH (DWR, 2008)

3.2.2 Results and discussions

Flood hydrograph separation and modification

Base flow separation was performed to extract annual maximum series hydrographs. The detailed procedure is presented in Chapter 2. Constant discharge 2,000 CFS was selected as base flow due to downstream channel capacity limitation. Figure 3.8 shows the selected 11 flood hydrographs as representative hydrographs. Table 3.1 presents the 11 hydrographs' annual peaks, volumes and duration time.

Table 3.1 Selected 11 Flood Hydrograph parameters in Mokelumne River Inflow

Year	Flood Peak	Volume	Duration of rising limbs	Duration of dropping limbs	Total Duration
	CFS	TAF	Day	Day	Day
1998	10,417	59.6	4	14	18
1997	76,137	289.1	7	13	20
1984	13,559	86.9	2	10	12
1980	31,924	155.2	3	9	12
1965	36,137	236.7	3	10	13
1963	29,861	118.6	3	7	10
1956	34,657	196.4	4	11	15
1943	11,380	82.1	6	14	20
1938	22,970	75.6	2	3	5
1936	15,034	56.1	2	6	8
1928	20,300	130.9	3	23	26

Table 3.1: Flood duration D, Volume V and Peaks Q in Mokelumne River

Separated flood hydrographs have to convert to dimensionless unit scalable hydrograph for PDF shapes fitting. Volume, duration and peak values calculated and presented in Table 2.1 in Chapter 2 for all hydrographs after the base flow separation. Using the estimated values of D/V and D in Table, the DSFHs were converted to dimensionless hydrographs. Modified hydrographs for PDFs fitting were shown

in Figure 8b. The shapes from four PDF including Lognormal, Beta, Gamma and Weibull distributions will be used for fitting.

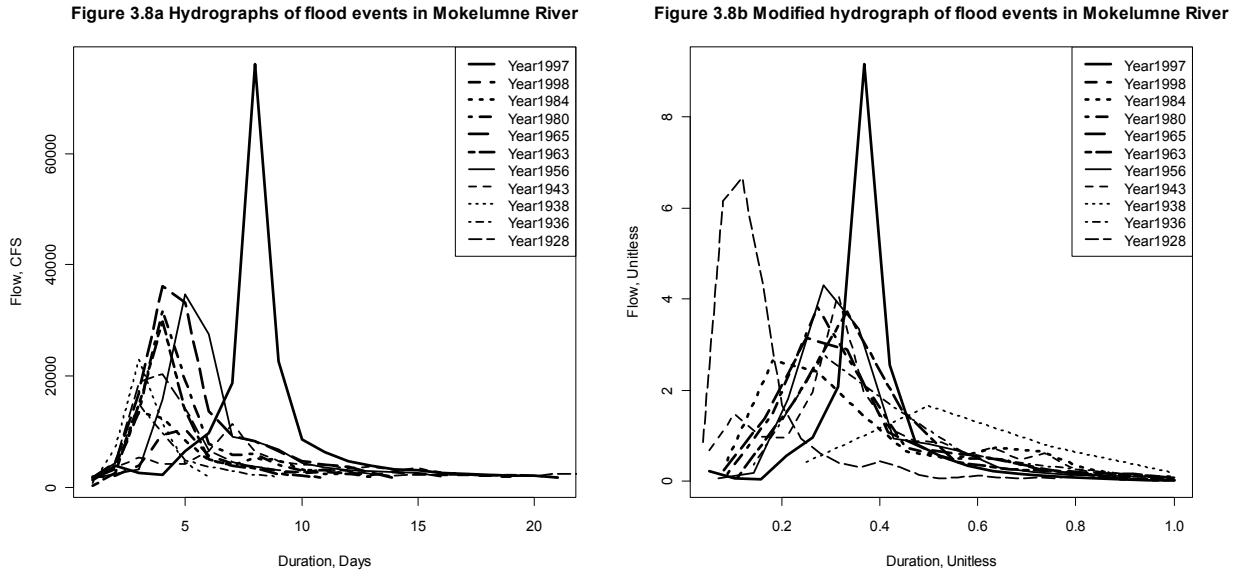


FIGURE 3.8 (A-B) ORIGINAL AND MODIFIED HYDROGRAPHS OF FLOOD EVENTS IN MOKELUMNE RIVER

Shapes of PDFs fitting

Lognormal, Beta, Gamma and Weibull distribution forms were used to fit the shape each modified hydrographs. An R program was employed and least square methods were used for this fitting. The values of the RMSE and R^2 for all the forms appear in Table 3.1. All three PDFs show comparable and satisfactory results except Lognormal. Two years' flood hydrographs including 1997 and 1965 can not to be fit by Lognormal due to limitations of the iteration algorithm. If considering both minimal RMSE and maximum R^2 as criteria for the best form, Gamma is the highest with number 9 of minimal RMSE and maximum R^2. The second highest is Lognormal with number 7. However, none of R^2 in Lognormal exceeds 0.9 while all other three forms have R^2 greater than 0.9. So the rank of best fit forms is Gamma > Weibull > Beta > Lognormal. The best forms for each year, which maps the shape of the hydrographs closest to the corresponding observed hydrographs, is highlighted with bold font.

Figure 3.9 shows all the hydrographs and fitted hydrograph in four forms for all selected years. The graphs are consistent with the results from Table 3.2. From these figures, Lognormal is most unfit to most observed hydrographs. It appears Gamma is best fitted forms in all other threes.

Table 3.2 Performance indices obtained from fitting the PDF functions with observed hydrographs

Year	Log Normal		Beta		Gamma		Weibull	
	RMSE	R2	RMSE	R2	RMSE	R2	RMSE	R2
1998	**0.1426**	0.7026	0.2356	0.9139	0.2352	**0.9256**	0.2354	0.8739
1997	NA	NA	0.2221	0.9433	0.2218	0.9454	**0.2010**	**0.9581**
1984	**0.03918**	0.7951	0.2935	0.0615	0.2913	**0.8326**	3.0204	0.2033
1980	**0.01282**	0.7557	0.2958	0.9227	0.2951	**0.9328**	0.2970	0.9024
1965	NA	NA	0.2847	0.9122	0.2824	**0.9262**	**0.2813**	0.8923
1963	**0.1056**	0.7796	**0.3279**	0.9708	0.3301	**0.9764**	0.3299	0.9618
1956	**0.0777**	0.6445	0.2612	**0.9343**	0.2612	0.5883	0.2614	0.9147
1943	0.3071	0.5035	0.2169	0.5855	**0.1828**	**0.9414**	0.2096	0.6010
1938	**0.4168**	0.8065	0.5049	0.9030	**0.4520**	0.9834	0.4875	0.9187
1936	**0.1689**	0.8577	0.3807	**0.9067**	0.3749	0.4750	0.3703	0.8598
1928	0.3213	0.6558	0.1987	0.9770	0.1976	**0.9798**	**0.1796**	0.9543
Total Best of fit	**7**		**3**		**9**		**4**	
Rank of Best of fit	**4**		**3**		**1**		**2**	

Figure 3.9a Comparison of observed and fitted hydrograph: 1998

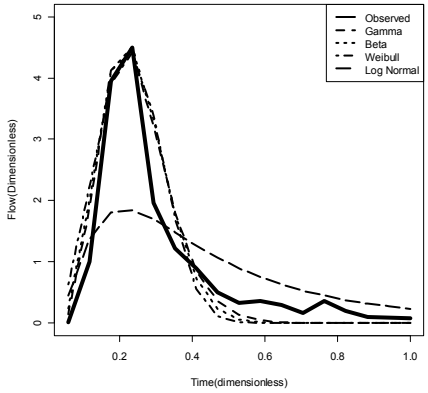


Figure 3.9b Comparison of observed and fitted hydrograph: 1997

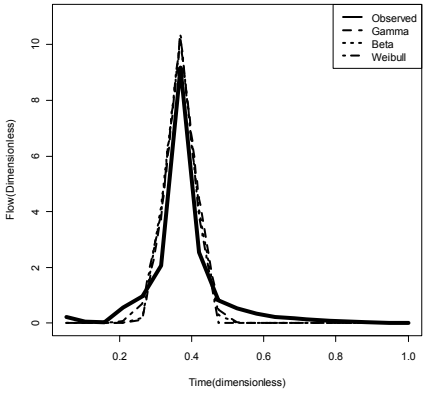


Figure 3.9c Comparison of observed and fitted hydrograph: 1984

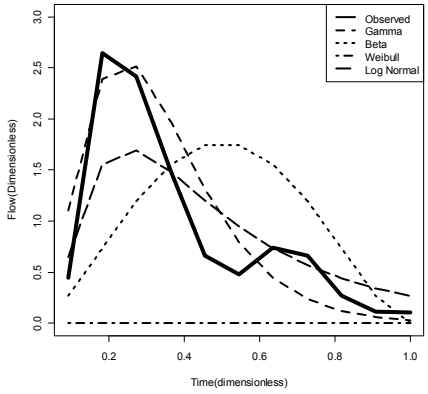


Figure 3.9d Comparison of observed and fitted hydrograph: 1980

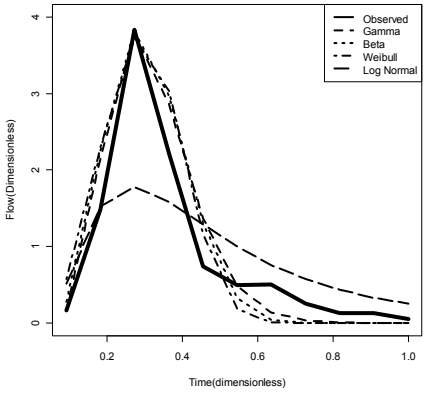


FIGURE 3.9 (A-K) COMPARISON OF OBSERVED AND FITTED HYDROGRAPH FOR DIFFERENT FLOOD EVENTS

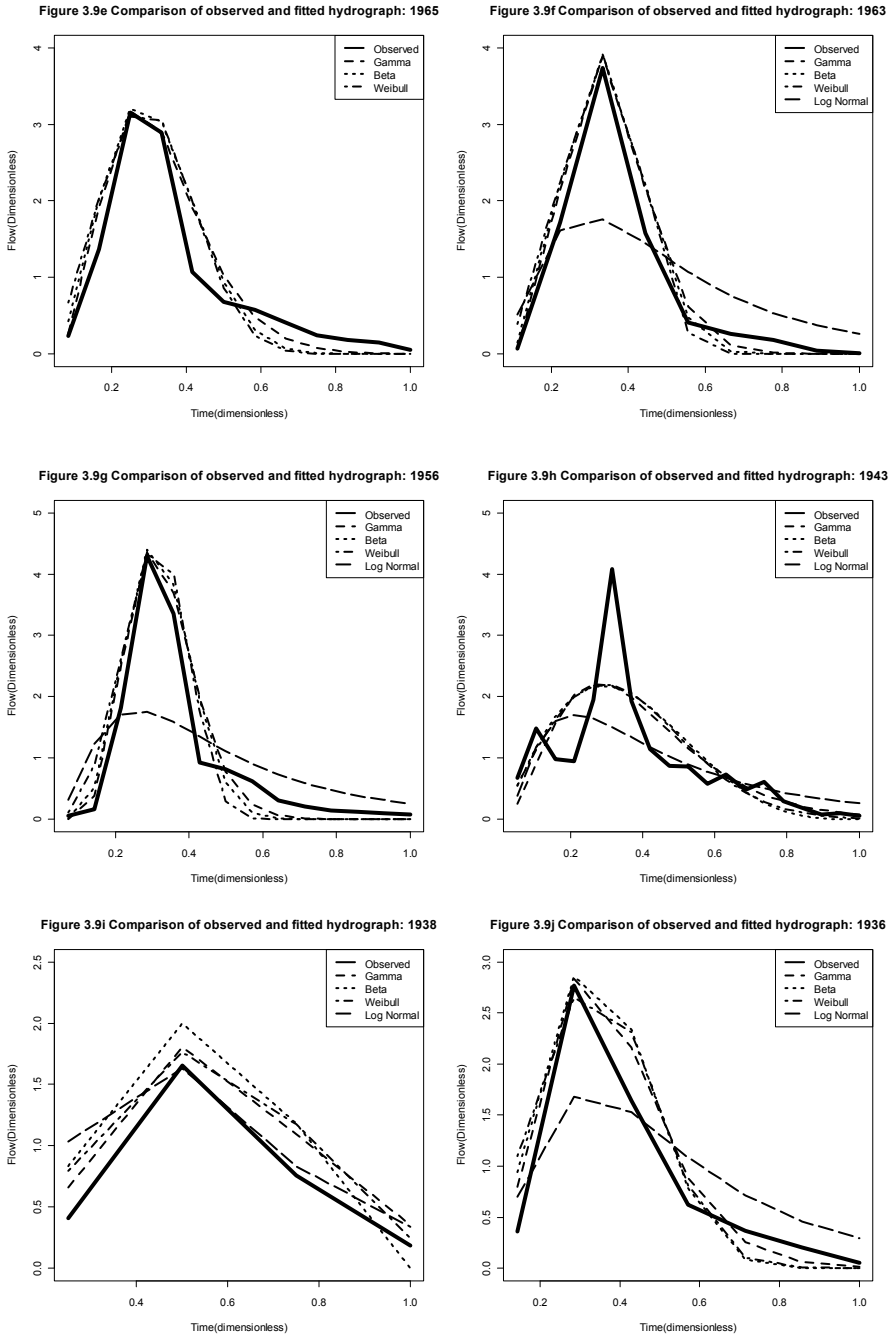


FIGURE 3.9 (A-K) COMPARISON OF OBSERVED AND FITTED HYDROGRAPH FOR DIFFERENT FLOOD EVENTS

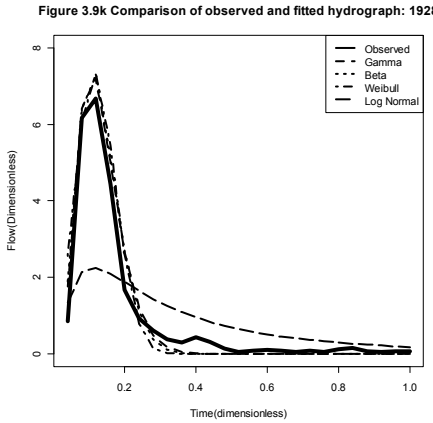


FIGURE 3.9 (A-K) COMPARISON OF OBSERVED AND FITTED HYDROGRAPH FOR DIFFERENT FLOOD EVENTS

Frequency analysis of selected hydrographs parameters to develop of design flood hydrograph

The design flood hydrographs were developed using the design values of the shape and scale parameters of all forms. As per the method described earlier, the shape parameters of Gamma, mean and standard deviation were subjected to frequency analysis. The Lognormal distribution was used for this analysis. Figure 3.10a and 10b show the observed and fitted lognormal distributions of shape mean and shape standard deviation, respectively. From these two figures, one outlier representing year 1997's flood hydrograph appears in normal paper. This is due to the uniqueness of year 1997's hydrograph shape. On Figure 3.7, the 1997's flood event has an over 70,000 CFS peak and less than 10 days duration. This flood event also is called New Year's Day 1997 [DWR, 2007] and due to a pineapple connection storm. This type of storm can cause a sharp spike in hydrograph and has not happened in this watershed historically. Although it's a unique event, the method used in this study still treat it as an effect event, not an outlier.

From this figure, μ_T and σ_T for 10-, 20-, 50-, 100 and 200 year return periods will be determined. Same as μ_T and σ_T , the values of V_T and D_T for10-, 20-, 50-, 100- and 200-year return periods were estimated. Chapter 2 has more detailed results. Using the above computed design parameters of various return periods, dimensionless design hydrographs were constructed. Figure 3.11a presents the developed flood hydrographs shapes of 10-, 20-, 50-, 100- and 200-year return periods using Gamma. The dimensionless hydrographs were then converted to original design flood hydrographs by multiplying V_T/D_T with the ordinates and D_T with the abscissas. Figure 11b presents the developed flood hydrographs of 10-, 20-, 50-, 100- and 200-year return periods using Gamma.

Figure 3.10a Distribution of shape mean on normal paper

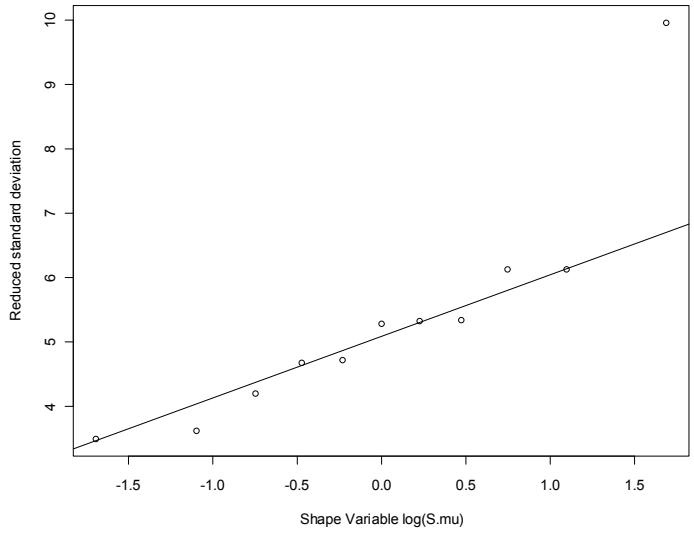


Figure 3.10b Distribution of shape standard deviation on normal paper

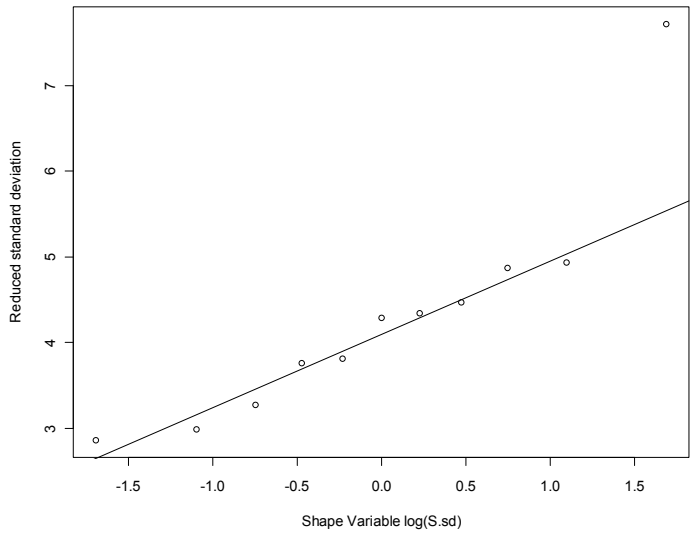


FIGURE 3.10 (A-B) DISTRIBUTIONS OF SHAPE MEAN AND STANDARD DEVIATION ON NORMAL PAPER

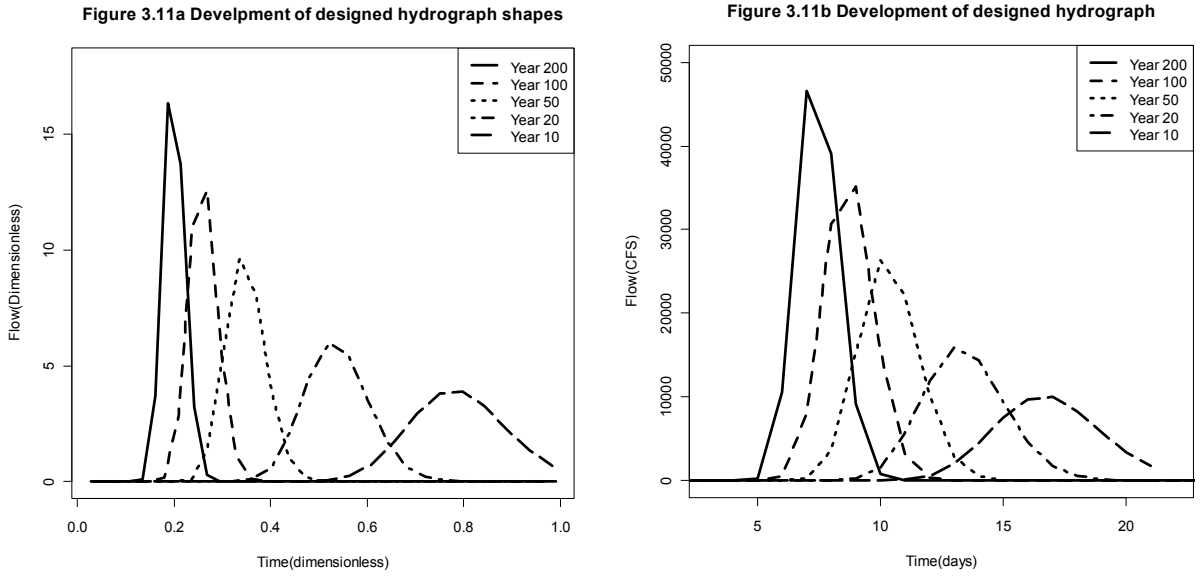


FIGURE 3.11 (A-B) COMPARISON OF DESIGNED FLOOD HYDROGRAPHS OF 10, 20, 50, 100 AND 200 YEAR

3.5 CONCLUSIONS AND LIMITATIONS

This chapter introduces a framework to develop design flood hydrographs for reservoir reoperation. Three steps are presented.

1) Flood hydrographs separation and modification: 11 historical flood hydrographs were selected, separated and converted to dimensionless ones;

2) Hydrograph forms fitting and selection: Beta, Gamma. Lognormal and Weibull distributions were selected and compared to fit modified hydrographs shapes based on goodness of fit criteria including RMSE and coefficients of determination.

3) Development of design flood hydrographs: The design shape variables were estimated from frequency analysis and finally, the design flood hydrographs including 10-, 20-, 50-, 100- and 200-year return periods were derived from the combinations of hydrographs shape, flood volume and durations. The Gamma form is most suitable foe Mokelumne River's floods.

Some shortcomings in this study include the independence assumption for shape factors, neglect of multi-peak hydrographs, and underestimation of flood peak values.

1) Assumption of independence: In this study, two shape parameters, mean and standard deviation, are considered independent. There may be correlation between these parameters. Also, flood volume and duration may be positively correlated. In such case, a bivariate joint probability distribution should be applied to represent the joint statistical properties of correlated variables.

2) Underestimation of peak values: Since the purpose of this study is to construct flood hydrographs for reservoir reoperation, the essential parameters of one flood hydrograph are volume and duration. Flood peak values are somewhat ignored in the process of converting of original hydrographs to dimensionless hydrographs making them less suitable for estimating unimpaired hydrographs. More work is needed to adjust flood peak values on the designed flood hydrographs.

REFERENCES

Archer D, Foster M, Faulkner D, Mawdsley H. (2000). “The synthesis of design flood hydrographs. Proceedings Flooding: Risks and Reactions”. CIWEM/ICE Conference, London.

Arnold JG, Allen PM, Muttiah R, Bernhardt G. 2005. Automated baseflow separation techniques. Ground Water 33: 1010–1018.

Bhunya PK, Berndtsson R, Ojha CSP, Mishra SK. 2007. Suitability of Gamma, Chi-square, Weibull, and Beta distributions as synthetic unit hydrographs. Journal of Hydrology 334: 28–38.

Bhunya PK, Berndtsson R, Singh PK, Hubert P. 2008. Comparison between Weibull and gamma distributions to derive synthetic unit hydrograph using Horton ratios. Water Resources Research 44: W04421. DOI: 10.1029/2007WR006031.

Bhunya PK, Mishra SK, Berndtsson R. (2003). “Simplified two-Parameter Gamma distribution for derivation of synthetic unit hydrograph”. Journal of Hydrologic Engineering 8(4): 26–230.

Bhunya PK, Mishra SK, Ojha CSP, Berndtsson R. (2004). “Parameter estimation of Beta-distribution for unit hydrograph derivation”. Journal of Hydrologic Engineering 9(4): 325–332.

California Department of Water Resources, (2007). “Historic rainstorms in California” , California, U.S.A.

Chapman TG, Maxwell A. (1996). “Baseflow separation—comparison of numerical methods with tracer experiments”. 23rd Hydrology and Water Resources Symposium. Institution of Engineers Australia, Hobart, Tasmania: Hobart; 539–545.

Chow, V. T. (1953). ‘‘Frequency analysis of hydrological data with special application to rainfall intensities.’’ Eng. Exper. Stat. Bull. 414, Univ. of Illinois, Urbana, In.

Chow VT, Maidment DR, Mays LW. (1988) “Applied Hydrology”. McGraw-Hill: New York.

Cruise, J.; Singh, V. (1996). “A stochastic model for reservoir flood operation”. Civil.Eng.Syst. Vol19: 141-155.

Ciepielowski A. (1987). “Statistical methods of determining typical winter and summer hydrographs of ungauged catchments”. In Flood Hydrology, Singh VP (ed.). Riedel: Dordrecht, The Netherlands.

DeGroot, M., Schervish, M. (2002).”Probability and Statistics”. Addison-Wesley: New York.

Dooge, J. C. I. (1959). ‘‘A general theory of the unit hydrograph.’’ J.Geophys. Res., 64(1), 241–256.

Eckhardt K. (2005). “How to construct recursive digital filters for baseflow separation”. Hydrological Processes 19: 507–515.

Ergish, N. (2010) “Flood frequency analysis for regulated watershed.” Master thesis in Department of Civil and Environmental Engineering, University of California, Davis.

Gray DM. (1961). “Synthetic hydrographs for small drainage areas”. Proceedings of the ASCE 87: 33–54.

Haktanir T, Sezen N. (1990). “Suitability of two-parameter Gamma distribution and three-parameter Beta distribution as synthetic hydrographs in Anatolia”. Hydrological Sciences Journal 35(2): 167–184.

Haan CT. (2002). “Statistical Methods in Hydrology”, 2nd edn. Iowa State Press: Ames, IA.

Haan CT, Barfield BJ, Hayes JC. 1994. Design Hydrology and Sedimentology for Small Catchments. Academics: New York.

Herrmann A. 1997. In Global Review of Isotope Hydrological Investigations, Oberlin G, Desbos E (eds). FRIEND 3rd Report. UNESCO: Paris.

Hickey, J. et al. (2002). “Synthetic Rain Flood Hydrology for the Sacramento and San Joaquin River Basins”. Journal of Hydrologic Engineering Vol. 7, No.3, May 1, 2002.

Johnson NL, Kotz S. (1970). “Distributions in Statistics: Continuous Univariate Distribution-II” . Wiley: New York.

Koutsoyiannis D, Xanthopoulos T. 1989. On the parametric approach to unit hydrograph identification. Water Resources Management 3:107–128.

Lim KJ, Engel BA, Tang Z, Choi J, Kim K-S, Muthukrishnan S, Tripathy D. 2005. Automated web GIS based hydrograph analysis tool, WHAT. Journal of the American Water Resources Association 41(6): 1407–1416.

Nadarajah S. (2007). "Probability models for unit hydrograph derivation". Journal of Hydrology 344: 185–189.

Nathan RJ, McMahan TA. 1990. Evaluation of automated techniques for baseflow and recession analysis. Water Resources Research 26(7): 1465–1473.

Nezhikhovsky RA. (1971). "Channel network of the basin and runoff formation". Hydrometeorological . Leningrad: Russia.

Pilgrim, D. H., and Cordery, I. (1993). ''Flood runoff.'' Handbook of hydrology, D. R. Maidment, ed., McGraw-Hill, New York.

R Core Team. (2010) "An Introduction to R". Website: http://cran.r-project.org/doc/manuals/R-intro.html

Rai RK, Sarkar S, Gundekar HG. (2008). "Adequacy of two-parameter beta distribution for deriving the unit hydrograph". Hydrology Research 39(3): 201–208.

Rai RK, Sarkar S, Singh VP. (2009). "Evaluation of the adequacy of statistical distribution functions for deriving unit hydrograph". Water Resources Management. 23: 899–929.

Rao AR, Hamed KH. (2000). "Flood Frequency Analysis". CRC Press: Boca Raton, FL; 350.

Rieger WA, Olive LJ. (1984). "The behaviour of suspended sediment concentrations during storm events". In Drainage Basin Erosion and Sedimentation, Loughran RJ (ed.). University of Newcastle, NSW:Australia; 121–126.

Rosin P, Rammler E. (1933). "The laws governing the fineness of powdered coal". Journal of the Institute of Fuel 7: 2936.

Sherman, L. K. (1932). ''Streamflow from rainfall by the unit-graph method.'' Eng. News-Rec., 108, 501–505.

Singh SK. (2000). "Transmuting synthetic unit hydrographs into gamma distribution". Journal of Hydrologic Engineering 5(4): 380–385.

Snyder FF. (1938). "Synthetic unit hydrographs". Transactions-American Geophysical Union 19: 447–454.

Sokolov AA, Rantz SE, Roche M. (1976). "Methods of developing design-flood hydrographs". Flood Computation Methods Compiled from World Experience. NESCO: Paris.

Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E. (1993). ''Frequency analysis of extreme events.'' Handbook of Hydrology, D. R. Maidment, ed., McGraw-Hill, New York.

Weibull, W. (1939). ''A statistical theory of the strength of materials" Proc., The Royal Swedish Institute for Engineering Research, Ingeniors Vetenskaps Akademiens-Handlingar, Stockholm, Sweden, 151.

Yue S, Taha BMJ, Bobee B, Legendre P, Bruneau P. 2002. Approach for describing statistical properties of flood hydrograph. Journal of Hydrologic Engineering 7(2): 147–153.Copyright

Yue, S., and Hashino, M. (2000). ''Unit hydrographs to model quick and slow runoff components of streamflow.'' J. Hydrol., 227, 195–206.

Yue, S., Ouarda, T. B. M. J., Bobe´e, B., Legendre, P., and Bruneau, P. (1999). ''The Gumbel mixed model for flood frequency analysis.'' J. Hydrol., 226, 88–100.

Yue, S., Ouarda, T. B. M. J., Bobe´e, B., Legendre, P., and Bruneau, P. (2000). ''Corrigendum to 'The Gumbel mixed model for flood frequency analysis'' J. Hydrol., 228 - 283.

Explanation of Variables and simplified terms

AD test – Anderson-Darling test

AMS – annual maximum series

CDF – cumulative density function

DFH – design flood hydrograph

DWR – Department of Water Resources

ECDF – Empirical Cumulative Distribution Function

KS test – Kolmogorov-Smirnov test

MSL – mean sea level

PAM – peak amplitude method

PDF – probability of density function

RMSE – root mean square error

SM – statistical method

TH – Typical Hydrograph

UH – Unit Hydrograph

USACE – U.S. Army of Corps of Engineer

USGS – United States of Geological Services

VAM – volume amplitude method

CHAPTER 4 REGULATED FLOOD FLOWS FREQUENCY CALCULATION THROUGH RESERVOIR RE-OPERATION

Not like his father, Gun, who just built levees to block floods, Da Yu constructed channels to lead floods to detention areas.

- Si Ma Qian, Han Dynasty Historian,100 BC

SUMMARY

Reservoirs transform unregulated flow to regulated flow with different operation rules. Regulated versus unregulated flow curves represent results of reservoir flood control operation. However, it is complicated to find an accurate curve due to the complicated physical setting and uncertainties of operations. This chapter introduces a framework to estimate regulated flow frequency for a reservoir's flood storage allocation. Three main steps including unregulated flow frequency analysis, unregulated/regulated flow transformation and regulated flow frequency estimation are presented. The main contributions include separating flood pulses from daily inflow time series by base flow criteria, modification of unregulated flow calculations, and fitting unregulated and regulated flow to appropriate probability distributions. Unregulated versus regulated flow curves are found using USACE's ResSim software. The statistical package R software is used to fit probability distributions and perform statistical inference. Camanche and Pardee reservoirs in Northern California are used as example applications.

4.1 INTRODUCTION

This chapter discusses estimation methods for regulated flow frequency for major flood control reservoirs with reallocation of reservoir flood storage. In most flood reservoir protected areas, such as Central Valley of California, regulated flow frequency curves are a basis of flood risk analysis for floodplain planning and management. Different reservoir flood storage volumes are key parameters for regulated flow frequency. This chapter presents a framework to estimate regulated flow frequency for a reservoir's reoperation rules.

First, unregulated flow frequency should be quantified with reasonable accuracy through mathematical models. Numerous papers examine flood flow frequency distributions which reflect the likelihood or probability of flood events [Stedinger, et al, 1993]. With a sufficiently and stationary long record of flood flows, a frequency distribution for a site could be accurately described. Much research has been devoted to developing generally applicable methods for estimating unregulated frequency curves [Wurbs, 1996]. The federal government recommends methods for estimating unregulated frequency curves at gage locations [IACWD, 1982].

However, reservoir regulated flow frequency is rarely discussed in literature. This is mainly due to the site and operation specific nature of regulated flow. Regulated frequency curves are functions of the at-site storage characteristics of the reservoir and the operational characteristics of the dam outlet works [Goldman, 2001]. Factors include the duration of flood volumes to determining the peak flows, antecedent conditions prior to major events, the relationship between regulated and unregulated flow values, operational contingencies and interagency and community involvement. The problem becomes extremely complex if a water control system involves multiple reservoirs, diversions and levees.

To describe uncertainty in reservoir discharge, the U.S. Army of Corps of Engineers (USACE) developed an inflow and outflow frequency procedure. Given an inflow peaks, outflow peaks were calculated based on stated reservoir operation rules [USACE, 1996]. For the same purpose, Goldman developed a basic procedure to find regulated frequency curves and applied in Folsom Dam at American River, California [Goldman, 2001]. The basic steps are as follows (see Figure 4.1): 1) Develop an unregulated flow record based on routing studies; 2) Estimate volume-duration-frequency curves describing the likelihood of reservoir inflow volumes from the unregulated period of record; 3) Estimate the critical duration or characteristic time (CT) for flood inflows based on routing studies or by examining the period record; 4)

Develop a relationship between regulated peak outflow and unregulated inflow volumes for the critical duration identified in step 3; 5) Combine the volume-duration-frequency curve estimated for the critical duration with the regulated versus unregulated relationship to obtain the regulated frequency curves.

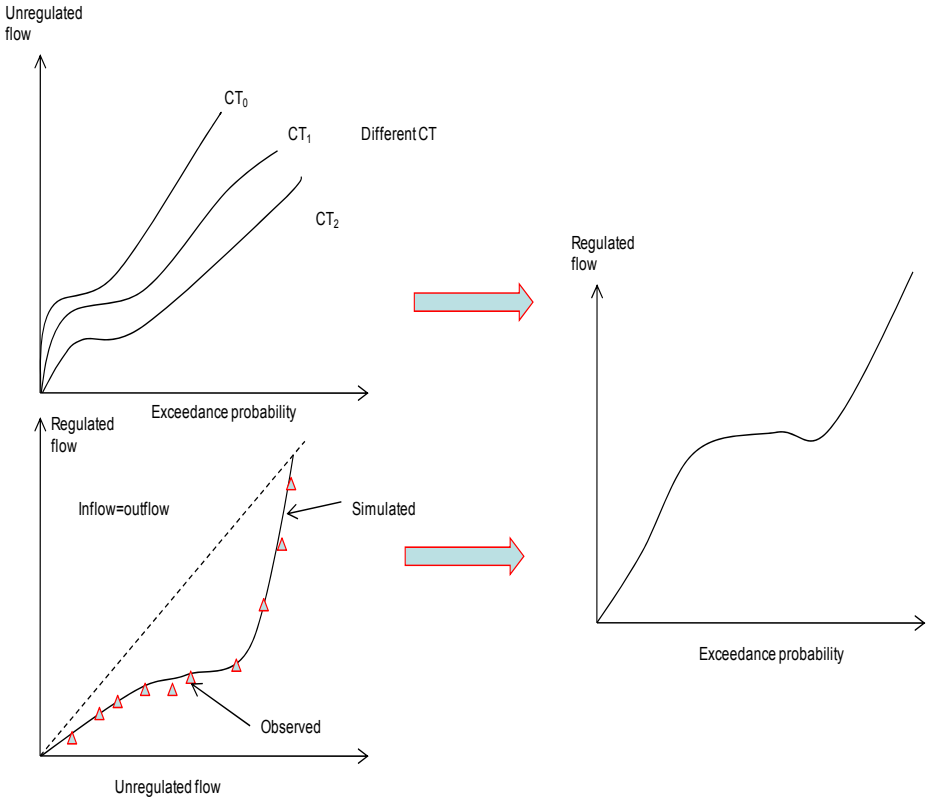


FIGURE 4.1 TRADITIONAL METHOD FOR ESTIMATING REGULATED FREQUENCY CURVE (USACE, 1996)

These steps result in an estimate of the regulated flood frequency curve. However, one big issue, the pre-defined characteristic time (CT) or flood duration in step 3, exists in this approach. To calculate the peak flow, USACE and Goldman used a Volume Duration Frequency (VDF) curve to construct an unregulated inflow's exceedence probability curve [USACE, 1996, Goldman, 2001]. VDF analysis involves estimating the frequency distribution of consecutive n-day annual maximum stream flow volumes. However, the selection of the value of n for n days of CT is complicated. Goldman used 3 days as CT for Folsom Reservoir's flood critical duration in California, and used 60 days as CT in Saylorville Reservoir in Iowa State. Typically CT is affected by peak volumes of high flows, flood control volume in reservoirs, outlet capacities (gates and spillway) and downstream channel capacity. The main disadvantage of the CT method is the rigidity of flood duration. If 3 days duration is selected in reservoir operation, the largest 3-days flood volume in a dry year may be smaller than even one day's flood in a wet year, or so small that calling them floods is misleading. To overcome this issue, other methods such as a base flow method will be considered to separate flood pulse from the inflow time series.

Another modification to Goldman's approach is to test how flood storage reallocation could affect regulated flow frequency. Based on step 4, the relationship will be constructed between unregulated flow and regulated flow for different levels of flood storage. Figure 4.2 shows the flow chart of a modified method for estimating the regulated frequency curve in this study. In figure 4.2, the unregulated flow PDF and unregulated/regulated flow relationship curves are calculated first. Then the regulated flow PDF will be established.

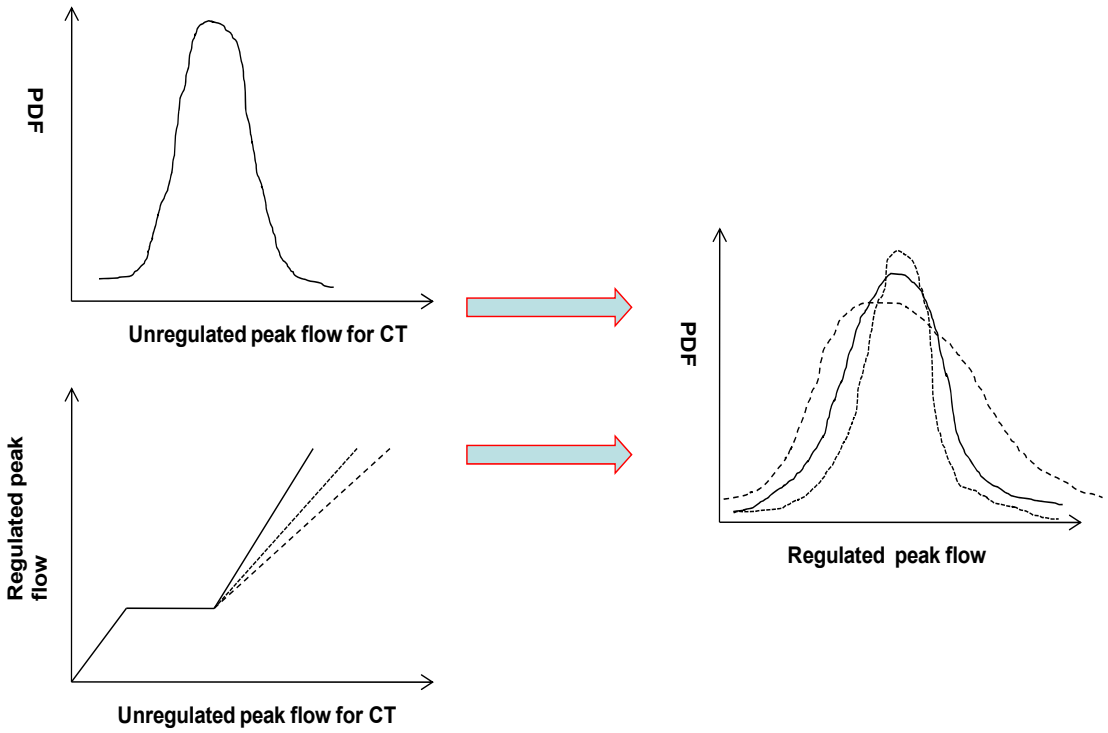


FIGURE 4.2 MODIFIED METHOD FOR ESTIMATING REGULATED FREQUENCY CURVE

This chapter is organized as follows. Section 2 presents some basic concepts and background of regulated flow frequency calculation. A description of the method is given in section 3. Section 4 presents the application for Carmanche/Pardee reservoirs on the Mokelumne River in Northern California. A summary and conclusions are given in section 5.

4.2 LITERATURE REVIEW AND CONCEPTS DEVELOPMENT

This section presents some basic concepts from the literature on unregulated frequency determination, unregulated and regulated flow relationship curve construction and some statistical methods used.

4.2.1 Unregulated flow frequency analysis

Unregulated flow frequency is the first step in this study. Some literature reviews and concepts development on general flow frequency calculation are described as follows.

- Physical or statistical models to simulate flood pulses

In reservoir operation studies, synthetic stream flow hydrographs are common to simulate historical flow time series. To generate representative inflows, various empirical, mathematically/physically based, mathematically/stochastically based, analog/physically based, and physically/laboratory-scale models and approaches have been proposed and developed [Salas, et al, 2003]. From all these approaches, there are two basic methods to generate hydrographs, precipitation/runoff modeling and historical flow analysis. Precipitation-runoff modeling, also called watershed modeling, simulate hydrologic processes by which precipitation is converted to stream flow [Linsley et al., 1982; McCuen, 1989; Singh, 1992; Bedient and Huber, 1992]. Historical flow analysis generates stream flows from historical observations using statistical inference such stochastic methods or frequency analysis of flood flows. In stochastic methods, time series theory through as auto regressive and moving average (ARMA) has been used to generate longer time series the inflows [Salas, et al, 2003]. Regression models are applied to analyze flood frequency.
In precipitation-runoff modeling, the watershed is modeled, with precipitation hyetograph input and runoff output. The source of water is rainfall or snow melt. Most precipitation is lost through the natural

interception, depression storage, infiltration, evaporation, and transpiration. The remaining precipitation flows overland and through the soil, collects as flow in swales, small channels, and eventually becomes runoff to stream. Groundwater also contributes to stream flow, largely independent of a particular precipitation-runoff event. Land use, drainage improvements, storage facilities, and other development activities significantly affect how precipitation is converted to stream flow. River basins are divided into smaller, more hydrologically homogeneous sub-watersheds for modeling purposes. The runoff hydrographs from individual sub-watersheds are routed through stream reaches and combined at appropriate locations. Such modeling can overcome shortages of stream flow records and provide flexibility to examine changes of conditions and locations. Also, in support of real-time reservoir operations, watershed models are used to forecast stream flows expected in the near future based on current precipitation measurement [Bedient & Huber, 1992].

Statistical methods to generate flows based on stream gage data are often used, as long as enough flow records exist. Loucks et al [1981], Linsley et al [1982], and Salas [1993] provide concise overviews of synthetic streamflow generation. Bras and Rodriquez-Iturbe provide an in-depth theoretical treatment of stochastic hydrology The most common stochastic model utilizes lag-1 autoregressive Markov model for synthesizing monthly flows for a single site [Bras and Rodriquez-Iturbe, 1985]. To model short time step flows, many methods based on time series theory have been applied. However, these methods are employed mostly in academic studies.

Recently, some more convenient methods based on flow statistics for design flood hydrographs were developed. Hickey, et al. [2000] constructed design flood hydrographs for the Central Valley, California by frequency analysis. This method translates frequencies to hourly flood hydrographs for use in reservoir simulations. Three steps are involved: 1) Obtain the average flood flow rates from unregulated frequency curves; 2) separate these average flow rates into flood volumes; and 3) distribute volumes into storm time series.

In this study, to estimate flow frequency, Hickey's method will be modified and applied to analyze flood pulses.

- Annual Maximum Series (AMS) and Partial Duration Series (PDS)

Two approaches exist to catch or separate flood episodes from historical daily inflow time series. One is Annual Maximum Series (AMS) which is constructed by selecting the annual maximum value of each year, i.e., only one event per year is retained. In the AMS approach, the individual peak is regarded as identically and independently distributed (IID). The other approach is Partial Duration Series (PDS). PDS consists of all values that exceed a specified threshold, i.e., the PDS approach is not confined to only one event per year and allows additional large events to be considered [Langbein, 1949; Stedinger, 1993].

PDS has been recognized in the past 30 years because it considers both the largest event in each year other large events. Moreover, the second largest event in many years usually exceeds the largest event of some other year. Also, the largest annual flood flow in a dry year in some arid or semiarid regions may be zero, or so small that calling them floods is misleading [Stedinger et al, 1993]. The classical PDS model assumes a Poisson distributed number of threshold exceedance and independent exponentially distributed exceedance magnitudes (PDS/ESP) [Shane and Lynn, 1964; Todorovic and Zelenhastic, 1970]. Many arguments favor PDS over AMS. If the arrival rate for peaks over the threshhold is large enough, a PDS estimator contains more than 1.65 exceedances on average per year, and if the return periods are more frequently than about 20 years, the PDS estimator has a smaller variance than the AMS estimator [Cunnne, 1973]. In recent years, the generalized Pareto (GP) distribution has replaced the exponential distribution to quantify exceedance magnitude [Madsen et al, 1994]. A PDS model is preferred in most cases [Madsen, 1997].

PDS methods have been applied in many areas in water resources area, especially in flood reservoir operation research. Todorovic [1978] derived the mathematical expression of PDS models and found the theoretical results agreed well with observed flood records. Cruise and Singh [1990] used a PDS model to model inflow peak, duration and inter-arrival times for flood reservoir analysis. Xu and others [1997] applied a PDS model to inflow for Sanba reservoir in China and simulated a 2,000 year time series.

Nevertheless, the PDS approach has been much less applied in hydrologic studies than AMS methods, mainly due to difficulties in defining the PDS. First, one must model the arrival rate and magnitudes of events larger than the threshold level. For large return-period events like floods in California, the actual probabilistic model for arrivals is not important, provided different models yield the same average number of arrivals per year. Also, in PDS models, consecutive peaks should be independent, and hence some criteria to identify independent events must be defined. For instance, if multiple peaks correspond to the same hydrologic event, only the largest peak should be included in the PDS. Last, the method involves selection of an appropriate threshold level, that is, a level which ensures as much relevant information as possible in the analysis without violating basic statistical assumptions [Madsen, 1997].

General relationships exist between the probability distribution for AMS and the frequency of events in a PDS [Stedinger, 1993]. Cunnane (1987) found AMS and PDS are similar for long-term time series. In this study, AMS was used to estimate inflow frequency.

- Separating floods from flow time series and inflow calculation

Separating floods from flow time series is very important in reservoir flood control. Figure 4.3 shows a typical flood including peak rate, flood volume, and flood duration. A flood hydrograph consists of two flow components, a peak component due to surface runoff/snowmelt and base flow. The shape of a flood hydrograph is dominated by the peak component. These two parts have to be distinguished. The separation requires identification of three features of the total hydrograph: the start and end of peak part and the shape of the base flow part. Generally, time boundaries of a flood are marked by a rise in stage and discharge from base flow and a return to base flow. For flood reservoir studies, the start of peak is usually marked by the abrupt rise of the hydrograph. The end of the peak is identified by the flattening of the recession limb of the hydrograph. Due to the different characteristics of the peak part and the base part, there exists a significant change in the slope of hydrograph as the transition occurs from peak to base. There are two criteria to separate flood pulse from time series: the critical duration method and the base flow method.

USACE has used a critical duration method to separate floods in flood control analysis. Critical duration is determined by peak volumes of high flows, flood control volume in reservoirs, outlet capacities (gates and spillway) and downstream channel capacity [Goldman, 201]. For example, 3 days is defined as Folsom Reservoir's flood critical duration in California State while 60 days is defined in Saylorville Reservoir in Iowa State. The main disadvantage using CT method is its rigidity of flood duration.

The other method to separate flood pulse separates the portion of hydrograph above a base flow. For simplicity, the base flow is set as a constant flow. Therefore, the start date and end date of peak will be identified as s and e. The duration (D) will be e-s. The flood volume is the area under the hydrograph bounded by e and s.

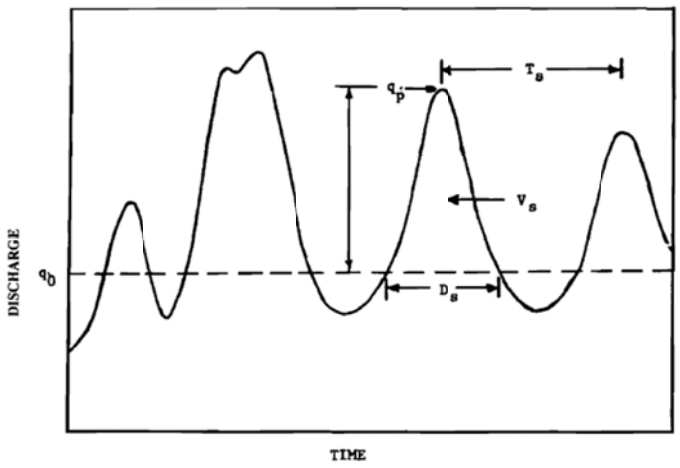


FIGURE 4.3 SEPARATING FLOOD PULSES BY BASE FLOW FROM FLOW TIME SERIES [MAIDMENT, ET AL. 1993]

Several significant articles on flood separation from inflow time series were published from the 1970s to 2009. Todorovic (1978) selected base flow x_0 to compare AFS (annual flood series) and PDS approaches. Singh, et al (1990) also selected base flows to separate flood hydrographs from flow time series to derive the probability distribution of associated flood volumes in a PDS approach. Similar to these two articles, Yue (2002) and Pramanik, et al. (2010) also used base flow to separate flood hydrographs.

After the flood was separated, average flood inflow will be calculated by volume divided by duration. Goldman used a Volume duration frequency (VDF) curve to construct an unregulated inflow exceedence probability curve. VDF analysis involves estimating the frequency distribution of consecutive n-day annual maximum stream flow volumes. [Goldman, 2001] The procedures are: 1) pick the maximum volume with duration of n-days in each year; 2) calculate the flow rate, i.e., Q = V/D, where D=n*24*3600 (seconds); 3) plot normal probability paper.

For flood reservoir operation, three key components of a typical inflow hydrograph are peak rate, volume and duration. The inflow peak is less important for reservoir operation. According to USACE' operating rule guidance, to minimize exceedence of downstream channel capacity, when the inflow is less than the downstream channel capacity, all inflow is released. As inflow exceeds channel capacity, the reservoir begins to store the excess flow, releasing a constant outflow equaling the channel capacity. When the reservoir has reached its storage capacity, the outflow again equals inflow. Therefore the modified inflow, particularly volume and duration, rather than peak flow rate is more important in reservoir flood control [Goldman, 2001].

- Frequency analysis and Flood-volume-frequency analysis

Estimating probabilities of flood flows is routine for reservoir management studies and flood frequency is covered in most hydrology textbooks [Cunnane, 1978; Linsley, Kohler, and Paulhus,1982; Linsley et al., 1992]. McCuen [1989], and Bedient and Huber [1992]. Bulletin 17B of the Interagency Committee on Water Data [1982] specifies flood flow frequency analysis procedures for federal water agencies. The HEC Flood Frequency Analysis (HEC-FFA) program performs the computations outlined in Bulletin 17B [HEC, 1992].

In flood frequency analysis, the annual exceedance frequency or probability (P) is the probability that a specified flow magnitude will be equaled or exceeded in a given year. The recurrence interval or return period (T) is the average interval, in years, between events equaling or exceeding a specified magnitude. Traditionally, probabilities are assigned by a probability distribution function, with parameters estimated from observed data. Most probability distribution functions for continuous random variables can be expressed as

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{K}\boldsymbol{\sigma} \quad (4.1)$$

Where μ and σ are the mean and standard deviation of the random variable X, and K is a frequency factor defined by a specific distribution as a function of the probability level F(x) or P(x) of X. Other common probability distributions are based on a logarithmic transformation, with equation 4-1 expressed as

$$\mathbf{log(X)} = \boldsymbol{\mu}_{\mathbf{logX}} + \mathbf{K}\boldsymbol{\sigma}_{\mathbf{logX}} \quad (4.2)$$

Equation 4-2 is also used for the log-Pearson Type III distribution used by U.S. federal agencies for performing flood flow frequency analyses [Bulletin 17B].

The other way to provide the frequency is plotting position formulas. A plotting position formula is required to plot observed peak flows versus exceedance probability. Various formulas have been applied. The most commonly used is the Weibull formula [Cunnane, 1978; Linsley et al, 1982; McCuen et al, 1989; Bedient and Huber, 1992; Sigh, 1992].

$$\mathbf{P} = \mathbf{m}/(\mathbf{N}+\mathbf{1}) \quad (4.3)$$

$$\mathbf{T} = (\mathbf{N}+\mathbf{1})/\mathbf{m} \quad (4.4)$$

Where N is the number of years of observation and m is the rank of an event's magnitude, with the largest annual peak flow having m=1. Plotting position formulas provide a visual display of the closeness

of fit of the analytical probability distribution to the observed data. However, plots of observed data should not be extrapolated to estimate infrequent events. Estimates of exceedance probability assigned by a plotting position formula to the largest floods in the observed data can be highly inaccurate.

To acquire the best-fitted PDF shapes to design flood hydrographs by statistical fitting techniques, Yue predefined the shape of flood hydrograph as a two-parameter beta PDF [Yue et al, 2002]. The two parameters are shape mean (S_m) and shape variance (S_v) which correspond to the combination of flood peak, volume and duration. Based Yue's findings, Pramanik adopted four PDF distributions including two parameter beta, Weibull, Gamma and lognormal to fit hydrograph shape [Pramanik et al 2010]. The conclusion is that the ranking of the PDFs based on estimation of peak of design flood hydrograph for 50-, 100- and 200-year return periods have the order: Weibull > Beta > Lognormal > Gamma. Chapter 3 of this dissertation introduces a framework to develop design flood hydrograph based on best fitted PDF shapes and to find the best fitted PDF is the Gamma distribution shape in Mokelumne River watershed.

4.2.2 Statistical methods to fit flow probability

Selecting an appropriate probability distribution is important for unregulated flow frequency analysis. Three steps are included in fitting a distribution: model/function choice, estimating parameters, and goodness of fit tests. R program package is used due to it's power and flexibility for statistical computing and graphics.

- Model choice

The first step in fitting inflow distributions is choosing the mathematical model or function to represent observed data. Normally a histogram/PDF of raw data is plotted and compared to candidate distributions. However, graphics & histograms can be quite subjective. It's easy to identify the location and variability characteristics of inflow data in a graph. But the skewness and kurtosis and general shape of the extremes are hard to determine. Skew is a measure of symmetry, or rather, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center point. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution [DeGroot & Schervish, 2003].

Common candidate distributions in hydrology include Normal, Log Normal, Gumbel, Weibull, Log-Pearson III, Gamma and Beta distribution [Stedinger, 1993]. In this study five distributions are compared for KS values. The distributions are Normal, Lognormal, Weibull, Gumbel, and Gamma. These five probability density function (PDF) formulas are as follows, respectively:

$$\mathbf{f(x)} = \frac{\mathbf{1}}{\sqrt{\mathbf{2\pi\sigma^2}}}\mathbf{exp}\left(\frac{\mathbf{-(x-\mu)^2}}{\mathbf{2\sigma^2}}\right) \quad \textbf{(Normal)} \qquad (4.5)$$

$$\mathbf{f(x;\mu,\sigma)} = \frac{\mathbf{1}}{\mathbf{x\sigma\sqrt{2\pi}}}\mathbf{exp}\left(\frac{\mathbf{-(lnx-\mu)^2}}{\mathbf{2\sigma^2}}\right),\ \boldsymbol{x} > 0 \quad \textbf{(Log Normal)} \quad (4.6)$$

$$\mathbf{f(x;\lambda,k)} = \frac{\mathbf{k}}{\boldsymbol{\lambda}}\left(\frac{\mathbf{x}}{\boldsymbol{\lambda}}\right)^{\mathbf{k-1}},\ \boldsymbol{for\ x} \geq \mathbf{0} \qquad \textbf{(Weibull)} \qquad (4.7)$$

$$\mathbf{f(x;\mu,\beta)} = \mathbf{e}^{-\mathbf{e}^{(\mu-x)/\beta}} \qquad \textbf{(Gumbel)} \qquad (4.8)$$

$$\mathbf{f(x;k,\theta)} = \mathbf{x}^{\mathbf{k-1}}\frac{\mathbf{e}^{-\mathbf{x}/\boldsymbol{\theta}}}{\boldsymbol{\theta}^{\mathbf{k}}\boldsymbol{\Gamma}(\mathbf{k})},\ \textbf{for } \mathbf{x} > 0\ and\ k,\theta > 0 \quad \textbf{(Gamma)} \quad (4.9)$$

- Parameters estimation

After choosing a model that can mathematically represent the data, parameters for the model must be estimated. There are many estimation methods, but the common methods in hydrology are anagogic, moment generation, and maximum likelihood.

Analogic methods estimate model parameters using empirical data for each parameter, i.e., we estimate the unknown mean of a normal population using the sample mean. This method is fast and simple, but it is only good for an assumed normal distribution and the parameters are required only for location and variability. No probability distribution is required.

The method of moments (MOM) estimates parameter values based on matching the sample moments with the corresponding distribution moments. This method equates sample moments to population (theoretical) moments. When moment methods are available, they have the advantage of simplicity.

The method of maximum likelihood estimate (MLE) is often used in statistical inference to estimate parameters in hydrology. For a known PDF f(x, θ) describing a population, x is a random variable and θ are unknown parameter values according to sampling data, x_i (i=1,2,3…n). MLE begins with the mathematical expression known as a likelihood function of the sample data given this distribution and this expression contains the unknown parameters. The parameter values that maximize the sample likelihood are known as the maximum likelihood estimates. In this study, MLE is selected to estimate parameter values.

- Goodness of fit statistical tests

After parameter estimation, the next step of a statistical method is goodness of fit testing. Goodness of fit tests indicate if it is reasonable to assume that a random sample comes from a specific distribution. They are a form of hypothesis testing where the null and alternative hypothesis are: H_0: Sample data are from the stated distribution; H_a: Sample data are not from the stated distribution (Also see Chapter 3 in this dissertation for KS test) These tests are sometimes called a omnibus test and they are distribution free, meaning they don't depend on a particular PDF. Null hypothesis H_0 will be rejected at a significant level of alpha if the calculated testing p-value is lower than alpha. Several methods for goodness of fit test commonly used in hydrology are chi-square, Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests (Chapter 3 in this dissertation). In this study, a KS test is used due to its flexibility.

4.2.3 Unregulated/regulated flow transformation through reservoir operation and reoperation

- Reservoir operation and reoperation

A reservoir operation plan is a set of rules establishing the quantities of water to be stored, released, or withdrawn from a reservoir or system of reservoirs under various conditions [Wurbs, 1993]. Reservoir operating decisions within these rules involve allocating storage capacity and streamflow among water users, managing the risks and consequences of water storage and flooding, generating beneficial uses of water, energy and land resources and managing environmental resources. Classically, reservoir operation is mainly based on the conflicting objectives of maximizing water available for conservation purposes and maximizing empty space available for storing floods to reduce downstream damages.

Typical reservoir storage plans include an inactive pool, a buffer pool, a conservative water supply pool and flood pool [Wurbs, 1993]. A typical reservoir is divided into one or more elevation zones, or pools, illustrated by Figure 4.4. Common practice is to operate a reservoir for either flood control only, water supply only, or a combination of flood control and water supply depending on the zone or storage pool. The water supply and flood-control pools, or volumes are fixed by a designed top of conservation-pool (bottom of flood-control pool) elevation. Conservation pools may serve various purposes, including urban and agricultural water supply, hydropower and recreation which themselves often involve both complementary and conflicting interactions.

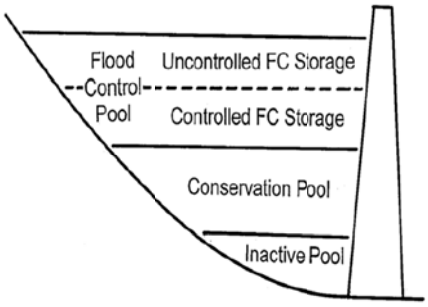


FIGURE 4.4 A RESERVOIR STORAGE ALLOCATION [FORD, D. 1990]

When a reservoir is constructed, water-storage volumes are authorized for different purposes, and operating policies are established to control water releases. Storage allocations and operating policies are sometimes altered to improve water-control management, to adapt to changes in the original purposes, and to accommodate new purposes. This is called reservoir reoperation. Reallocation of storage capacity between flood control and conservation typically involves a long-term or seasonal change in the designated top of conservation-pool elevation. Reallocations between conservation purposes can be achieved by modifications of operation policies.

With rapid urbanization, numerous papers on reservoir re-operation examine storage reallocation to accommodate new demands for water supply and hydropower or flood control [Wurbs, 1990; Duren, 1971]. Recently, greater environmental consciousness has brought concerns for environmental flows downstream of dams [Richter, 2006]. Among these environmental concerns, threatened river-floodplain ecosystems call for reservoir reallocation to restore original floodplains to support downstream ecosystems.

- Unregulated/regulated flow transformation curves

Regulated versus unregulated peak flow curves operationally describe a reservoir flood control system. They represent the transformation from unregulated flood frequency to regulated frequency. These curves need to reflect relationships between the critical inflow flood volume and the peak regulated flow. The latter will be used for downstream flood damage calculations. However, it is complicated to find an accurate unregulated/regulated curve. Indeed, an accurate curve may not exist. The difficulty is not only from a reservoir's physical condition (including reservoir storage, outlet capacity and downstream channel capacity), but also from institutional conflicts and operations.

Physical factors to consider in obtaining a unregulated/regulated relationship are: 1) reservoir storage and the capacity of dam outlet works, 2) constraints on operation for downstream flood protection (i.e., the need to meet the objective release, which is usually an estimated channel capacity), and usefulness of flood of hydrograph forecasts [Ergish, 2010].

Besides the physical factors, the institutional and operational factors also affect the relationships of unregulated and regulated curves. Since most reservoirs serve several purposes including flood control, water supply, hydropower and environmental purposes. There are conflicts among these purposes. The conflicts are mainly how to determine the initial reservoir storage for the flood. Also, operation does not always follow the planned rules [Goldman, 2001]. USACE guidance describes "… In constructing frequency curves of regulated flows, it should be recognized that actual operation is rarely perfect and that release will frequently be curtailed or diminished because of unforeseen operation contingencies." [USACE, 1993].

Here, only flood storage changes are considered, with other factors constant. Figure 4.5 shows a typical outflow-elevation-inflow rule.

- Random variable generation

To construct the unregulated versus regulated flow curves, reservoir operation needs to be simulated with a great number of unregulated flows. Simulation is a process of replicating the real world based on a set of assumptions and conceived models of reality [Ang and Tang, 1984]. To perform simulation, a large number of random numbers following a known probability distribution need be generated. Two algorithms are common to generate random numbers including CDF-Inverse method and Acceptance-rejection method.

CDF-inverse method requires a random variable x to have a cumulative distribution function (CDF) F(x). F(x) is a non-decreasing function with respect to the value x, and $0 \le F(x) \le 1$. Therefore, $F_x^{-1}(u)$ may be defined for any value of u between 0 and 1 as $F_x^{-1}(u)$ is the smallest x satisfying $F_x(x) \ge u$ [DeGroot & Schervish, 2003].

F(x) is strictly increasing functions of x. Hence, a unique relationship exists between F(x) and u. Furthermore, it can be shown that if U is a standard uniform random variable defined over the unit interval [0,1], denoted by U~U(0,1), the following relationship holds:

$$\mathbf{X} = \mathbf{F_x^{-1}}(\mathbf{U}) \qquad (4.10)$$

Where X is target random variable. Since the CDF-inverse method requires an explicit expression between X and U, X can be obtained analytically from the generated U. It is efficient to generate random variables with inverse forms of distributions including exponential, uniform, Weibull, and Gumbel [Tung, 2006]. However, for unavailable CDF inverse distributions such as normal, lognormal and Gamma, the CDF-inverse method may not be convenient.

Acceptance-rejection (AR) methods overcome the shortcoming of the CDF-inverse method. It only requires the random variable's specified PDF $f_x(x)$. This method replaces the original $f_x(x)$ by an appropriate PDF $h_x(x)$ from which random variate can be produced easily and efficiently. The generated random variate from h(x), then, is subject to testing before it is accepted as one from the original $f_x(x)$ distribution. In AR method, the PDF $f_x(x)$ from which a random variate x to be generated is represented, in terms of $h_x(x)$ by $f_x(x) = \varepsilon h_x(x)g(x)$, which $\varepsilon \geq 1$ and $0 < g(x) \leq 1$. $\varepsilon \geq 1$ is chosen such that $\psi(x) = \varepsilon h_x(x)g(x)$ over the sample space of the random variate X. The problem then is to find a function $\psi(x) = \varepsilon h_x(x)$ such that $\psi(x) \geq f_x(x)$ and a function $h_x(x) = \psi(x)/\varepsilon$, from which random variates are generated. The constant ε that satisfies $\psi(x) \geq f_x(x)$ can be obtained from

$$\varepsilon = \max\left[\frac{f_x(x)}{h_x(x)}\right] \qquad (4.11)$$

The AR method can be very fast compared with CDF-inverse method for distribution models lacking analytical forms of CDF inverse. This approach has been applied to Gamma distribution resulting in extremely simple and efficient algorithms [Dagpunar, 1988].

4.3 METHODS

In this study, the procedures include three major steps. They are: Step 1: Unregulated flow probability distribution estimation; Step 2: Unregulated versus regulated flow relationship curves calculation and Step 3: Regulated flow frequency calculation. The errors will be introduced into the curves above.

In first step, four sub-steps are involved: 1) Extract flood pulse annual maximum series volume data from inflow time series. Two approaches are available to separate flood hydrographs, base flow approach and characteristic time approach. 2) Calculate the average flood flow, volume divided by duration (V/D) of each flood pulse; 3) Fit the inflow (V/D) for different probability distributions. After data transformation, Normal, lognormal, Weibull, Gumbel, and Gamma form distributions were used; and 4) lastly, select the appropriate probability distribution based on goodness of fit statistical tests techniques.

In the second step, two sub-steps are involved including using ResSim software to simulate outflow and construct unregulated versus regulated flow curves. They are: 1) Simulate peak flood outflow using Ressim software under different reservoir storage scenarios; and 2) Construct regulated flows versus unregulated flows relationships and characterize errors in this curve.

Finally, in step three, regulated flow frequency was determined by three sub-steps. They include 1) Generate unregulated inflow following the selected probability distribution; 2) Calculate corresponding regulated flow based on Step two's curves and reintroduce error in the curve; and 3) Estimate regulated flow frequency. Figure 4.6 shows the flow chart of this procedure.

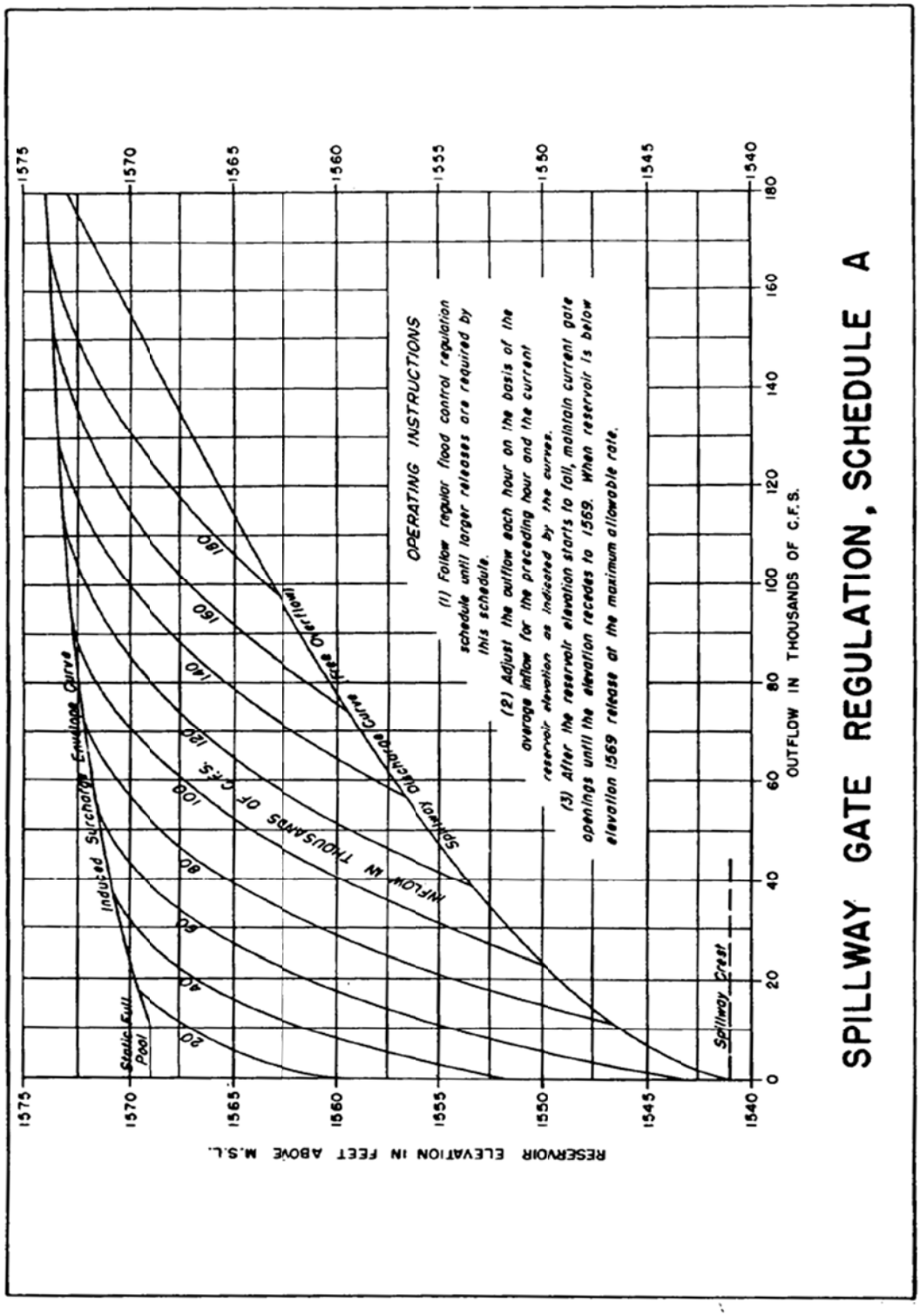


FIGURE 4.5 TYPICAL RESERVOIR OPERATION FOR INFLOW/OUTFLOW RELATIONSHIPS [USACE, 1993]

Next section presents an application for this procedure. In this section, base flow method is used to separate flood pulse. Q_b = 2,000, 3,000, 4,000 and 5,000 CFS were selected and compared. For comparison purpose, Section 4.5 present different characteristic time, i.e., flood duration including 3 days, 5 days, 10 days and 20 days to separate flood pulses. All the other steps in the procedure are identical.

4.4 APPLICATION IN CAMANCHE/PARDEE RESERVOIRS

4.4.1 Site, data and software description

This study uses the Mokelumne River, a major tributary of the San Joaquin River in California as an application. Two reservoirs, Pardee and Camanche on Mokelumne River are re-operated. Constructed in 1927 and 1963 respectively, Pardee and Camanche reservoirs provide storage for flood control, water supply, irrigation, hydropower and recreation and downstream environmental purposes. The maximum flood control space is about 200,000 acre-feet (200 TAF) during winter months in Camanche.

The Mokelumne River watershed is divided into an upper watershed and lower watershed by Camanche/Pardee reservoirs. The reservoirs on the upper Mokelumne River are only for water supply and commercial hydropower purposes, without flood space. During winter, inflow to Pardee reservoir is regarded as natural flow. The watershed below Pardee and Camanche, is mainly agriculture land includes over 70,800 acres of cropland and 60,300 acres of orchards and vineyards protected by levees and flood control space. The communities of Lodi, Woodbridge, Lockford and Clements are also on the lower Mokelumne River.

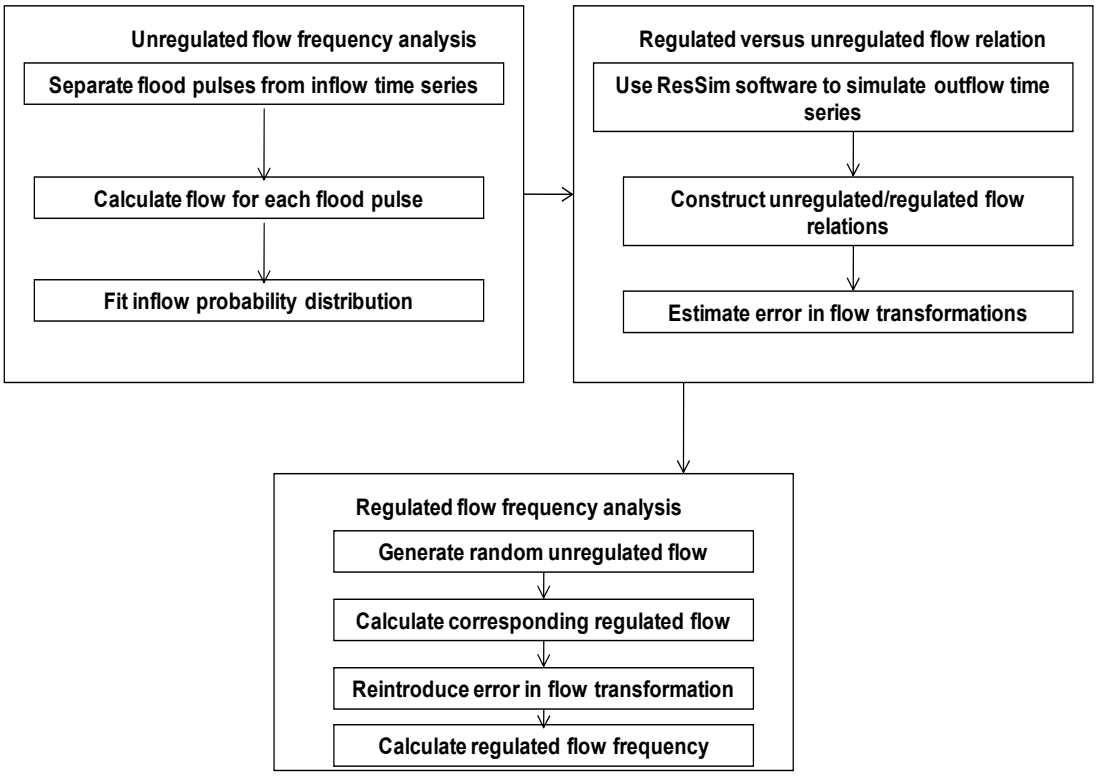
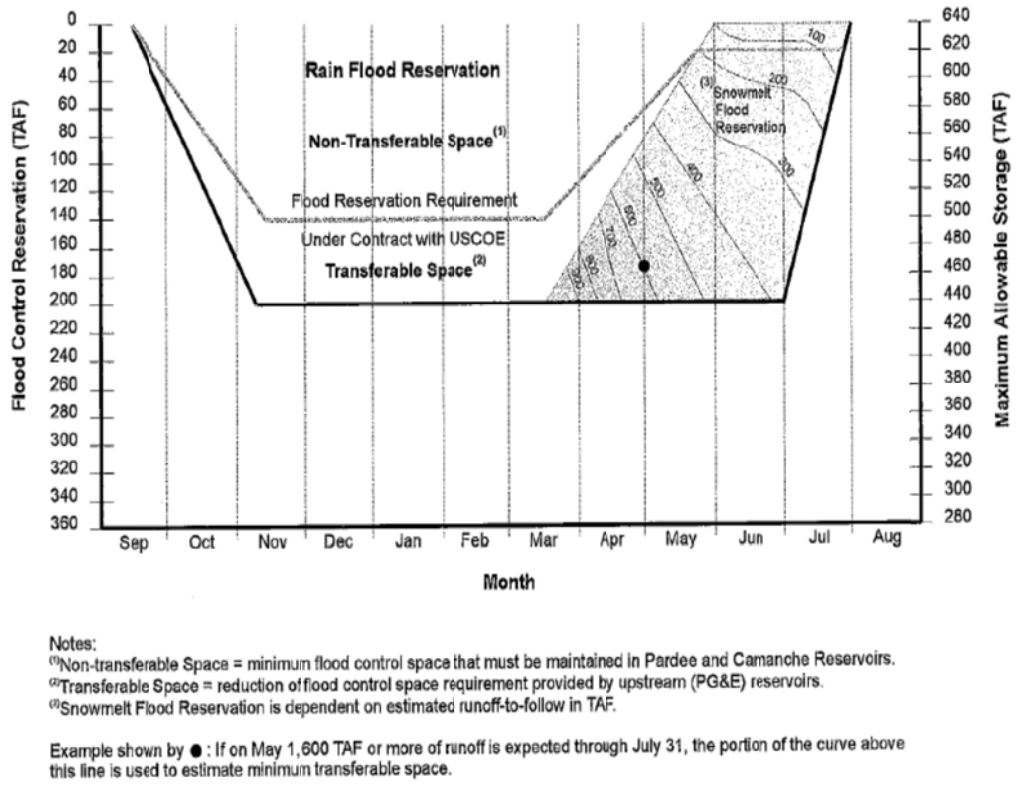


FIGURE 4.6 REGULATED FLOOD FLOW FREQUENCY ANALYSIS FLOW CHART

During winter, inflows to Pardee and Camanche are from rainfall and snowmelt. Flood storage is reserved from Oct.15 to end of next June. The reservoir flood control system has worked well. However, ten major flood events have occurred in the past 50 years.

Required data for the analysis includes hydrologic data, physical setting, operation data and economic data. The hydrologic data is mainly daily time series flow data on both upstream and downstream of the reservoirs. Physical setting is the reservoir's geometric data including size, storage, elevation and channels' geometric shape. Operation parameters include the flood control guide curve, operation rule. Figure 4.7 shows Camanche/Pardee Reservoir Operation Rule for flood control [USACE, 2006]. Figure 2.4 from Chapter 2 shows the inflow time series. USACE's HEC-ResSim was used to simulate reservoir operation to calculate outflow [USACE, 2009]. R program was used for statistical analysis [R Core Team, 2010].



Notes:
(1)Non-transferable Space = minimum flood control space that must be maintained in Pardee and Camanche Reservoirs.
(2)Transferable Space = reduction of flood control space requirement provided by upstream (PG&E) reservoirs.
(3)Snowmelt Flood Reservation is dependent on estimated runoff-to-follow in TAF.

Example shown by ● : If on May 1,600 TAF or more of runoff is expected through July 31, the portion of the curve above this line is used to estimate minimum transferable space.

FIGURE 4.7 CAMANCHE/PARDEE RESERVOIR OPERATION RULE FOR FLOOD CONTROL [USACE, 2006]

The re-operation plans examined reduce the flood storage to various levels. Figure 4.8 shows Camanche/Pardee Reservoir Re-operation Rule. The winter flood space was reduced to 180, 160, 140, 120 and 100 TAF.

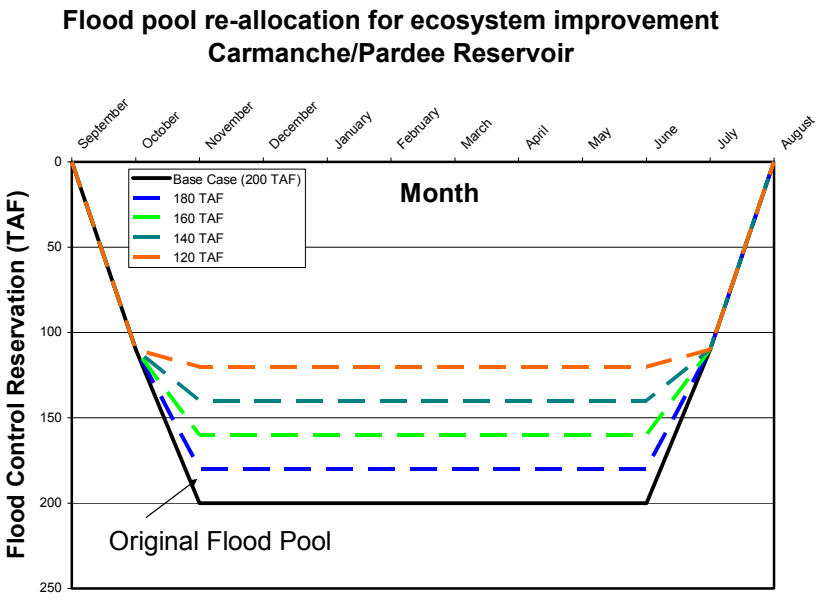


FIGURE 4.8 RESERVOIR REALLOCATION (BASED ON USACE 2006)

4.4.2 Selection of Unregulated Flow Probability Distribution

This step finds the most appropriate probability distribution of unregulated flow. Three sub-steps include separating flood pulse based on base flow, flood volume, duration and inflow comparison and fitting probability distribution are presented as below.

- Separating flood pulse from different base flow

Annual maximum flood pulses series will be separated from daily inflow time series. In this study, 2,000 to 5,000 CFS with 1,000 CFS increment was selected as base flow criteria. In the selected flood pulses, there are 12 events with multi-peaks. In this study, all multi-peaks flood pulses are treated as single peak floods because the modified inflow, i.e., V/D will be calculated to determine the regulated flow's probability distribution.

- Flood volume (V), duration (D) and inflow (Q_{in}) calculations.

After separating the annual 86 flood pulses from 86 years of recorded time series of inflow, statistical methods were employed to analyze their flood volume, duration and inflow. Figure 4.9a, 4.9b and 4.9c present flood volume, duration and inflow frequency plots on normal probability axes under different base flows. When base flow is 2,000 CFS, flood volumes and durations appears have normal distributions. However, when base flows are 3,000 to 5,000 CFS, flood volumes and durations don't fit normal distributions. For all inflows (Figure 4.9c), all of them don't follow normal distribution, especially larger inflows deviate greatly from normal. This is mainly due to uncertainty of extreme flows. In this study, base flow 2,000 CFS is used for next step.

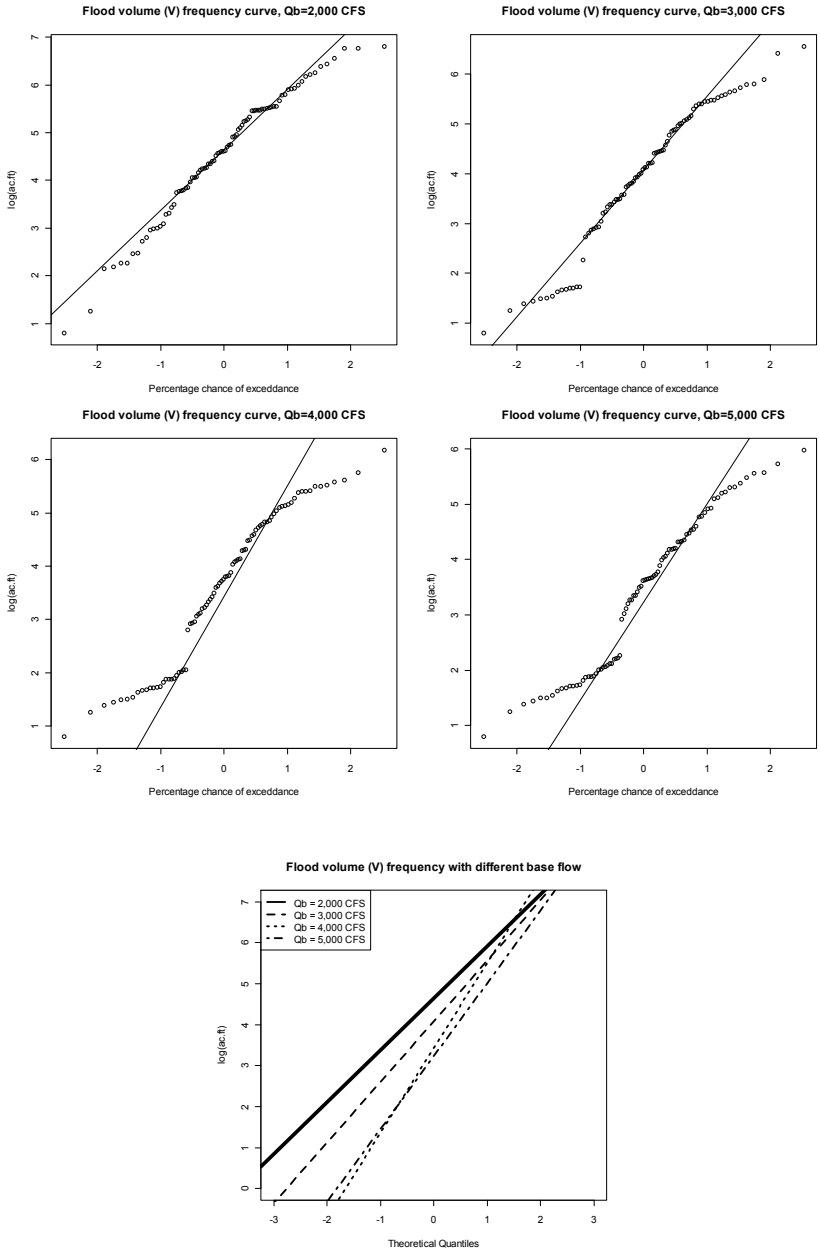


FIGURE 4.9A FLOOD PULSE VOLUME FREQUENCY UNDER DIFFERENT BASE FLOWS (A: Q_B=2,000 CFS; B: Q_B=3,000 CFS; C: Q_B=4,000 CFS AND D: Q_B=5,000 CFS)

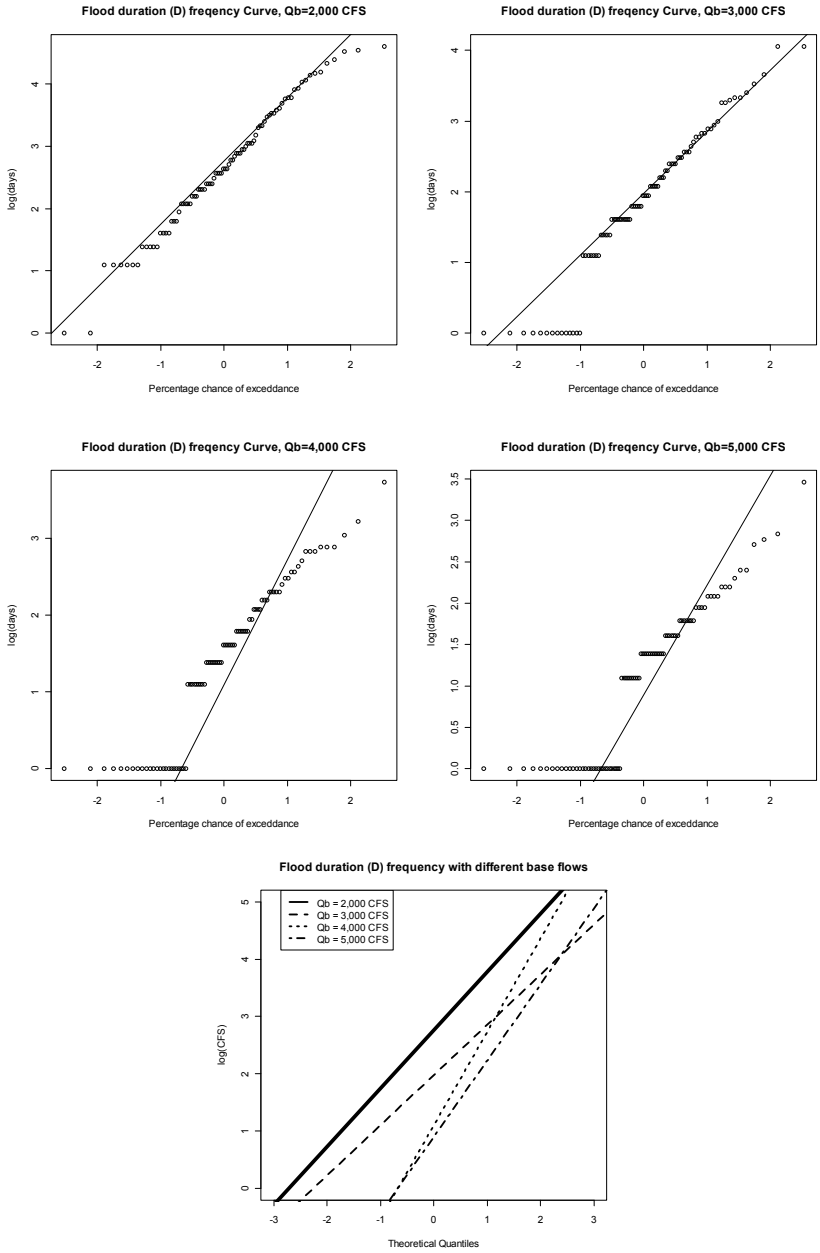


FIGURE 4.9B FLOOD PULSE DURATION FREQUENCY UNDER DIFFERENT BASE FLOWS (A: Q_B=2,000 CFS; B: Q_B=3,000 CFS; C: Q_B=4,000 CFS AND D: Q_B=5,000 CFS)

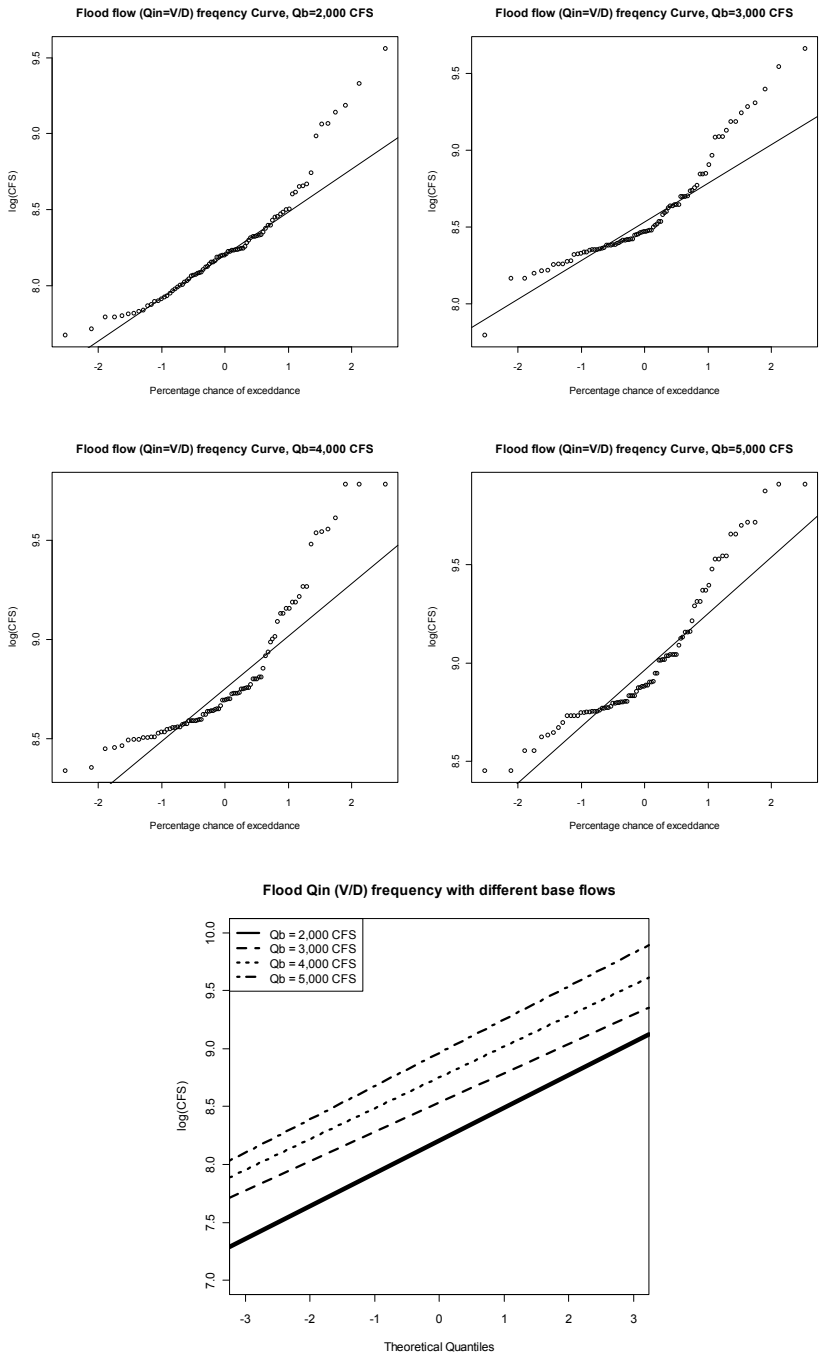


FIGURE 4.9C FLOOD PULSE FLOW (V/D) FREQUENCY UNDER DIFFERENT BASE FLOWS (A: Q_B=2,000 CFS; B: Q_B=3,000 CFS; C: Q_B=4,000 CFS AND D: Q_B=5,000 CFS)

- Fit Q_{in} (V/D) probability distribution

Normal, lognormal, loggamma and Weibull distributions are common for flood flows. Figure 4.11 presents four candidate distributions comparing with distribution of the observed unregulated flow. It appears the shapes of lognormal and loggamma distributions are closer to the observed distribution than normal and Weibull distribution. It is proved by K-S tests. Table 4.1 shows the statistics of four candidate distributions.

Lognormal is fittest distribution for unregulated flood flows. Therefore, a lognormal distribution will be used as flood inflow distribution in this site.

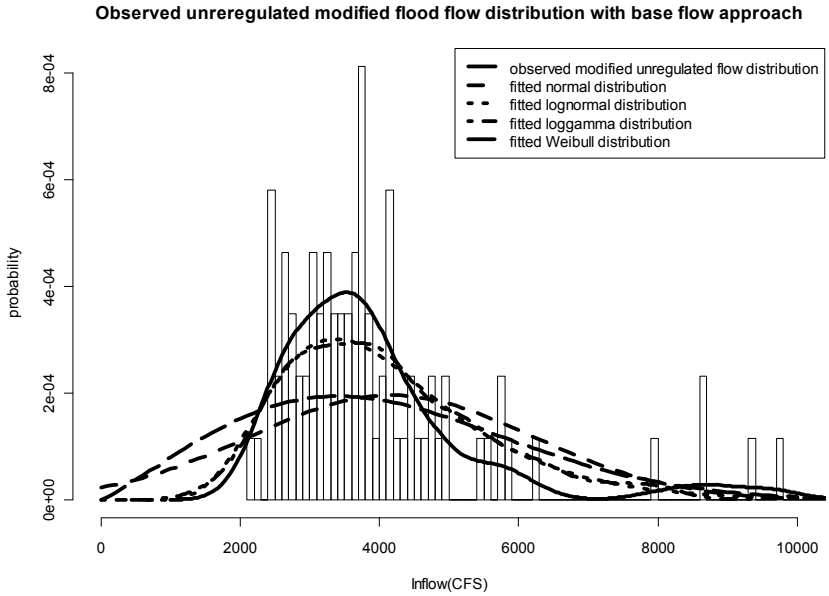


FIGURE 4.11 UNREGULATED FLOOD FLOW HISTOGRAMS WITH BASE FLOW 2,000 CFS

Table 4.1 Statistics of candidate distributions for modified unregulated flow with base flow approach

Distribution	unit	parameters		Goodness of fit values
				k-s test p value
Normal	CFS	mean=4,172	sd=2,019	0.0005104
lognormal	log(CFS)	mean=8.2573	sd=0.3663	**0.09268**
loggamma	log(CFS)	shape=525.95	rate=63.59	0.06361
Weibull	CFS	shape=2.1277	rate=4595.7995	0.00157

4.4.1 Reservoir Re-operation and Changes of Transformation Curves

This task is to construct unregulated flow and regulated flow relationship curves for different flood storage levels. In this application, besides flood control, Camanche and Pardee reservoirs serve water supply, hydropower and irrigation objectives. Operation rules at this site were developed and operated by USACE Sacramento District [USACE, 1981]. ResSim software was employed to simulated reservoir operations. Three storage options including 200 TAF, 180 TAF and 150 TAF were selected. AR method was used to generate unregulated flows. The unregulated modified flow is calculated by flood volume divided by duration, i.e., Q_{in}=V/D.

- Unregulated-regulated flow transform under deterministic scenarios

Figure 4.12 Illustration of unregulated-regulated flow transform curves at Camanche/Pardee Reservoirs for deterministic scenario. It appears that the regulated flow increase more rapid with smaller flood storage. The deterministic relations between unregulated (inflow) and regulated flows are in Table 4.2.

Table 4.2 The Deterministic Relationship between Inflow and Regulated Flows

Storage = 200 TAF		Storage = 180 TAF		Storage = 150 TAF	
Q_{in} (CFS)	Q_{reg} (CFS)	Q_{in} (CFS)	Q_{reg} (CFS)	Q_{in} (CFS)	Q_{reg} (CFS)
Q_{in} < 4,500 CFS	$Q_{reg}=Q_{in}$	Q_{in} < 3,300 CFS	$Q_{reg}=Q_{in}$	Q_{in} < 2,500 CFS	$Q_{reg}=Q_{in}$
4,500 < Q_{in} < 8,000	Q_{reg}=4,500	3,300 < Q_{in} < 6,000	Q_{reg}=3,300	2,500 < Q_{in} < 4,500	Q_{reg}=2,500
8,000 < Q_{in} < 10,000	Q_{reg}=2.75*Q_{in} - 17,500	6,000 < Q_{in} < 8,000	Q_{reg}=2.35*Q_{in} - 10,800	4,500 < Q_{in} < 6,500	Q_{reg}=2*Q_{in} - 6,500
Q_{in} > 10,000	$Q_{reg}=Q_{in}$	Q_{in} > 8,000	$Q_{reg}=Q_{in}$	Q_{in} > 6,500	$Q_{reg}=Q_{in}$

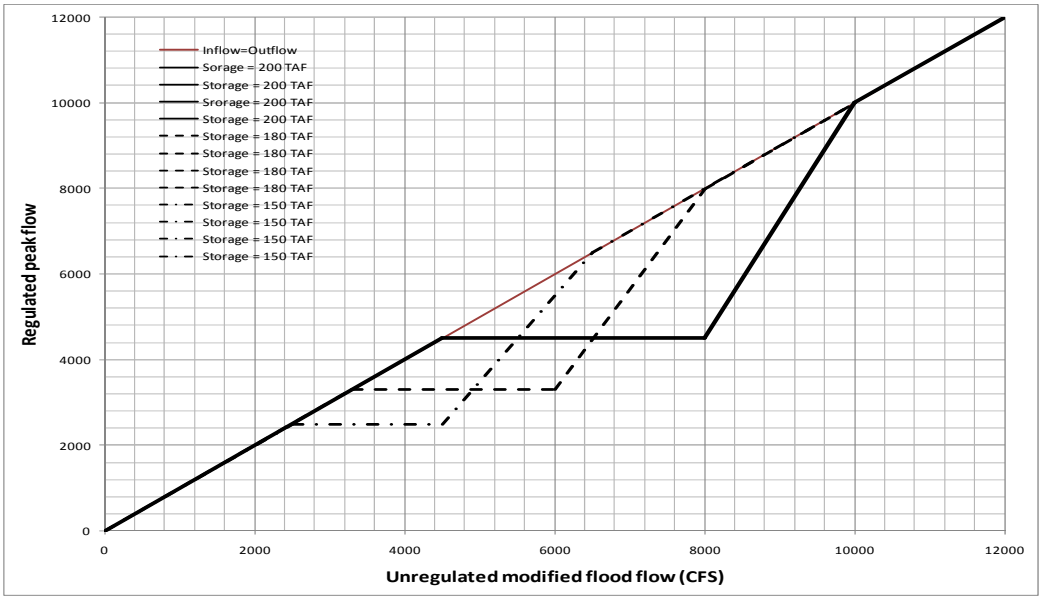


Figure 4.12 Illustration of unregulated-regulated flow transform curves at Camanche/Pardee Reservoirs for deterministic scenario

- Incorporating uncertainty into unregulated/regulated transform curves

In real world, the uncertainty is unavoidable. In this study, the error is incorporated into regulated flow. The equation is as followed:

$$\mathbf{Q_{reg}} = \overline{\mathbf{Q_{reg}}} + \boldsymbol{\varepsilon} \quad (4.12)$$

Where $\overline{\mathbf{Q_{reg}}}$ is deterministic part, i.e., mean of regulated flow. ε is an error part and ε follow normal distribution with mean 0, i.e., $Q_{reg} \sim N(\overline{Q_{reg}}, \varepsilon)$. The Table 4.2 presents the error part ε in three storage scenarios. It appears the error increase with the less flood storage size. Figure 4.13 illustrated the unregulated-regulated flow transform curves at Camanche/Pardee Reservoirs for these storage levels.

Table 4.3 Summary of Variance Coefficients (error part of ε)

Storage = 200 TAF		Storage = 180 TAF		Storage = 150 TAF	
Q_{in} (CFS)	error in Q_{reg} (CFS)	Q_{in} (CFS)	error in Q_{reg} (CFS)	Q_{in} (CFS)	error in Q_{reg} (CFS)
Q_{in} < 4,500 CFS	500	Q_{in} < 3,300 CFS	750	Q_{in} < 2,500 CFS	1,000
4,500 < Q_{in} < 8,000	800	3,300 < Q_{in} < 6,000	1,000	2,500 < Q_{in} < 4,500	1,500
8,000 < Q_{in} < 10,000	1,200	6,000 < Q_{in} < 8,000	1,500	4,500 < Q_{in} < 6,500	2,000
Q_{in} > 10,000	1,500	Q_{in} > 8,000	2,000	Q_{in} > 6,500	2,500

4.4.2 Regulated Flow Frequency Calculation

- Regulated flow distribution with/without incorporated error

The goal of this section is to find the impact of uncertainty to the inflow/regulated transform curves. The deterministic regulated flows, i.e., no consideration of uncertainty in transform curves are calculated from Table 4.2. Figure 4.13 present the comparison of with/without uncertainty in transform curves in three storage scenarios. It appears that the incorporated error have a great impact to transform curves under three storage scenarios. The variance of regulated flow increases when the error is incorporated and the value of variance increases with the storage size decreases.

- Determining regulated flow distribution

This section develops the regulated flow frequencies with three flood storage scenarios. Three probability distributions including Normal, Lognormal and LogGamma are selected as candidate distributions. Figure 4.14 presents the simulated regulated flood flow comparing to three fitted distributions. It appears none of them are very close to the shapes of simulated regulated density shape. However, Lognormal distribution is fitter than the other two distributions for size 180 and 150 TAF while normal distribution is fittest for the size 200 TAF. Table 4.4 summarized the statistical analysis of regulated flow for different flood storages. Lognormal distribution has higher K-S test p values when the flood sizes are 180 and 150 TAF while normal distribution has higher p value when the flood storage is 200 TAF. The mean and standard deviation of regulated flow increase as flood storage capacity decreases. Finally, Lognormal distribution will be used for regulated flow.

Table 4.4 Regulated flows statistics with base flow approach

Flood Storage Size	mean	sd	Goodness of fit(K-S test p value)		
	CFS	CFS	Normal	Lognormal	LogGamma
150 TAF	3,549	2,002	0.01203	**0.02317**	0.0005253
180 TAF	3,169	1,707	3.66E-07	**7.89E-06**	5.04E-07
200 TAF	2,795	1,058	**0.00368**	3.39E-13	2.11E-15

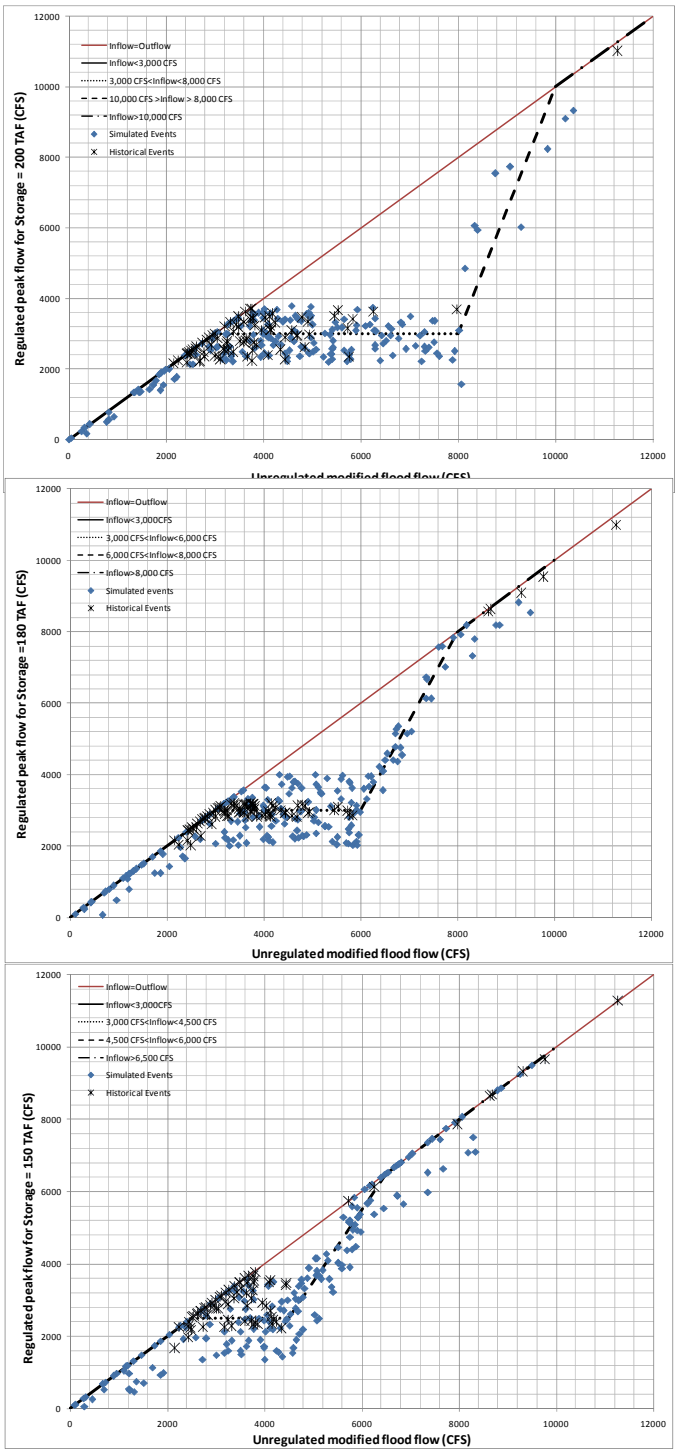


FIGURE 4.12 ILLUSTRATION OF UNREGULATED-REGULATED FLOW TRANSFORM CURVES AT CAMANCHE/PARDEE RESERVOIRS FOR VARIOUS STORAGES

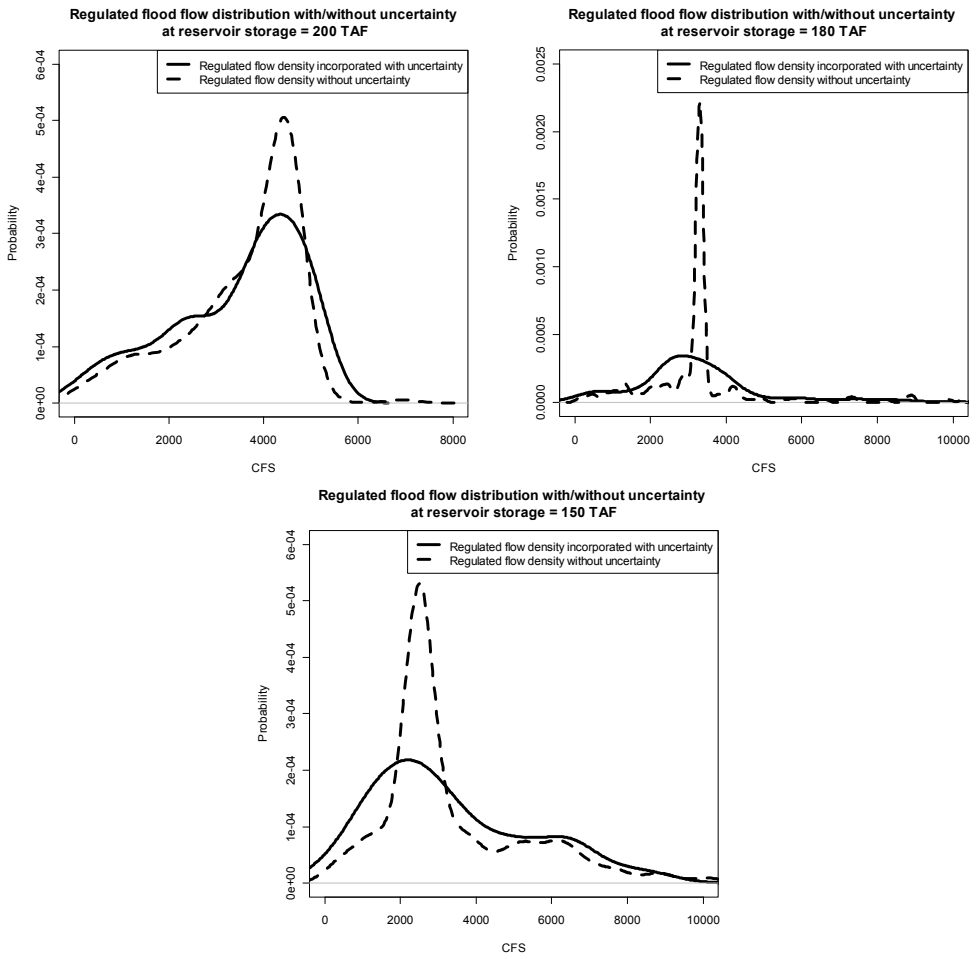


FIGURE 4.13 REGULATED FLOW DENSITY PLOTS WITH/WITHOUT UNCERTAINTY COMPARISON

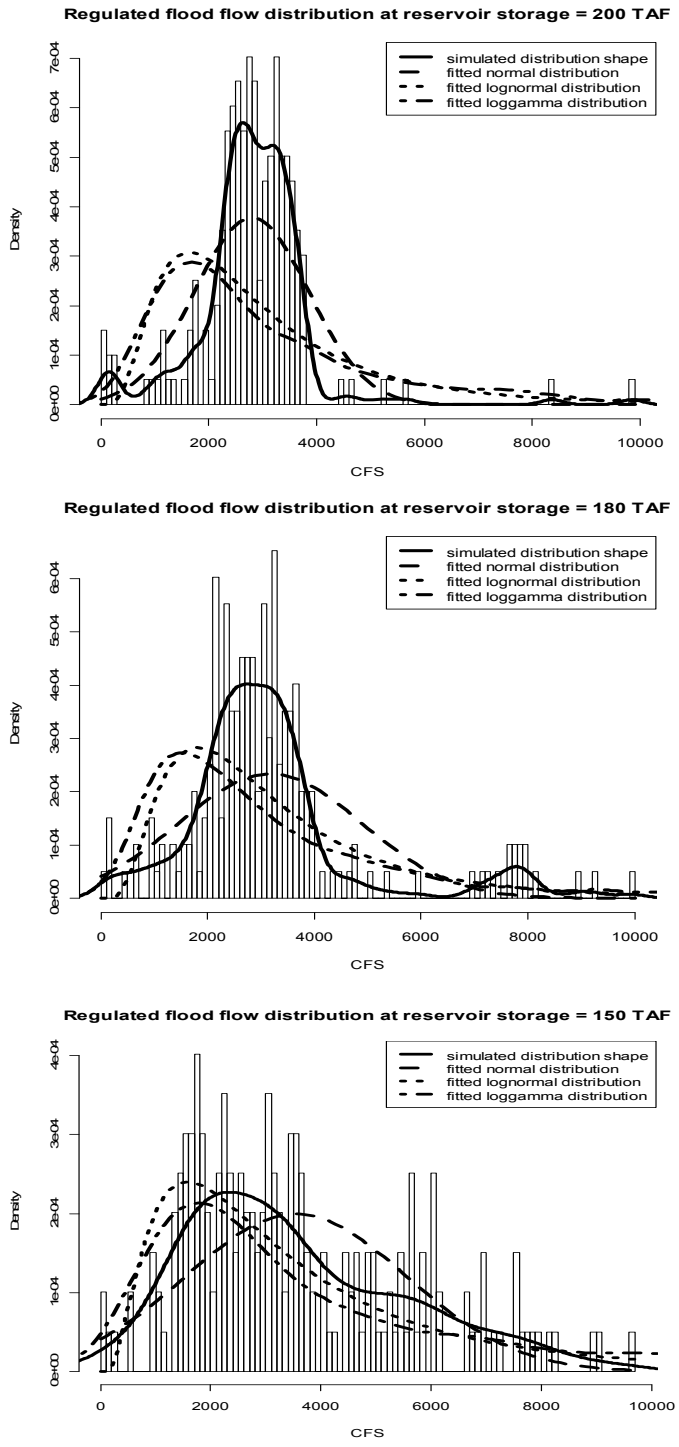


FIGURE 4.14 REGULATED FLOW HISTOGRAMS AT THREE STORAGE SIZE OPTIONS

4.5 COMPARISON WITH USACE APPROACH – CHARACTERISTIC TIME (CT) TO SEPARATE FLOODS

For the comparison purpose, this section starts with characteristic time approach to select flood pulses. All other steps are identical to the ones in Section 4.4.

According to USACE [2006], 3 days, 5 days, 10 days and 20 days are used to separate flood pulses. Figure 4.14a and 4.14b present flood volume (V) and inflow, V/D or Q_{in} frequency plots on normal probability axes under different CT. When CT is 10 days, flood volumes and inflows appear have normal distributions. Therefore, in this study, CT 10 days is used for next steps.

- Comparison in unregulated flow

Similar to Section 4.4, Normal, lognormal, loggamma and Weibull distributions are common for flood flows and selected as candidate distributions. Figure 4.15 presents four candidate distributions comparing with distribution of the observed unregulated flow. It appears the shapes of lognormal and loggamma distributions are closer to the observed distribution than normal and Weibull distribution. It is proved by K-S tests in Table 4.3. Table 4.3 shows the statistics of four candidate distributions. Therefore, a lognormal distribution will be used as flood inflow distribution in this site. The results are same as base flow approach, which indicate lognormal is fittest unregulated flow distribution. However, the goodness of fit value, K-S test p value in this approach is only 0.0005353, is much less than the one in base flow approach, 0.09268. It appears the base flow method to separate flood pulse is super than duration method in unregulated flow calculation.

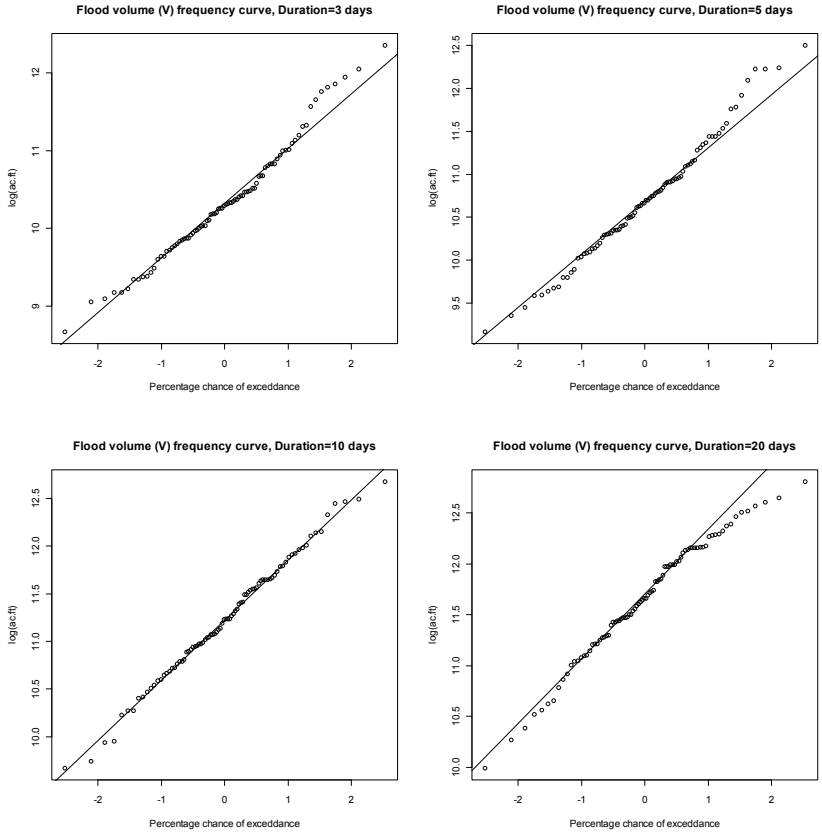


FIGURE 4.15A FLOOD PULSE VOLUME (V) FREQUENCY UNDER DIFFERENT DURATION (A: D = 3 DAYS; B: D = 5 DAYS; C: D = 10 DAYS AND D: D = 20 DAYS)

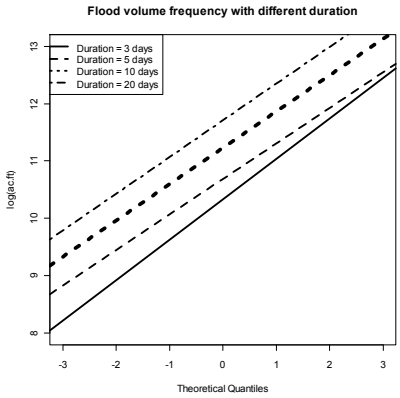


FIGURE 4.15A FLOOD PULSE VOLUME (V) FREQUENCY UNDER DIFFERENT DURATION (A: D = 3 DAYS; B: D = 5 DAYS; C: D = 10 DAYS AND D: D = 20 DAYS)

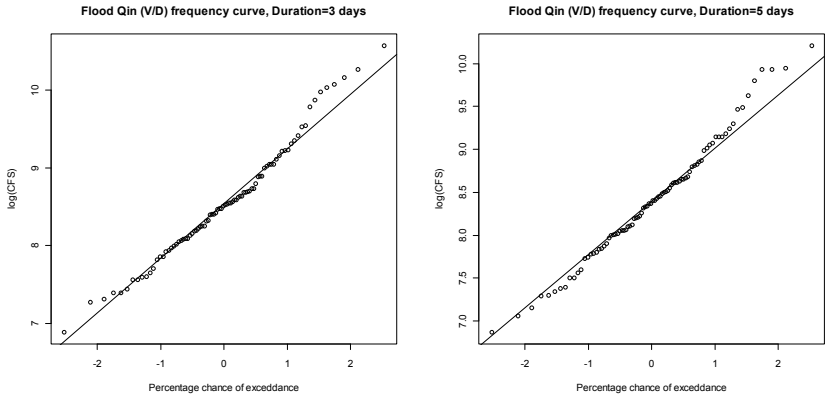


FIGURE 4.15B FLOOD PULSE INFLOW (V/D) FREQUENCY UNDER DIFFERENT DURATION (A: D = 3 DAYS; B: D = 5 DAYS; C: D = 10 DAYS AND D: D = 20 DAYS)

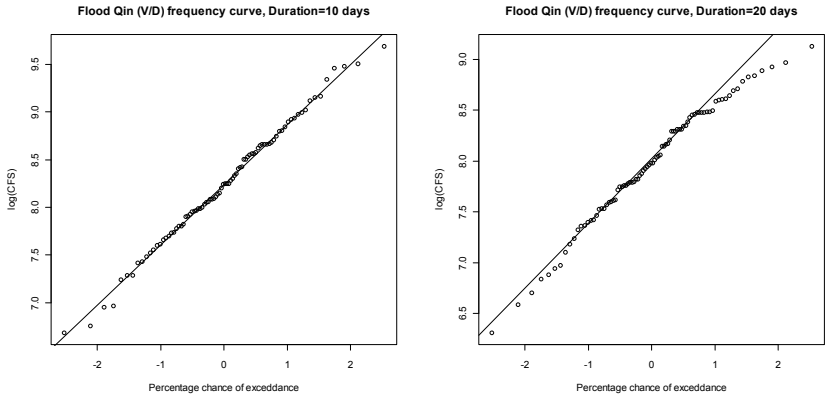


FIGURE 4.15B FLOOD PULSE INFLOW (V/D) FREQUENCY UNDER DIFFERENT DURATION (A: D = 3 DAYS; B: D = 5 DAYS; C: D = 10 DAYS AND D: D = 20 DAYS)

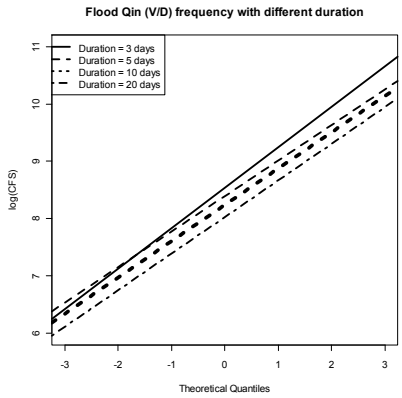


FIGURE 4.15B FLOOD PULSE INFLOW (V/D) FREQUENCY UNDER DIFFERENT DURATION (A: D = 3 DAYS; B: D = 5 DAYS; C: D = 10 DAYS AND D: D = 20 DAYS)

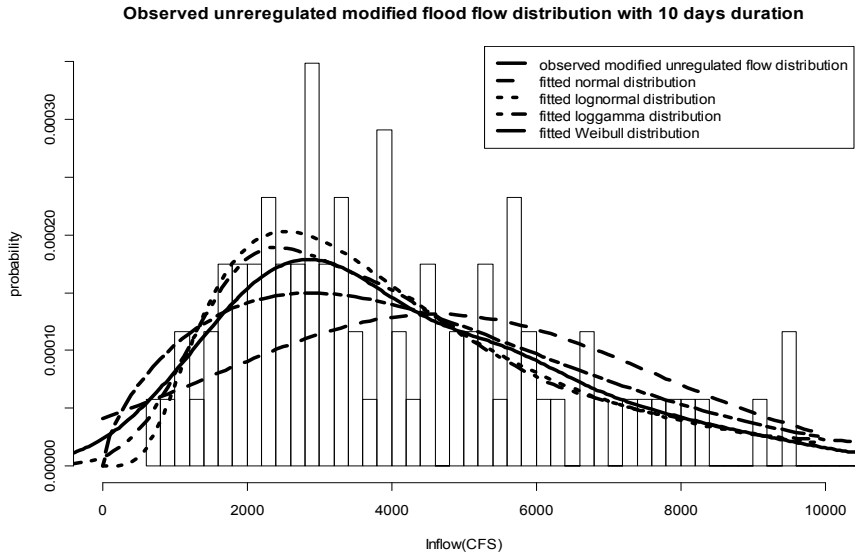


FIGURE 4.16 UNREGULATED FLOOD FLOW HISTOGRAMS WITH CT 10 DAYS

Table 4.5 Statistics of candidate distributions for unregulated flow with USACE approach

Distribution	unit	parameters		Goodness of fit values
				k-s test p value
Normal	CFS	mean=4,610	sd=3,037	2.37E-07
lognormal	log(CFS)	mean=8.2398	sd=0.6346	**0.0005353**
loggamma	log(CFS)	shape=166.5787	rate=20.2164	0.0003543
Weibull	CFS	shape=1.6392	rate=5136.2942	0.0001063

- Comparison in regulated flow

Same as Section 4.4, this section develops the regulated flow frequencies with three flood storage scenarios. Three probability distributions including Normal, Lognormal and LogGamma are selected as candidate distributions. Figure 4.17 presents the simulated regulated flood flow comparing to three fitted

distributions. It appears none of them are very close to the shapes of simulated regulated density shape. However, Normal distribution is fitter than the other two distributions for size 180 and 200 TAF while lognormal distribution is fittest for the size 150 TAF. Table 4.7 summarized the statistical analysis of regulated flow for different flood storages. Normal distribution has higher K-S test p values when the flood sizes are 180 and 200 TAF while Lognormal distribution has higher p value when the flood storage is 150 TAF. Same as base flow approach, the mean and standard deviation of regulated flow increase as flood storage capacity decreases.

Table 4.6 Regulated flows statistics with USACE approach

Flood Storage Size	mean	sd	Goodness of fit(K-S test p value)		
	CFS	CFS	Normal	Lognormal	LogGamma
150 TAF	3,346	2,107	0.005604	**0.254**	0.08696
180 TAF	3,215	1,753	**0.00268**	1.24E-08	1.57E-06
200 TAF	3,420	1,383	**0.0005355**	2.08E-07	4.18E-08

Figure 4.18 present the comparisons of two approaches to select flood pulses including USACE approach by duration and this study's base flow approach. When the flood storage is 200 TAF, the regulated flow frequency has more occurrences in low flows and less occurrences in extreme high flows than USACE approach. When the storage is 180 and 150 TAF, there is no big difference of regulated flow distributions between these two approaches.

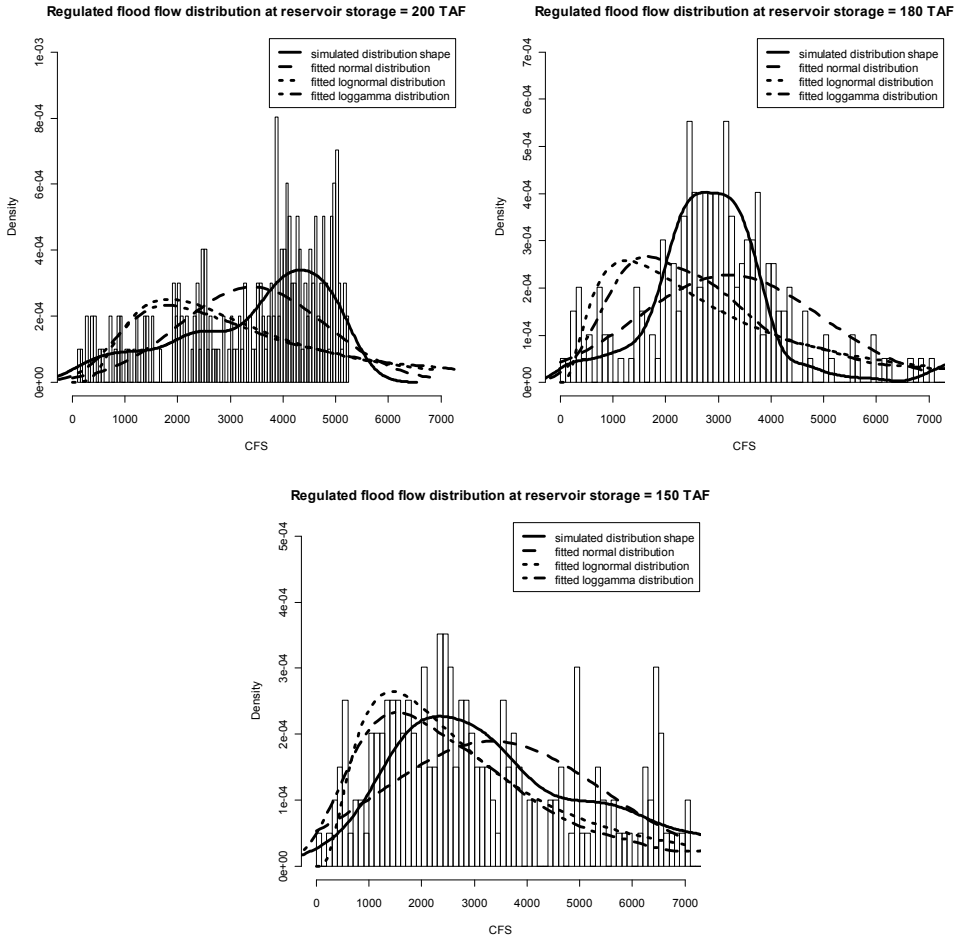


FIGURE 4.17 REGULATED FLOW HISTOGRAMS AT THREE STORAGE SIZE OPTIONS WITH USACE APPROACH

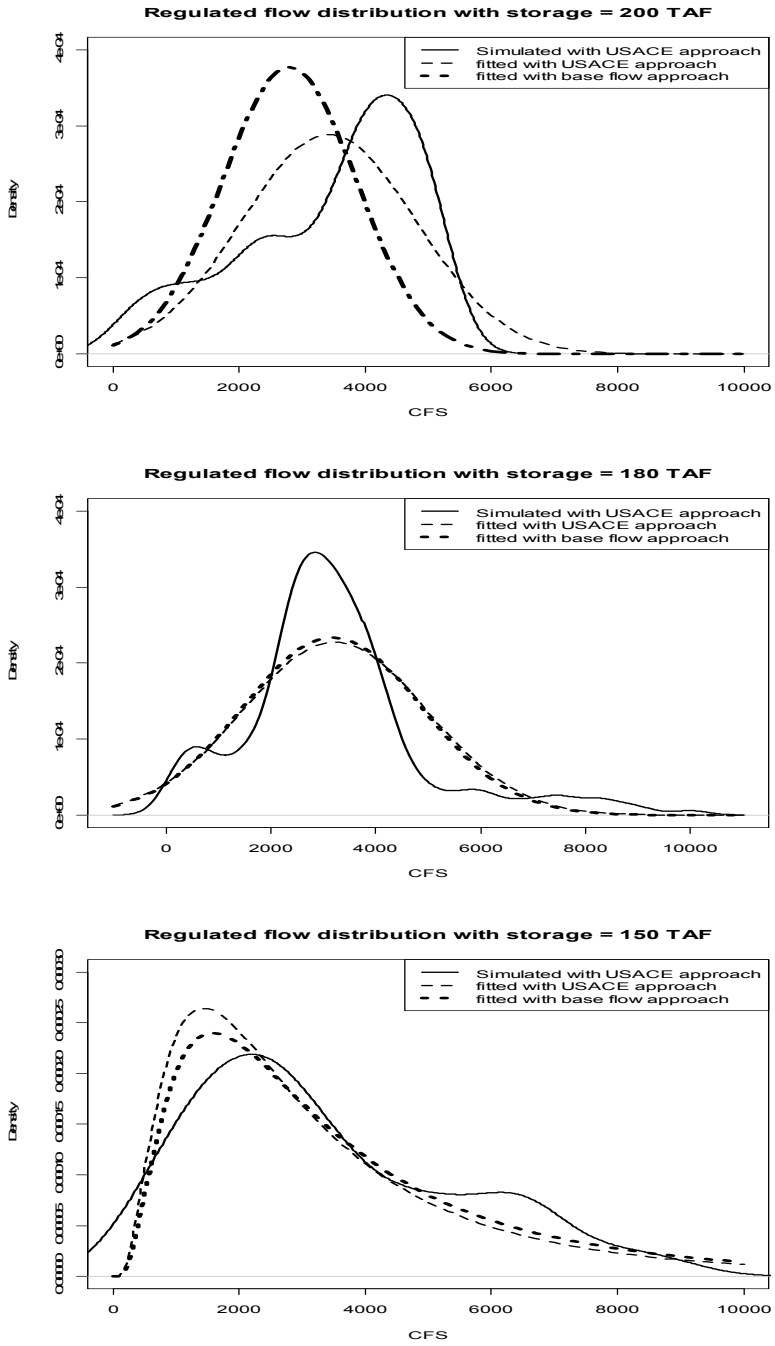


FIGURE 4.18 REGULATED FLOW HISTOGRAMS AT THREE STORAGE SIZE OPTIONS WITH TWO APPROACHES

4.6 CONCLUSION AND FUTURE RESEARCH SUGGESTIONS

This chapter proposed a conceptual framework to estimate regulated flow frequency with changes in flood storage space. Three main steps including unregulated flow frequency analysis, unregulated/regulated flow transformation and regulated flow frequency determination are presented. The main contributions include separating flood pulses from daily flow time series by base flow, modification of unregulated flow calculation to accommodating to reservoir's flood control, and fitting unregulated and regulated flow to appropriate probability distributions. The results will be inputs for levee failure analysis in next chapter.

Some limitations of this study arise from frequency methods and reservoir operation modeling.

1) Simplicity of Inflow/outflow relationship curves: Scores of other factors exist in regulated versus unregulated flow relationship such as initial water surface elevation at the beginning of flood, outlet capacity, operation uncertainty. However, this study only set flood storage as factor. This would greatly simplify the whole process and bring the bigger uncertainty for regulated flood frequency.

2) Extrapolation problem: This is a big issue for flood frequency analysis. To estimate low frequency events, one must rely on extrapolation from short of records. Viessman et al (1977) note that "as a general rule, frequency analysis should be avoided…in estimating frequencies of expected hydrologic events greater than twice the record length." In this study, only 86 years records are available. Theoretically, generated regulated flow distribution will be used to determine the extreme event such as 100 year and 200 year.

3) Assumption of stationary: Historical records only provide the past records. With the climate change, the hydrology would change. Klemes noted many known causes for non-stationary ranging from the dynamics of the earth's motion to human cause changes in land use [Klemes, 1986]. This study assumes stationary.

Future study suggestions:

1) In inflow frequency, PDS should be considered to analyze flood pulse statistics since PDS has advantage over AMS.

2) In regulated/unregulated flow relations, more factors including initial storage situation, downstream capacity changes and operation uncertainty will be considered and quantified.

REFERENCES

Ang, A.H.-S., and Tang, W.H.(1984), "Probability Concepts in Engineering Planning and Design, Vol.II: Decision, Risk, and Reliability", John Wiley and Sons, New York.

Bedient, P., Huber, W. (1992), "Hydrology and Floodplain Analysis", Addison-Wesley: New York.

Cunnane, C., (1978), "Unbiased Plotting Positions – A Review", J. of Hydrology, 37(1978) 205-222.

Cunnane, C., (1974), "A Particular Comparison of Annual Maxima and Partial Duration Series Methods of Flood Frequency Prediction", J. of Hydrology, 18(1973) 257-271.

Ford, D.(1990), "Reservoir Storage Reallocation Analysis with PC", J. Water Resources Planning and Management.Vol.116(3) .

Dagpunar, J. (1988), "Principles of Random Variate Generation", New York: Oxford University Press.

DeGroot, M., Schervish, M. (2002)."Probability and Statistics". Addison-Wesley: New York.

Goldman, David M. (2001). "Quantifying Uncertainty in Estimates of Regulated Flood Frequency Curves." State of the Practice – Proceedings of the World Water and Environmental Resources Congress, ASCE, Reston, Va.

Ergish, N. (2010) “Flood frequency analysis for regulated watershed.” Master thesis in Department of Civil and Environmental Engineering, University of California, Davis.

Hickey, J.T., Collins, R.F., High, J.M., Richardson, K.A., White, L.L., and Pugner, P.E.(2000) “Synthetic rain flood hydrology for the Sacramento and San Joaquin River Basins.” J.Hydrologic Eng., 7(3), 195-208.

Hickey, J.T., Collins, R.F., High, J.M., Richardson, K.A., White, L.L., and Pugner, P.E.(2003) “Reservoir simulations of synthetic rain floods for the Sacramento and San Joaquin River Basins.” J.of Water Resources Planning and Management, 2003, 443-457.

IACWD (Interagency Advisory Committee on Water Data), 1982. Guidelines for determining flood flow frequency, Bulletin #17B, US Department of the Interior, Geological Survey.

Linsley, R., Kohler, M., Paulhus, J. (1982), “Hydrology for Engineers”, McGraw-Hill Companies, New York.

Langbein, W.B.(1949), "Annual Floods and the Partial Duration Flood Series", EOS,, Trans. AGU, vol.30, no.6, pp.879-881, 1949.

Loucks,D.P., Stedinger, J., Haith, D.(1981), “Water Resource Systems Planning and Analysis”, Prentice-Hall, Englewood Cliffs, New Hersey.

Madsen, H., Rosbjerg,D, and Harremos, P. (1994) “PDS-modeling and regional Bayesian estimation of extreme rainfalls” Nordic Hydrol., Vol 25(4), 279-300.

Madsen, H., Rasmussen, P., Rosbjerg, D. (1997), “Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events”, Water Resources Research, Vol. 33, No.4, Page 747-757.

McCuen,M. (1989), “Hydrologic Analysis and Design”, Prentice-Hall, Englewood Cliffs, New Hersey.

Pramanik, N., Panda, R., and Sen, D. (2010) “Development of design flood hydrographs using probability density functions”, Hydrological Process: 24, 415-428.

R Core Team. (2010) “An Introduction to R”. Website: http://cran.r-project.org/doc/manuals/R-intro.html

Salas, J. (1993) "Analysis and modeling of hydrologic time series, in D.R. Maidment(Ed) Handbook of hydrology. MaGraw-Hill, Inc, New York, Chapter 19..

Shane, R., and W. Lynn, “Mathematical model for flood risk evaluation”, J. Hydraul Div., Amer. Soc. Civil Eng., 90(HY6), 1-20, 1964.

Singh, V. P. and Fiorentino, M. (1992), “A historical perspective of entropy applications in water resources”, in Singh, V. P. and Fiorentino, M. (Eds), Entropy and Energy Dissipation in Water Resources. Kluwer Academic Publishers, Dordrecht. pp. 155 to 173.

Stedinger JR., Vogel RM., Foufoula-Georgiou E. (1993) “Frequency analysis of extreme events” In Handbook of Hydrology, Maidment DR. (ed) from 18-1 to 19-65. McGraw Hill: New York.

Todorovic, P., and Zelenhasic, A (1970) “Stochastic model for flood analysis”, Water Resource Research, 6 (6), 1641-1648.

Tung, Y.K., Yen, B.C., Melching, C.S. (2006) “Hydrosystems Engineering Reliability Assessment and Risk Analysis”. McGraw-Hill.

USACE (1981) "USACE Water control manual for Camanche dam and reservoir, September 1981".

USACE (1996). “Risk-Based Analysis for Flood Damage Reduction Studies”, USACE Manual No. 1110-2-1619.

USACE (1993) “Hydrologic Frequency Analysis”, EM 1110-2-1415, USACE, Washington, D.C.

USACE (1997). “Hydrologic Engineering Requirements for Reservoirs”, EM 1110-2-1420, USACE, Washington, D.C.

Valdes, J. and Marco J. (1995), “Managing Reservoir for Flood Control”, U.S.- Italy Research Workshop on the Hydrometeorology, Impacts, and Management of Extreme Floods, Perugia, Italy.

Wurbs. R. (1996) “Modeling and Analysis of Reservoir System Operations”, Prentice-Hall, Inc., New Jersey.

Wurbs, R., (2002) “State-frequency Analysis for Urban Flood Control Reservoir”. J. Water Resources Planning and Management.

Wurbs, R., (1993) “Reservoir-system Simulation and Optimization Models”. J. Water Resources Planning and Management.Vol.119, No.4, July/August, 1993.

Wurbs, R., Cabezas, L., (1987) “Analysis of Reservoir Storage Reallocation”. Journal of Hydrology. No.92(1987) 77-95

Xu, Z.X., Ito, K., Liao, S and Wang, L (1997), “Incorporating Inflow Uncertainty Into Risk Assessment For Reservoir Operation”, Stochastic Hydrology and Hydraulics 11:433-448.

Yeh, W. (1985). “Reservoir management and operation models: A state of the art review”. Water Resources Research, Vol. 21, NO. 12, pp. 1797-1818.

Yue, S., Ouarda, T., Bobee, B., Legendre, P, and Bruneau, P.(2002). “Approach for describing statistical properties of flood hydrograph”. Journal of Hydrological Engineering, Vol. 7, NO. 2, pp. 147-153.

Explanation of Variables and simplified terms

AFS – annual flood series

AMS – Annual Maximum Series

AR – Acceptance-rejection

ARMA – auto regressive and moving average

CFS – cubic feet per second

CT – characteristic time

DWR – Department of Water Resources

FFA – Flood Frequency Analysis

GP – generalized Pareto

HEC – Hydrologic Engineer Center)

IID – identically and independently distributed

KS test – Kolmogorov-Smirnov test

MLE – maximum likelihood estimate

MOM – method of moments

PDS – Partial Duration Series

SD – Standard Deviation

TAF – thousand ac.ft

USGS – United States Geological Service

USACE – U.S. Army of Corps of Engineer

VDF – Volume Duration Frequency

CHAPTER 5 FLOOD LEVEE FAILURE ANALYSIS WITH HYDROLOGIC, HYDRAULIC AND GEOTECHNICAL UNCERTAINTIES

China, also called Nine States, is based Da Yu's delineation of watersheds.

- Si Ma Qian, Han Dynasty Historian,100 BC

SUMMARY

Levee failure has drawn more attention recently due to urbanization behind levees and climate change increasing hydrological extremes. Levees can fail by several mechanisms. This chapter introduces a framework to assess levee overall failure probability from overtopping and erosion, incorporating several hydrologic, hydraulic and geotechnical uncertainties. Two main contributions include overall failure probability estimation and load-resistance analysis. Overtopping and erosion failure analysis are usually performed separately in water resources and geotechnical engineering. This chapter presents a more integrated analysis combining these two failure mechanisms. Also load-resistance analysis is introduced to consider overtopping between flood magnitude and levee capacity and erosion failure between velocity and soil strength. Both analyses are performed by Monte Carlo simulation to estimate overall levee failure probability. Failure probability can be more sensitive to geotechnical variables and less sensitive to reservoirs reoperation. The Lower Mokelumne River levee system below Camanche/Pardee reservoirs in Northern California is the application site.

5.1 INTRODUCTION

Levee systems have been built for flood protection worldwide over human history. In the Netherlands, half of the country is protected by primary dikes and other water-retaining structures since the middle ages [Vrijling, 2001]. In China, with high dense populations living along two major rivers, Yangze River and Yellow River, hundreds of millions of people's lives depend on levee systems along these rivers [China Information Almanac, 2009]. Levee failures have occurred quite often recently.

In California's Central Valley, approximately 1,600 miles of State/federal levees and 520 miles private levees protect some of the country's most productive farmland and rapidly growing areas in the country's most populous state [DWR, 2009]. The levee system, key part of Central Valley Project, protects people, property and infrastructure from flooding on the Sacramento and San Joaquin rivers. Over 1 million people and 1.9 million acres of cultivated land with an annual production value of over $2 billion are protected behind project levees. Approximately 200,000 structures with an estimated value of $64 billion rely on this system, as well as the last remnants of extensive riparian forests and wetlands that once flourished along California's rivers prior to European settlement. Finally, the water supply for 23 million people, farms and industry relies on a vast network of 1,100 miles of De
earthquake damage and flooding. Economic losses from a catastrophic f
the tens of billions of dollars [DWR, 2009]. However, because of c
hydrology, urbanization, inadequate design and construction, deferred maintenance, erosion and other reasons, levees in California have lost capability. Over the years, major storms and flooding by levee failure have taken lives, caused significant property losses and caused extensive damage to public infrastructure.

Levee failure mechanisms are complicated and involve several fields of expertise including hydrological, hydraulic, geotechnical and structural engineering. Overall, levee failure mechanisms include bearing, sliding, slump/spread, seepage, erosion and overtopping. Generally, failure of infrastructure can be classified broadly into two types [Yen and Ang, 1971; Yen et al., 1986]: structural failure and functional failure (performance). Figure 5.1 presents the common levee failure mechanisms. Structural failure involves damage or change of the structure or facility, therefore hindering its ability to function as desired. Performance failure doesn't necessarily involve structure damage. However, the performance limit of the structure is exceeded, and undesirable consequences occur. California's levee failures are mainly due to overtopping and levee breach [DWR, 2008]. In this study, levee failure from overtopping is categorized as a function failure while failure due to breach is a structural failure. When overtopping happens, the levee's structure may be not destroyed at all. Flood damage will be considered by the flooded area. After the

flood recedes, the levees can still function. However, if the levee fails due to breach, the levee loses both its function and its structural stability.

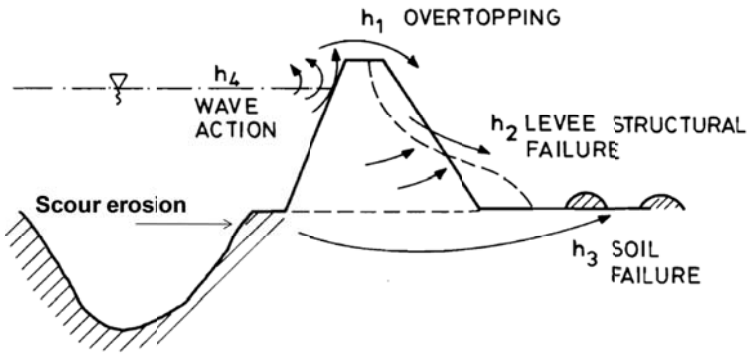


FIGURE 5.1 COMMON LEVEE FAILURE MECHANISMS [WOOD, 1977]

Risk and reliability analysis can be important in decision-making. It can give decision makers a better understanding of the levee and flood control system. However, gaps in knowledge remain. The quantification of levee failure mechanisms is facilitated by several methods ranging from empirical equations to more physical process-based models. Currently, there is growing interest to quantify the probability of levee failure using probabilistic approaches [Vrijling et al., 2001, Voortman 2002]. The probabilistic method allows designers to account for uncertainties of input parameters, treating them as random variables and to better estimate probability of failure. Overtopping involves hydrologic and hydraulic uncertainties. Levee breach processes involve hydraulic and geotechnical uncertainties.

Two main approaches exist to examine levee failure. One approach is to estimate overtopping failure probability mainly from hydrologic uncertainty. Numerous papers in water resources engineering identify and quantify most water-related variables [Tung & Mays, 1980]. The other approach for levee failure is from geotechnical engineering. Their goal is to find structure failure probability including erosion, seepage, slope instability and others failure mechanisms. Gaps remain in combining these two approaches. Numerous articles in the literature concern hydrologic models for describing flood magnitudes help to describe the probability of overtopping. Bras [1978] presented an excellent review of many of these models which only account for hydrologic uncertainties. In this approach, levee failure is only due to overtopping. To define levee structural failure, Duncan and Houston estimate failure probabilities for California levees constructed from a heterogeneous mixture of sand, silt, and peat, and founded on peat of uncertain strength. However, in this approach, overtopping failure is ignored [Duncan and Houston, 1983]. Traditionally, overtopping failure determination focuses on hydrologic uncertainty while other variables such as soil properties are defined as deterministic. To find structural levee failure such as erosion or slope stability analysis, researchers treat geotechnical variables as random variables while hydrologic variables of flow or velocity are deterministic.

Some investigators have attempted to define stochastic models for levee failure that combine several mechanisms. Bogardi, et al. [1975] attempted to define a probability of failure consider overtopping, seepage, slope stability failure and wind wave causing scouring. This is an interesting approach in that the probabilities of overtopping and subsoil failure are related to flood peaks, slope stability failure is related to both flood peaks and flood duration, and wind wave action is related to flood peaks and water levels. However, no geotechnical variables are included such as soil properties like soil strength and erosion. Wood [1977] attempts to develop flood levee reliability models to consider both structural failure and overtopping. The probability of structural failure was described using uniform and quadratic distributions of an exccedance discharge at which the levee structurally fails. This simplistic approach does not consider work from the geotechnical literature. As pointed out in the geotechnical literature, the probability of slope stability failure must be related to the soil properties and also on a long-term basis [Yuceman et al., 1976] because of the progressive failure effects and changing pore pressure due to the wetting and drying. Even though the probability of slope stability failure has been defined in terms of changing soil properties, this has never been combined with risk and reliability models for overtopping.

Bras [1978] pointed out that even though considerable progress has been made in quantifying hydrologic uncertainties, there several problems with the present day handling of reliability in water resources. (1) present practice is limited and ignores tremendous hydrologic, hydraulic and geotechnical uncertainties; (2) hydraulic and geotechnical uncertainty analysis remains largely unexplored, and (3) a unified approach to quantifying reliability is lacking in that water resource design and analysis should be objective dependent and the modeling of loads (hydrology) and response (hydraulics and geotechnical variables) should be compatible with the ultimate goals of the project.

This study aims to narrow this gap between water resources and geotechnical engineering and construct overall levee failure probabilities including overtopping and structural failures.

Since hydraulic uncertainties are involved in overtopping, Tung [1981] constructed models to analyze both hydrologic and hydraulic uncertainties to determine levee overtopping risk. First-order analysis was employed to quantify levee capacity's uncertainty. The probability distribution of the levee capacity, Q_r, is assumed to be lognormal. Finally, risk-safety relationships were built with different return periods.

Most levee systems do not fail by overtopping but by structural weakness, either in the levee or in the soil near it [Wood, 1977]. The geotechnical literature defines structural levee failures, with most uncertainties assigned to geotechnical variables such as soil properties, while loads like hydrologic and hydraulic variables are treated as deterministic [Wolfe, 2008]. This approach underestimates the structural failure probability. In this study, both geotechnical variables and hydrologic and hydraulic variables including flow and duration are defined as uncertainties to define levee failure due to erosion.

In summary, the objective of this chapter is therefore to investigate levee overall failure probability from overtopping and erosion and the effects on probability distributions of hydrologic, hydraulic and geotechnical variables. To determine overtopping failure, both load and resistance variables, i.e. hydrologic and hydraulic, uncertainties are considered. To define failure due to erosion, hydrologic, hydraulic and geotechnical uncertainties are identified and quantified. This is performed by Monte Carlo simulation. Finally, overall levee failure probability is computed assuming of independent failure mechanisms.

The organization of this chapter is as follows. Section 2 describes the literature review and concepts development including levee failure mechanisms, reliability of levees, uncertainties in overtopping and erosion. Section 3 presents the methods flow chart used in this study. Section 4 gives the application to the Lower Mokelumne River levee system in Northern California. Section 5 shows the conclusions and limitations.

5.2 LITERATURE AND CONCEPTS DEVELOPMENT

This section presents literature review and some concepts development in levee failure probability analysis.

5.2.1. Levee failure analysis

Failure mechanisms in overall

Levee failures are mainly by overtopping and breach. For each case, the results can be significant from acres of land flooded to loss of lives. Generally, loading situations that can cause failures falls into four categories including: high water from large flows, waves, ground shaking from seismic activity, and/or high tides and static stress conditions. From these four loading situations, six related failure mechanisms have been observed, including overtopping, erosion, bearing, sliding, slumping/spreading, and seepage. Figure 5.1 shows a schematic of loading functions, failure mechanisms and how they interrelate [Seed, et al., 2006]

Overtopping occurs when high water exceeds the levee crest elevation. Also the flow's energy loss is concentrated at the levee's inboard toe leading to soil erosion and decreased levee stability. Overtopping failure can be prevented by sandbags or raising the crest elevation to accommodate the high flow. Seepage is another common failures mechanism. Levees are built in a fluvial depositional environment

and it is common for levees to have a sandy layer under the foundation. The sandy layer can be a conduit for flow under the levee, resulting in critical conditions at the landward toe. This can erode the foundation during high flows or weaken the foundation over a long period. Plants and animals can also destabilize the levee and cause seepage, such as through rodents holes, tree roots, or other biological activities that create conduits for seepage [Wolfe, 2008].

Water-side erosion is caused by high energy gradient or velocity and the materials from the water side are scoured, leading to instability and failure. Erosion can occur at once or over time as a function of water velocity, duration and material properties [Wolfe, 2008].

Other failure mechanisms including bearing, sliding and slump/spread, are mainly caused by ground shaking such as an earthquake. A bearing failure in levees is typically deep-seated and is most likely induced by seismic ground shaking. Failure is commonly triggered by a seismic event that either causes a loss of soil strength or induces destabilizing inertial loading conditions. A sliding failure may occur if the foundation soil has a weak or brittle zone resulting in a preferred failure plane. Both seismic induced inertial loading and high water levels can cause sliding failures. Slumping and spreading can be generated by two loading conditions. Cyclic loading from earthquakes may increase pore pressures and reduce soil strength, leading to volumetric and/or deviatoric strains in the foundation. The same results also can occur due to increased pore pressure from high water levels and seepage [Wolfe, 2008].

Of six levee failure mechanisms, four involve high water and waves. Bogardi and Zoltan [1968] have identified four common modes of failure due to water: 1) The first is overtopping, in which the flood elevation exceeds the levee crest. 2) The second is structure failure by water saturation and loss of soil stability. Flood increases saturation of the levee and increase pressure gradient through the levee. Decrease in soil strength is from increased saturation, which, with the increase in pressure gradient from the height of flood, leads to levee failure through slumping. 3) The third mode of failure consists of boils and hydraulic soil failures. The flood height and its resulting pressure are transmitted through the foundation soil under the levee and can cause soil failure through rupturing. The resulting failure usually leads to large flows of water into the protected areas and the undermining of the levee's foundation. 4) The fourth mode is wave action. High flood levels give rise to waves which scours the top of the levee. Such scouring reduces levee strength and can cause failure.

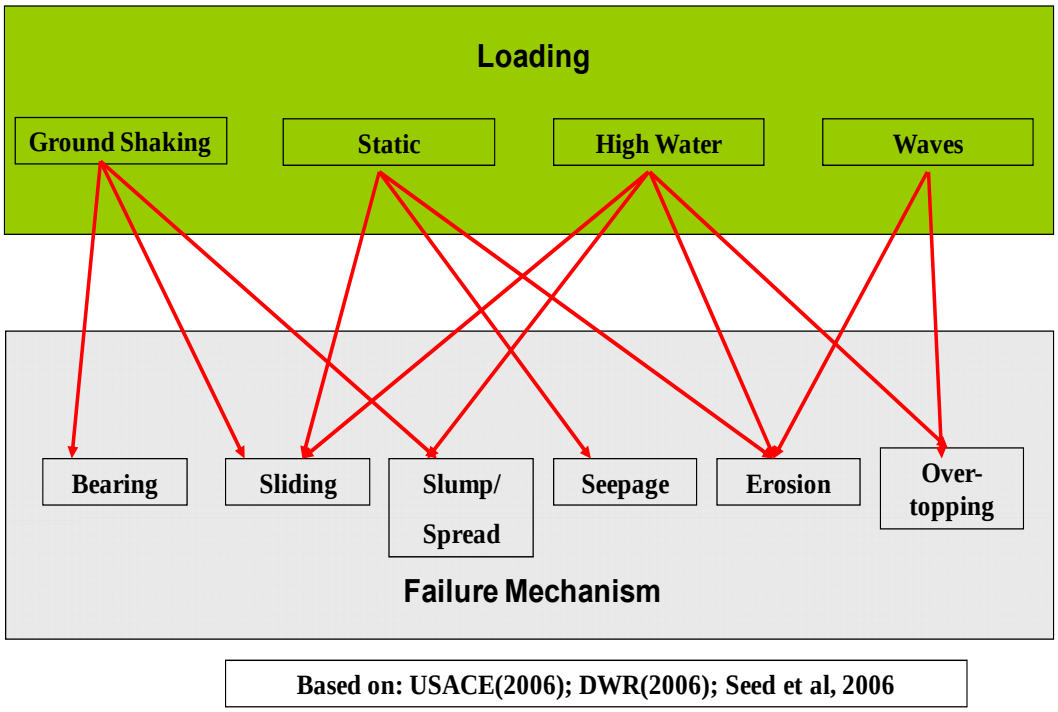


FIGURE 5.2 LEVEE FAILURE MECHANISMS

Overall, levee failure is complicated and the failure mechanism involves from structure engineering, geotechnical engineering, hydrological/hydraulic engineering as well as the levee's maintenance and operation, wildlife habitat and so on. No model can yet include all these mechanisms to analyze the failure probability. In this study, high water is defined as a loading factor and the erosion and overtopping failure mechanisms are considered.

Beyond overtopping

More levee failures seem to be related to levee breaches than overtopping [DWR, 2006]. Numerous articles address overtopping probabilities. Few papers address breach probabilities. Most levee failures involving breach are geotechnical, focusing on slope instability, seepage and erosion.

Duncan and Houston estimated failure probabilities for California levees constructed from a heterogeneous mixture of sand, silt, and peat, and founded on peat of uncertain strength. The factor of safety was expressed as a function of shear strength, which is a random variable due to its uncertainty, and water level, which has a defined annual exceedance probability. Values for the annual probability of failure for 18 islands in the levee system were calculated by integrating numerically over the joint events of high-water levels and insufficient shear strength [Duncan and Houston, 1983].

Vrouwenvelder (1987) provides a thorough treatise on a proposed probabilistic approach to design of dikes and levees in the Netherlands. It is recognized by Vrouwenvelder recognizes that the exceedance frequency of the crest elevation is not necessarily the frequency of failure. There is some probability of failure for lower elevations, and some probability of no failure or inundation above this level as an effort might be made to raise the protection through sandbagging or similar methods. Also, aside from overtopping, piping is found to be the governing mode for the section studies [Vrouwenvelder,1987].

The U.S. Army Corps of Engineers (USACE) introduced probabilistic concepts to levee evaluation in the U.S. in 1991. Prior to that time, planning studies for federally funded levee improvements to existing, non-federal levees were based on the assumption that the existing levee was essentially absent and provided no protection. Following 1991, it was assumed that the levee was present with some probability, which was a function of water elevations, defined as a straight line between two points. The probable failure point (PFP) was taken as the water elevation for which the probability of levee failure was estimated to be 0.85, and the probability of non-failure point (PNP) was defined as the water surface elevation for which the probability of failure was estimated as 0.15 [USACE, 1991].

In 1994, USACE published the report entitled Evaluating the Reliability of Existing Levees. Similar to Vrouwenvelder, the report considered slope stability, underseepage, through seepage and surface erosion. Also it incorporated other information through judgmental probabilities. The overall probability of failure considering multiple failure modes is treated by assuming the failure modes form a series system [USACE, 1994].

In 2000, the National Research Council published Risk and Uncertainty in Flood Damage Reduction Studies (National Research Council, 2000), a critical review of the Corps of Engineers' approach to the application of probabilistic methods in flood control planning. It states"...Such an analysis should consider multiple modes of levee failure (e.g. overtopping, embankment instability), correlation of embankment and foundation properties, hazards associated with flood stage (e.g. debris, waves, flood duration) and the potential for multiple levee section failures during a flood..." [NRC, 2000].

Voortman reviewed the history of dike design in the Netherland’s using probabilistic analysis and concluded the reliability based design of levees is still in a preliminary phase in the Netherlands. Voortman noted that a complete probabilistic analysis considering all variables was explored for dike design in the 1980s, but that "the legislative safety requirements are still prescribed as probabilities of exceedance of design water level and thus the full probabilistic approach has not been officially adopted to date. Probabilistic methods are sometimes applied, but a required failure probability for flood defense is not defined in Dutch law." Risk-based design should focus on other failure mechanisms rather than only overtopping [Voortman, 2002].

Van et al. analyzed the contributions of all levee failure mechanisms in Vietnam and concluded the importance of overtopping is about 52% and failure due to erosion including surface soil erosion and

foundation erosion is about 40%. All other failure mechanisms contributed for about 8 percent [Van. et al, 2007].

5.2.2 Uncertainty, failure probability and reliability in hydraulic engineering

Uncertainty analysis

Uncertainty is attributed to the lack of perfect information concerning the phenomena, processes, and data involved in problem definition and resolution [Tung and Yen, 2005]. Yen and Ang [1971] classified uncertainties into two types – objective uncertainties from any random process or deducible from statistical samples, and subjective uncertainties for which no quantitative factual information is available. Generally, in water resources engineering, uncertainties can be divided into five basic categories: geophysical, transmission, structural, operational, and economics [Tung and Yen, 2005]. Plate et al. classified the uncertainty in flood levee systems into two types: inherent type and knowledge type. Inherent uncertainties are come from natural variation, mainly including temporary flows, wave and stage. Knowledge uncertainties include statistical and model uncertainty. Statistical uncertainty is due to insufficient data from which the parameters in an assumed model are estimated. Model uncertainty resulted from a limited amount of data and limit of knowledge to represent the true random mechanisms of natural process [Plate & Ihringer, 1986].

To quantify uncertainty or statistical features of system outputs or responses affected by the stochastic process, statistical techniques are commonly used in water resources [Tung and Yen, 2006]. These methods include statistical moment methods, PDF estimates and confidence of interval (CI) of parameters and so on. General statistical techniques in uncertainty analysis can be classified into two categories: analytical approaches and approximate approaches [Tung and Yen, 2006]. Each technique in these two categories has different levels of mathematical complexity and data requirements. Analytic methods can be used to determine exact PDF formulas. Analytic methods include derived distribution and integral transform techniques. Integral transform techniques are some well-known integral transforms – the Fourier, Laplace, and exponential transforms and the less known Mellin transform [Epstein 1948]. Although these techniques can be powerful in determining complete information of distribution, analytical approaches are rather restrictive in practical application due to the complexity of most practical problems.

To overcome the drawbacks of analytical methods, approximate techniques in uncertainty analysis are been widely used. Methods for performing uncertainty analysis in hydrologic and hydraulic application are well summarized by Tung and Yen [Tung & Yen, 1993]. The common approximate methods are First-Order Variance Estimation (FOVE) method, probabilistic point estimation (PE) methods and Monte Carlo simulation.

The FOVE method is also called the variance propagation method [Berthouex, 1975]. It estimates uncertainty in terms of the variance of system output, which is evaluated on the basis of statistical properties of the system's stochastic variables. The method approximates the function involving stochastic variables by the Taylor series expansion [Tung & Yen, 1993]. The FOVE method doesn't require knowledge of the PDF of stochastic variables which simplifies the analysis. However, this advantage is also a disadvantage of the method because it is insensitive to the distributions of stochastic variables.

One of common probability point estimation (PE) methods is called Rosenblueth's PE (RPE). The basic idea of RPE is to approximate the original PDF of the random variable X by assuming that the entire probability mass of X is concentrated at two points x_- and x_+. The four unknowns including the locations of x_- and x_+ and the corresponding probability mass p^+ and p^-, are calculated such that the first three moments of the original random variable X are preserved. The drawback of RPE is the number of variables. For N given variables, the number of function evaluations is 2^N. To circumvent this drawback, Harr developed an alternative PE method that reduced the 2^N function evaluations to 2N [Harr, 1989].

Monte Carlo simulation has a long history and a rich literature in uncertainty analysis in water resource engineering benefiting from high speed computer and more sophisticated algorithms [Tung, Yen and Melching, 2006]. Each continuous variable is replaced by a large number of discrete values generated from an assumed underlying distribution; these values are used to compute a large number of values of

function and its distribution. The large number of computations once limited the use of this method, but inexpensive modern computers have largely removed this obstacle. There are also several serious questions of convergence and of randomness in the generated variables. Several so-called variance reduction schemes can be effective in improving convergence and reducing computational effort [Fishman, 1996]. Monte Carlo simulation with variance reduction is particularly helpful in improving the accuracy of the first order reliability method (FORM) results [Baecher and Christian 2003]. Another drawback of Monte Carlo simulation is generating multivariate variates. If the multivariate stochastic variables are correlated with a mixture of marginal distributions: the joint PDF is difficult to formulate.

Probability of failure and reliability analysis

The failure of an engineering component, subsystem, or system can be defined as when the load (L) (external forces or demands) on the system exceeds the system resistance R (strength, capacity, or supply). Risk is the product of probability of failure and consequences of that failure [Tung, Yen and Melching, 2006]. Reliability is defined as the probability of system resistance exceeding the load, i.e., the probability of survival. The mathematical representation of the probability of failure P_f can be expressed as:

$$\mathbf{P_f} = \mathbf{P_r}(\mathbf{L} > R) = \mathbf{P_r}(\mathbf{L} - \mathbf{R} > 0) \quad (5.1)$$

in which $P_r(.)$ denotes probability. The relationship between probability of failure and reliability, R_e is reliability

$$\boldsymbol{R_e} = \mathbf{1} - \boldsymbol{P_f} \quad (5.2)$$

If both load and resistance are uncertain, the probability of failure formula will be: [Kapur and Lamberson, 1977]

$$\mathbf{P_f} = \int_0^\infty \mathbf{f_r}(\mathbf{r})\left[\int_r^\infty \mathbf{f_l}(\mathbf{l})\right]\mathbf{dr} \quad (5.3)$$

Where $f_r()$ and $f_l()$ represent the probability density functions of resistance and loading, respectively.

Generally quantification of the probability of system failure starts with the definition of reliability functions for all potential failure modes of all system elements [Tung, Yen and Melching, 2006]. The general form of a reliability function can be written by:

$$\mathbf{g}(\mathbf{Y}, \mathbf{X}) = \boldsymbol{R}(\boldsymbol{Y}, \boldsymbol{X}) - \boldsymbol{S}(\boldsymbol{Y}, \boldsymbol{X}) \quad (5.4)$$

where R is the resistance of the component, S is the loading on the component, Y is a vector of design variables describing design decisions such as the structural geometry of the component and X is a vector load of random variables.

If the joint probability density function $f_{,R,S}(R, S)$ of the strength R and the load S is known, the probability of failure can be calculated by integration:

$$\mathbf{P_f} = \mathbf{P}\{\mathbf{R} < S\} = \iint \mathbf{f_{R,S}}(\mathbf{R}, \mathbf{S})\mathbf{dRdS} \quad (5.5)$$

If R and S are statistically independent, the following applies:

$$\mathbf{P_f} == \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{\mathbf{R}} \mathbf{f_R}(\mathbf{R})\mathbf{f_S}(\mathbf{S})\mathbf{dR}\right)\mathbf{dS} = \int_{-\infty}^{+\infty} \mathbf{F_R}(\mathbf{R})\mathbf{f_S} \quad (5.6)$$

Usually, the strength and the load are functions of one or more random variables. In such a case the reliability function can be rewritten as Z = R - S = $g(X_1, X_2 \ldots X_n)$. If the variables X_1, X_2, ..., X_n are statistically independent, the equation can be simplified to:

$$\boldsymbol{P_f} = \iiint_{Z<0} \boldsymbol{f_{X1}}(\boldsymbol{X_1})\boldsymbol{f_{X2}}(\boldsymbol{X_2}) \ldots \boldsymbol{f_{Xn}}(\boldsymbol{X_n})\boldsymbol{dX_1 dX_2} \ldots \boldsymbol{dX_n} \quad (5.7)$$

This integral can seldom be determined analytically. The solution is therefore usually calculated with numerical methods. The overall failure probability of a system component is then given combing the failure probability for all considered failure modes:

$$\mathbf{P_f^{Overall}} = \mathbf{P}(\mathbf{Z_1} < 0; \mathbf{Z_2} < 0; \ldots \mathbf{Z_i} < 0; \ldots \mathbf{Z_m} < 0) \quad (5.8)$$

where (Z_1<0; Z_2<0; ... Z_i<0;... Z_m<0) denotes at least one of m failure mechanisms occurs. The overall system failure probability is determined in a similar way as that of system components considering the correlation between components. Several methods are available in calculating exactly the system failure probability including fault-tree analysis with numerical integration and/ or Monte Carlo simulation [Tung, Yen and Melching, 2006].

Quantification of uncertainties

Computing the probability of failure (P_f) requires knowledge of probability distributions of load and resistance, or the performance function [Tung and Yen, 2005]. Three main methods have been used in hydraulic structures: risk analysis including direct integration method, Mean-Value First-Order Second-Moment (MFOSM) Method, and Monte Carlo simulation.

Direct integration requires PDFs of load and resistance or the performance function be known or derived. This information is seldom available in practice because of the complexity of hydrologic and hydraulic models. Explicit solution of direct integration can be obtained for only a few PDFs.

The MFSOSM method for failure probability analysis employs the FOVE method to estimate the mean and standard deviation of the performance function from which the reliability index is computed. Several studies have shown that failure probability is not greatly influenced by the choice of distribution for the performance function and the assumption of a normal distribution is quite satisfactory except in the tail portion of a distribution [Tung, Yen and Melching, 2006]. The MFSOSM method has been used widely in various hydraulic structures such as storm sewers, levees and open channels. However, this method has several drawbacks. It requires accurate extreme probabilities which are hard to estimate. Other drawbacks are inappropriate choice of the expansion point, general poor estimation of mean and variance for highly non-linear function, and inability to handle distribution with large skew coefficient.

5.2.3 Uncertainties in levee overtopping

Overtopping is when flood water stages exceed levee's crest, i.e., the flood flow q exceeds the levee capacity q. Therefore, the probability of overtopping is:

$$\mathbf{P_f} = \int_{\mathbf{0}}^{\infty} \mathbf{fq(q)}\,[\int_{\mathbf{q}}^{\infty} \mathbf{fQ(Q)}]\mathbf{dQ} \quad (5.9)$$

In this case, the resistance of a levee system is essentially the channel capacity of the levee, and the loading is the magnitude of flows passing the levee. Traditionally, the capacity of the structure has been taken as deterministic, or its uncertainty has been considered small enough to be ignored, as compared to the hydrologic uncertainties. Plate and Duckstein presented four different probability-based concepts for levee design and analysis. Level I, traditional concepts, uses the exceedance probability, PE as the performance index. Level II analysis considers both loading and resistance as random variables and the these random variables are assigned as Gaussian distributed. Level III is based on any given distributions of load and resistance. Level II and III are appropriate for evaluation of reliability. Level IV is also based on the joint probability density function for load and resistance; and also, it requires the assignment of a consequence function to each combination of resistance and loads [Plate and Duckstein, 1988].

Ignoring the uncertainty of the hydraulic structure may lead to an underestimation of the failure probability. To determine the failure probability or reliability of a structure, knowledge of probability distributions for loading and resistance is required. This is referred to as load-resistance inference [Tung and Mays, 1981a].

To find levee failure probability including both structure failure and overtopping, Wood considered the resistance of levee as a random variable. However, no hydraulic parameters are included, so levee failure was described using uniform and quadratic distributions of an exceedance discharge at which the levee structurally fails [Wood, 1977].

Tung and Mays analyzed the various uncertainties including hydrologic uncertainty and hydraulic uncertainty in levee design and defined the risk and reliability of overtopping. Hydrologic uncertainties were estimated by generalized values of parameters through regression analysis. Hydraulic uncertainties were estimated by first-order analysis through Manning's equation [Tung and Mays, 1981b]

Kuo et al. assessed the overtopping risk of Feitsui Dam by five uncertainty analysis methods including Rosenblueth's point estimation method, Harr's point estimation method, Monte Carlo simulation, Latin hypercube sampling, and the mean-value first order second moment method. The results show that values of overtopping failure probability computed by different methods are similar. The selection and application of the uncertainty methods depend upon the information available for the model parameters and model complexity [Kuo et al., 2007].

Load uncertainties: hydrologic uncertainty

Hydrologic uncertainties can be classified into three types [Wood, 1975; Bras, 1970]: 1) inherent uncertainty due to the stochastic nature of the hydrologic process, 2) model uncertainty resulting from a limited amount of data available for assessing the true random mechanism of the hydrologic process, and 3) parameter uncertainty due to an insufficient amount of data for estimating parameter values in the assumed model. Numerous articles in the literature concern hydrologic models for describing flood magnitudes to help describe the probability of overtopping. Bras [1979] presented an excellent review of many of these models which only account for hydrologic uncertainties.

Bayesian techniques have been developed for treating hydrologic uncertainties in levee analysis. Wood developed flood frequency curves by Bayesian methods [Wood, 1977]. To model hydrologic uncertainty, Tung used regional regression methods and Bayes theorem to estimate parameter and model uncertainty. Tung developed generalized values of the parameters (mean, standard deviation, and skew coefficient) through the use of a weighting between sample and regional parameter estimates [Tung and Mays, 1981a].

Generally in water resources engineering, hydrologic uncertainties are quantified by frequency analysis. Stedinger et al. presented an excellent review of many of these models of frequency methods [Stedinger et al, 1993]. In this study, the hydrologic variables come from reservoir's regulated flow. The regulated/unregulated relationship and both probability distributions are presented in previous chapter. From the model, the main hydrologic variables, flood flow magnitude and duration, are expressed in PDF formulas.

Resistance uncertainties: hydraulic uncertainty

Hydraulic uncertainties for the design/analysis of hydraulic structures may be divided into several types: model, construction materials conditions, and operational flow conditions. Model uncertainty results from using a deterministic hydraulic model to describe the flow conditions through or over the structure. The other uncertainties relate directly to parameters and variables in the hydraulic model. The capacity of a flood levee system is described as a function of several parameters which can be considered as random variables. These parameters could statistically dependent or independent of each other. In practice, it is difficult to determine the uncertainty aspects of the structure as a whole. Alternatively, it is easier to use more parameters or components and then derive uncertainty properties for the structure.

Johnson presented an overview of hydraulic uncertainties analysis in water resources engineering [Johnson, 1996]. To quantify various hydraulic parameter uncertainties in terms of the coefficients of variation and associated distributions, several methods were evaluated. From this overview, the first-order analysis is the primary method for including uncertainties in hydraulic parameters [Johnson, 1996]. Common hydraulic parameters are Manning value, channel slope, particle size, and flow velocity.

Lee and Mays use first-order to examine the uncertainty in levee capacity. The levee capacity was assumed to be a function of the Manning roughness, friction slope, cross-sectional area and wetted perimeter. The coefficients of variation used in this study were arbitrarily chosen and all distributions were assumed to be normal [Lee and Mays, 1986].

In another study by Tung and Mays, both hydrologic and hydraulic uncertainties in flood levee's design were analyzed to define the risk and reliability of overtopping [Tung and Mays, 1981]. In hydraulic uncertainty analysis, six hydraulic variables are considered random variables: channel top width, flow cross-sectional area of channel at full bank, width of encroachment of levees, traverse slope of the floodplain, slope of floodplain along the channel and channel bottom slope. Manning's equation was used to calculate the capacity's probability distribution from those 6 uncertain parameters. Levee capacity Q_r

was estimated by first-order analysis. Finally Q_r's probability distribution is assumed to be log-normal [Tung and Mays, 1981b].

Zhao and Mays used Rosenblueth's point-estimate to calculate the mean and standard deviation for a fan arc width in FEMA alluvial-fan delineation The result is used to estimate the flood probability for a given location in north Scottsdale, Arizona [Zhao and Mays, 1996]. USACE has studied hydrologic, hydraulic and economic uncertainties for flood damage reduction [USACE, 1996]. In this approach, hydraulic uncertainty is from discharge and water stage relationship curves. USACE used confidence intervals to quantify hydraulic uncertainties by Monte Carlo methods.

In this study, levee resistance capacity Q_r will be analyzed as a probability distribution using first order analysis. The first-order analysis yields estimates of the mean and the variances of the respective contributing variables in the deterministic flow equation.

The total levee capacity, Q, can be estimated by using Manning's equation given by Chow [1959] as

$$\boldsymbol{Q = 1.49\left(\frac{1}{N_c}A_c^{5/3}P_c^{-2/3} + \frac{2}{N_b}A_b^{5/3}P_b^{-2/3}\right)S_f^{1/2}} \qquad (5.10)$$

in which N_c and N_b are roughness coefficients for the channel and the flood plain, respectively; S_f is the friction slope; A_c, P_c, and A_b, P_b are the cross-sectional area and the wetted perimeter of the channel and the overbank flow, respectively.

The levee capacity is considered as a random variable related to independent random variables, N_c, N_b, Ac, P_c, A_b, P_b, and S_f. The coefficient of variation of the random variable serves as a plausible index to represent the variation of a random variable relative to its mean value. Applying the first-order analysis to Eq. 5.10, the coefficient of variation of Q according to Tung and Mays [Tung and May, 1981b] is:

$$\boldsymbol{\Omega_Q^2 = \frac{1}{4}\Omega_{sf}^2 + \frac{1}{\Psi^2}\left(\Omega_{NC}^2 + \frac{25}{9}\Omega_{AC}^2 + \frac{4}{9}\Omega_{PC}^2\right) + \frac{1}{\Phi^2}\left(\Omega_{Nb}^2 + \frac{25}{9}\Omega_{Ab}^2 + \frac{4}{9}\Omega_{Pb}^2\right)} \qquad (5.11)$$

$$\text{in which} \quad \boldsymbol{\Psi = 1 + 2\left(\frac{N_c}{N_b}\right)\left(\frac{A_b}{A_c}\right)^{5/3}\left(\frac{P_c}{P_b}\right)^{2/3}} \qquad (5.12)$$

$$\text{and} \quad \boldsymbol{\Phi = 1 + \frac{2\Psi}{\Psi - 1}} \qquad (5.13)$$

Ω represents the coefficient of variation of the random variables and the bar denotes the mean value. The coefficient of variation of Q, Ω_Q, is the measure of hydraulic uncertainty and is related to the individual uncertainty of each variable in Manning's equation.

According to Lee and Mays, only a few contributing variables, namely slope of energy line (S_f), Manning coefficient of bank (N_b), and Manning coefficient of channel (N_c) dominate the hydraulic capacity uncertainty in this example. Conversely, cross area of channel (A_c), wet perimeter of channel (P_c), cross area of bank (A_b), and wet perimeter of bank (P_b) have much less influence on the hydraulic uncertainty since they are estimated from direct measurements in field surveys. Treating random variables with negligible uncertainties as deterministic variables simplifies the procedure for risk and reliability evaluation [Lee and Mays, 1986].

The uncertainty of each contributing random variable varies from case to case. With sufficient knowledge of the characteristics of the individual random variables, an experienced engineer could treat hydraulic uncertainty in a simple manner by ignoring uncertainties of minor importance.

The probability distribution of the levee capacity Q_r, is assumed to be log-normal. It needs to be justified by the multiple relationship of the most uncertain parameters such as N_c and S_f. The mean and coefficient of variation of the levee capacity can be computed by equation 3.6 in Chapter 3. Knowing these statistical parameters, the probability density function of the levee capacity becomes:

$$\boldsymbol{f_r(Q_r) = \frac{1}{2\pi^{1/2}}\exp\left[-\frac{1}{2}\left(\frac{\ln Q_r - \mu_{Q_r}}{\sigma_{Q_r}}\right)^2\right]} \qquad (5.14)$$

Where μ_{Qr} σ_{Qr} are, respectively, the mean and standard deviation of the transformed levee capacity, $\ln(Q_r)$. The coefficient of variation of the levee capacity, obtained by equation (5.14), can be used to compute the standard deviation σ_{Qr}, for the distribution by using: [Tung and Mays, 1981]

$$\boldsymbol{\sigma_{Q_r} = [\ln(\Omega_Q^2 + 1)]^{1/2}} \quad (5.15)$$

5.2.4 Uncertainties in levee erosion

Levee erosion is from high velocity flows scouring the levee to reduce the levee cross section, especially near the crown. Figure 5.3 shows the general process of levee erosion. As flood stages increase, surface erosion potential increase from two sources: (1) excessive current velocities parallel to the levee slope and (2) erosion due to wave attack directly against the levee slope. Therefore, loads for erosion are velocity and wind speed and resistance is from levee materials and construction.

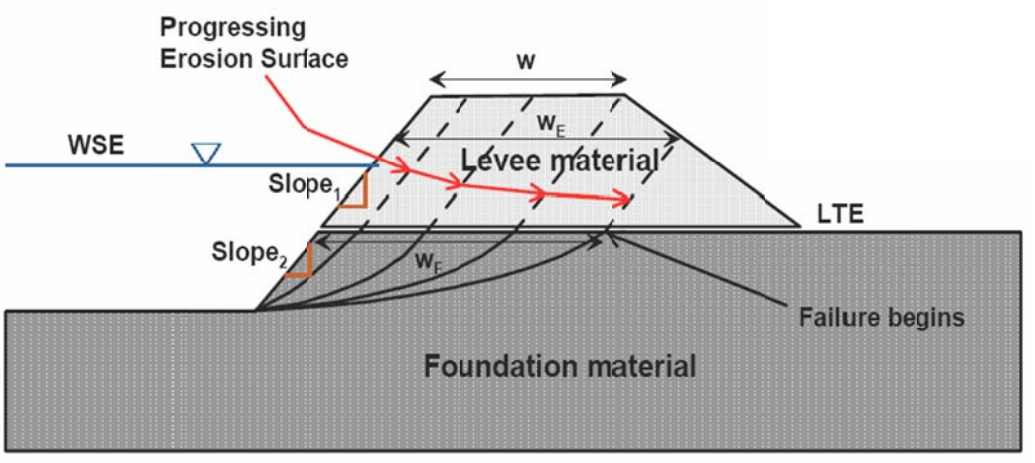


FIGURE 5.2 LEVEE EROSION PROCESS SKETCH (W_E IS EFFECTIVE LEVEE WIDTH, $SLOPE_1$ IS LEVEE WATERSIDE SLOPE, $SLOPE_2$ IS LEVEE FOUNDATION SLOPE, WSE IS WATER SURFACE ELEVATION)

Uncertainties in soil properties

As with hydrologic/hydraulic uncertainty, geotechnical uncertainties are from both natural variability and knowledge uncertainty [NRC, 2000]. However, since geotechnical parameters involve many soil properties, they will have unique characteristics [Christian et al. 1994]. Christian divided the geotechnical uncertainty into two types including data scatter and systematic error. Data scatter can be further divided into actual spatial or temporal variation and random measurement error. To remove the random measurement error, common approaches are the methods of moments and the method of maximum likelihood estimators. Systematic error can also be divided into systematic error in the trend and bias in measurement procedures.

The selection of random variables varies with failure mode. To evaluate levee failure probability, Wolfe used unit weight, drained strength of sand and drained and undrained strength of clays for instability analysis and permeability for seepage analysis [Wolfe, 2008]. However, in this study, since only erosion failure is considered, the soil's strength including shear stress and erodibility contributing to erosion rate will be considered as random variables.

Similar to hydraulic uncertainty quantification, geotechnical uncertainty analysis also includes four approximation methods: direct calculation, First-order method, point estimate method and other statistical methods [Christian, 2004]. The direct method formula is very simple M = R - Q where R is loading factor while Q is resistance. M indicates overall performance of hydraulic structure, i.e., when M is negative, the structure fails. The distribution of M indicates overall reliability. The mean of M is the difference in means of R and Q. The variance of M is the sum of variance of R and Q. From this theory, the failure probability index can be calculated as function of mean of variance of R and Q. Finally P_f is a normal distribution.

This method is straightforward. However, the assumption of this method is that all the random variables are normally distributed [Christian, 2004]. As in hydraulic engineering, First Order Variance estimation (FOVE), Point Estimation (PE) and Monte Carlo simulation are frequently to be used in geotechnical engineering [Christian, 2004].

Erosion failure definition

As in figure 5.3, levee resistance is a function of levee type, levee geometry, levee material characteristics, armoring and vegetation. The load is a function of hydraulic stress which is function of water surface elevation and velocity. Therefore, the levee erosion failure probability is represented as a function of erosion width (load) and levee width (resistance).

$$\boldsymbol{P_f = (\varepsilon D > W_E)} \qquad (5.16)$$

Where ε is erosion rate, D is duration of certain velocity, W_E is effective levee width. Failure is said to occur when the calculated total erosion exceeds the effective width.

Several erosions studies have been performed to identify the erosion parameters and correlate those parameters to formulate an expression (i.e., a physically-based model) for erosion rate [Hanson and Temple, 2002]. The overall equation is:

$$\boldsymbol{\varepsilon = k * (\tau - \tau_c) * T} \qquad (5.17)$$

Where k = erodibility coefficient or detachment rate coefficient (ft^3/lb-hr); τ = effective hydraulic stress on the soil boundary (psf); τ_c = critical shear stress (psf); T = Adjustment factor; ε = erosion rate ft/hr.

The erosion rate (ε) is function of both hydraulic (τ) and geotechnical (k, τ_c) parameters. *T* mainly depends on characteristics of water-soil boundary, current velocity and wind wave height and period. Both k and τ_c are functions of the engineering properties of the levee materials, which are inherently uncertain.

Erosion rate as a function of flow velocity can be measured in the laboratory using one of several devices such as the Erosion Function Apparatus (EFA) [Briaud et al., 2001a and 2001b]. The critical shear stress, τ_c is defined as the shear stress corresponding to a rate of erosion of 1 mm/hr in the EFA. While useful for analytical studies, this method is impractical in engineering fields. Alternatively, the critical shear stress can be estimated using empirical correlations between the critical shear stress and soil index properties. Several empirical correlations between critical shear stress and soil index properties such as grain size, plasticity index and shear strength are available to estimate the value of τ_c. Same as to measure shear stress, coefficients of eordibility k measurement is performed by jet testing in ASTM D 5852. However, site-specific tests will be impractical. Therefore, in a manner similar to the method used to evaluated critical shear stress, erodibility of levee materials has been estimated by empirical correlations with soil index properties. To simplify the calculation, erosion resistance of the levee materials has been divided into five broad classes related to their ASTM classifications, as shown in Table 5.1 [Briaud et al., 2001a, 2001b, Hanson and Simon 2003].

Table 5.1 Strength of levee soil characteristics [Briaud et al., 2001a, b; Hanson and Simon 2003]

Levee soil type	ASTM typical soil type	Critical Stress, $\tau_{s\,(psf)}$	Erodibility coefficient, k (ft^3/lb-hr)
Very Resistant	Boulders and Cobbles	4.869	0.005
Resistant	Gravel(GP-GW)	1.058	0.021
Moderately Resistant	Clay(CL,CH,SC,GC)	0.094	0.094
Erodible	Sand(SP,SM, and mixture)	0.014	0.409
Very Erodible	Silt(ML)	0.003	1.867

The hydraulic stress due to velocity is function of water density, current friction factor and velocity. The express is as followed:

$$\boldsymbol{\tau = \frac{1}{2}\rho * f_c * v^2} \qquad (5.18)$$

Where ρ is mass density of water (lbm/ft^3); and

$$\boldsymbol{f_c = 2(2.5(\ln(\frac{30h}{k_b}) - 1)^{-2})} \qquad (5.19)$$

Where f_c is current fiction factor (dimensionless) [DHI, 2007]; and h = water depth; k_b = bed roughness (ft); v = current velocity.

Also from first principles, shear stress can be calculated from:

$$\boldsymbol{\tau = \quad DS_w} \qquad (5.20)$$

Where τ = Shear Stress (N/m^2), = Weight Density of Water (N/m^{3}or lb/ft^3), D = Average water depth (m or ft), and S_w = Water Surface slope (m/m or ft/ft).

Again, the channel/levee velocity can be calculated [Chow, 1957] and downstream boundary conditions:

$$\boldsymbol{v = \frac{Q}{A} = Q/(A_c + A_L)} \quad (5.21)$$

5.3 METHODS

5.3.1 Flow chart of computation levee overall failure probability

Figure 5.4 shows the flow chart for computing of overall levee failure. Three steps are involved. First is to estimate hydrological, hydraulic and geotechnical uncertainty. Then overtopping and erosion caused levee failure are calculated. Finally the overall failure probability is calculated.

5.3.2 Illustrative example for levee overtopping and erosion failure

This section presents a simple example to demonstrate the above framework. The illustrative example includes the parameters of a simple levee, uncertainties in hydrology, hydraulics and soil properties, overtopping and erosion failure probabilities and overall levee failure probability.

1) Simple levee geometric parameters and uncertainty in hydrology, hydraulics and soil properties

Figure 5.5 shows a simple flood levee\channel system. A compound levee\channel system is assumed consisting of a irregular main channel with wet perimeter P_c and bank-full cross area A_c which normally contains flow within bank-full conditions, and two trapezoidal levees over each side of the main channel, carry overbank flow during floods. The levee heights on each side are H_{LL} and H_{LR} respectively. The full levee cross section area is A_L. The example reach is 1,000 ft long.

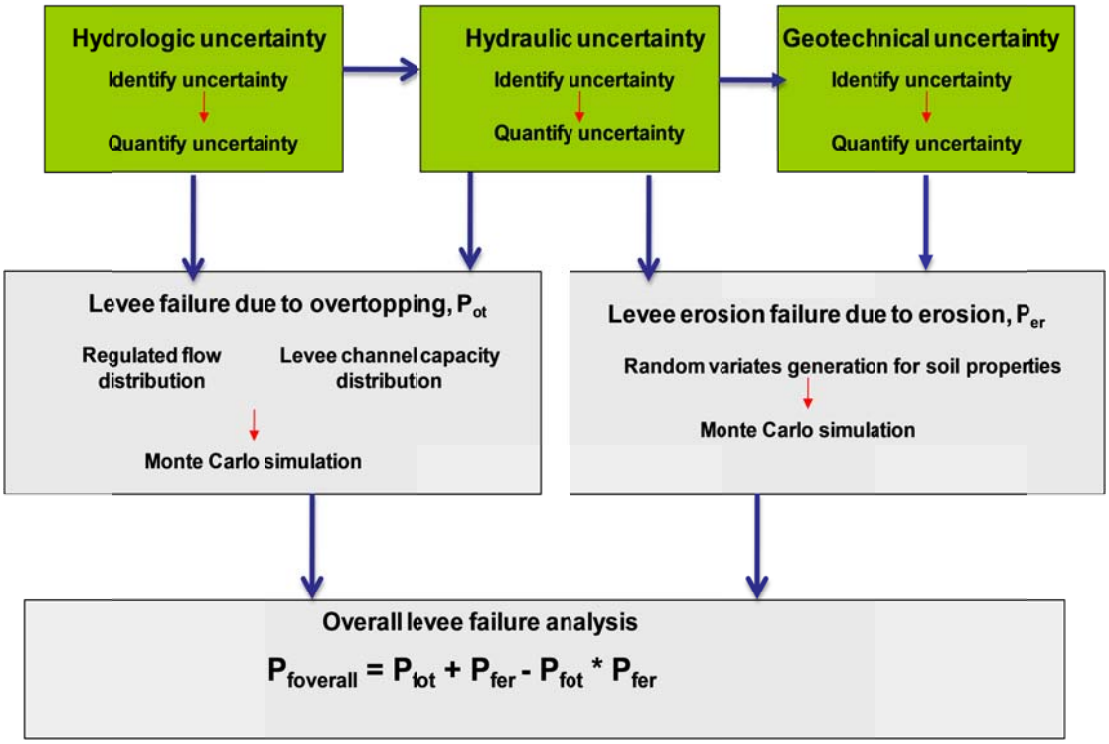


FIGURE 5.3 FLOW CHART OF COMPUTATION OF OVERALL LEVEE FAILURE

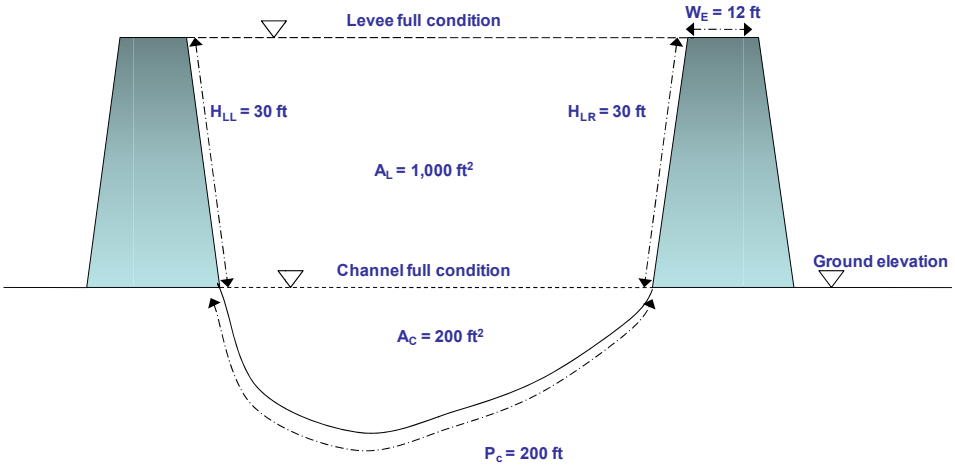


FIGURE 5.4 A SIMPLE LEVEE/CHANNEL SYSTEM SKETCH (NOT TO SCALE)

The key geometric parameters and associated probability distributions are shown in Table 5.1. For this simple example, all random variables have Log Normal distributions, with a mean μ and standard deviation σ . The mean M and standard deviation S of this lognormal distribution are：

$$\mathbf{M} = \frac{1}{2}\ln\left(\frac{\mu^2}{(\frac{\sigma}{\mu})^2+1}\right) \qquad (5.21)$$

$$\mathbf{S} = \sqrt{\ln\left((\frac{\sigma}{\mu})^2 + 1\right)} \qquad (5.22)$$

The PDF of random variable is shown in Eqn. 5.14 and Eqn. 5.23 is:

$$\mathbf{P(x)} = \int_{-\infty}^{+\infty} \left[\frac{\mathbf{1}}{\mathbf{xS\sqrt{2\pi}}} \mathbf{exp}\left(\frac{-[\mathbf{\ln(x)-M}]^{\mathbf{2}}}{\mathbf{2S^2}} \right) \right] \mathbf{dx} \qquad (5.23)$$

Table 5.2 Levee geometric variable and distribution

Levee geometric variables	Distribution type	Distribution parameters	
		mean (μ)	sd (σ)
Channel Cross area (A_c: ft^2)	LN	200	40
Levee Cross area (A_L: ft^2)	LN	300	100
Channel wet perimeter (P_c: ft)	LN	200	50
Levee wet perimeter (P_L: ft)	LN	260	40
Channel slope (S_f: ft/ft)	LN	0.001	0.0005
Channel Manning value	LN	0.045	0.005
Levee Manning value	LN	0.065	0.01

Flood flow in this example also follows a Log Normal distribution. The mean and standard deviation of flood flow are 1,000 CFS and 300 CFS, respectively. To simplify the computation, the flood duration is assumed as deterministic.

As with levee geometric and hydrologic variables, soil strength including critical stress and erodibility coefficients are assumed as Log Normal distributions. Table 5.3 presents the distribution parameters [Briaud,et al, 2001a, 2003, Hanson and Simon, 2003].

Table 5.3 Critical soil strength and distribution

	Critical Stress, τ_s (psf)			Erodibility coefficient, k (ft^3/lb-hr)		
Soil type	Distribution	mean (μ)	sd (σ)	Distribution	mean (μ)	sd (σ)
Very Resistant	LN	4.869	0.005	LN	0.446	0.952
Resistant	LN	1.058	0.021	LN	0.560	1.101
Moderately Resistant	LN	0.094	0.094	LN	0.917	0.800
Erodible	LN	0.014	0.409	LN	1.089	0.440
Very Erodible	LN	0.003	1.867	LN	0.785	0.473

Sources: Briaud,et al, 2001a, 2003, Hanson and Simon, 2003

Using Equation 5.21, the flood velocity can be calculated. Assume there is no hydraulic routing. Table 5.4 shows the velocity values with different water depths. Here the bank-full flow is 400 cfs.

Table 5.4 Summary of Hydraulic variables for example

WSE from bankfull level	Mean Velocity (ftp)	Hydraulic Radius, R (ft)	Energy Grade Line, S (slope)
0	2.00	0.77	0.0005
5	2.40	0.93	0.0005
10	3.33	1.07	0.0005
15	3.50	1.33	0.0005
20	3.00	1.76	0.0005
30	2.50	1.95	0.0005

2) Failure probability calculation

Overtopping failure probability

The levee capacity is calculated from Equation 5.10. Six variables shown in Table 5.2 are assumed as independent. Monte Carlo simulation is employed to find P_{fOT}.

$$\mathbf{P_{fOT}} = \frac{\mathbf{N_f}}{\mathbf{N}} \qquad (5.24)$$

Where N_f is the occurrences when levee capacity Q_c is less than flood flow Q, N is the total number of realizations. In this example, 100, 1,000 and 10,000 realizations are used to compare convergence. It appears the model will converge when the realization reach to 5,000. The result shows the P_{fOT} in this levee/channel system is 0.12 in each year, i.e., about 8.3 year recurrence interval.

Erosion failure probability

Levee erosion failure probability is estimated using Equation 5.16. To simplify the computation in this example, only river stream-induced erosion was considered. The soil property was assumed as moderately resistant. After flood erosion, if the levee's effective width is eroded by 25 %, the levee is declared as failed. 100, 1,000 and 10,000 realizations are used to compare convergence. The results show the P_{fer} in this levee system is 0.05.

Overall failure probability

The overall failure probability P_{fov} is 0.164. Figure 5.6 shows the Monte Carlo results with 100, 1,000 and 10,000 realization. It concludes 10,000 realizations can reach convergence.

5.4 APPLICATION ON LOWER MOKELUMNE RIVER LEVEE SYSTEM

This section presents the site description, data set and two types of failure probability calculation for the Lower Mokelumne River levees.

5.4.1 Site, Data Set and Software Description

Mokelumne River is a major tributary of San Joaquin River in California. Figure 5.6 shows the location of this reach and watershed. Ten major flood events have occurred in this river in the past 50 years with four occurring in the past 20 years. These four floods have accounted for an average per event flood damage value of $4 million [DWR, 2006]. The Mokelumne River watershed covers approximately 920 square miles of mountainous to valley floor terrain. Elevations range from a peak above 8,800 feet msl to slightly below sea level in the vicinity of the Delta. The Mokelumne River is highly regulated by reservoirs for waters supply and power generation, with Camanche Reservoir providing flood control capacity. The lower Mokelumne River watershed below Camanche Dam includes over 70,800 acres of cropland and nearly 60,300 acres of orchards and vineyards. This area also includes the communities of Clements, Lockeford, Lodi, and Woodbridge [Robinson and Bryon, 2006].

The Lower Mokelumne River watershed is protected by two reservoirs and about 22 mile of levees. Figure 5.7 shows the typical cross section of the levee in this reach. Historically, levees on the north side of river are higher than the south side. In this study, only south side of levees was analyzed. The required data in this study includes hydrologic data, geometric data, hydraulic data and soil data. Table 5.5 presents the levee geometry data. Hydrologic data are from the previous chapter. The flood regulated flow follows a normal distribution. Three reservoir operation scenarios with storages of 200, 180 and 150 TAF are selected. Table 4.2 in Chapter 4 presents the detailed hydrological data.

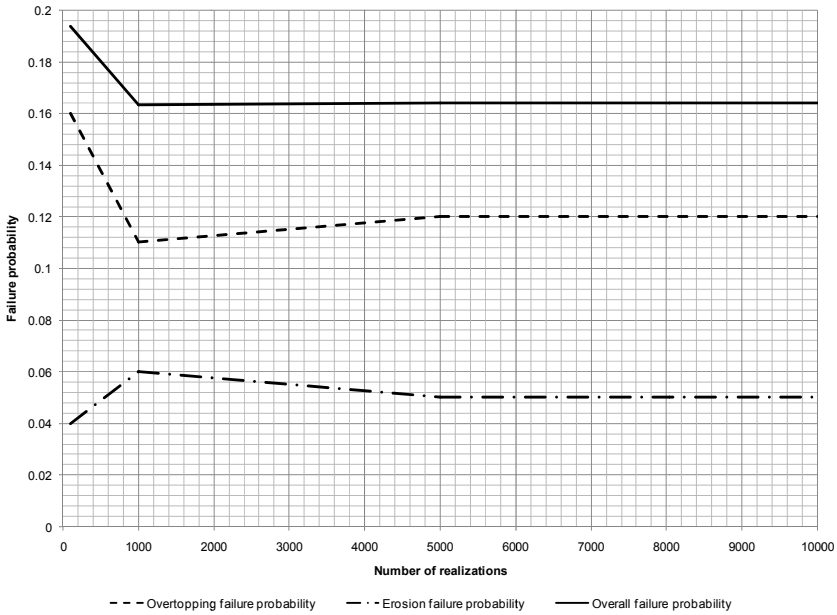


FIGURE 5.6 MONTE CARLO RESULTS WITH 100, 1,000 AND 10,000 REALIZATION

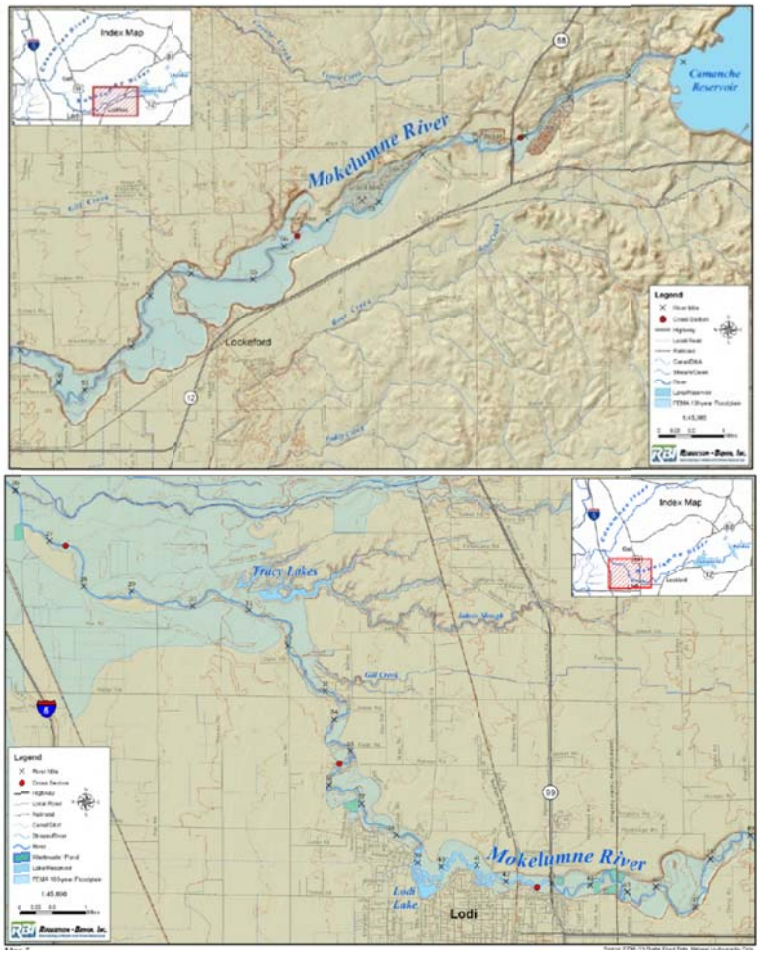


FIGURE 5.7 LOWER MOKELUMNE RIVER AND LEVEE SYSTEM (FROM USGS AND ROBINSON AND BRYON, 2006)

Table 5.5 Levee geometric variable and distribution in Lower Mokelumne River levee

Levee geometric variables	Distribution type	Distribution parameters	
		mean (μ)	sd (σ)
Channel Cross area (A_c: ft^2)	LN	500	200
Levee Cross area (A_L: ft^2)	LN	2,000	500
Channel wet perimeter (P_c: ft)	LN	400	100
Levee wet perimeter (P_L: ft)	LN	800	200
Channel slope (S_f: ft/ft)	LN	0.001	0.0005
Channel Manning value	LN	0.045	0.005
Levee Manning value	LN	0.065	0.01

Table 5.5 Levee geometric variable and distribution in Lower Mokelumne River levee

As in the illustrative example, six hydraulic variables are selected as random variables including S_f, A_c, P_c, N_c, P_b, A_b and N_b. All these six variables are assumed to be independent. Levee capacity Q_r is calculated by Manning's equation in Equation 5.10. The other hydraulic variable statistics are estimated using EBMUD's existing HEC-RAS model [EBMUD, 2009]. The HEC-RAS model is used to calculate hydraulic conditions given flow data and existing geometric condition. Through modeling of full bank run and full levee runs, all variables can be tabulated. There are about 500 cross sections in this reach. From cross section to cross section, some variables vary greatly, such as A_c, P_c, A_b and P_b. Some variables vary slightly, such as N_c, N_b, and S_f. All variables are assumed to have log-normal distributions. Soil properties are acquired from Department of Water Resources Levee Repair program [DWR, 2009]. The levee soils mainly fall into the moderately resistant category.

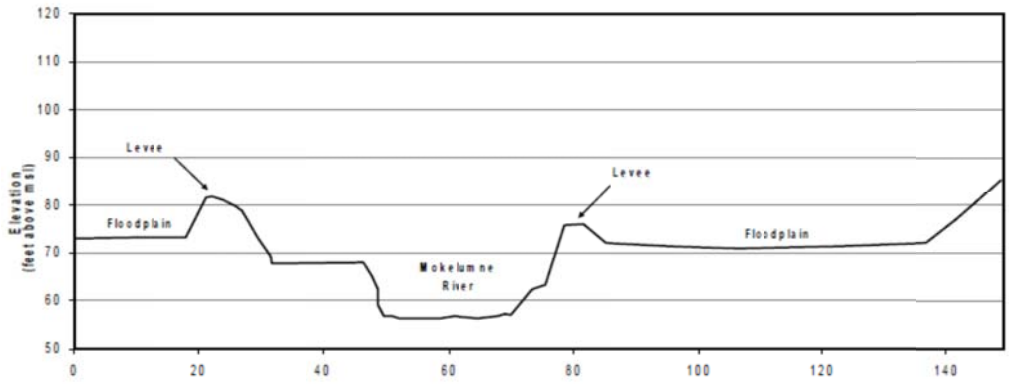


FIGURE 5.8 ONE TYPICAL LEVEE CROSS SECTION IN LOWER MOKELUMNE RIVER (ROBINSON AND BRYON, 2006)

5.4.2 Levee failure probability

Similar to the example above, overtopping risk here is defined as load-resistance interaction and Monte Carlo simulation is employed to find the failure probability. Here, 10,000 realizations are selected. The result shows the P_{fOT} in this levee/channel system is 0.10, 0.12 and 0.15 per year under flood storage scenarios of 200, 180 and 150 TAF.

The failure due to erosion criteria is $P_f = P(\varepsilon D > 0.25 * W_e)$, i.e., if the effective levee is eroded by 25% in width, the levee is declared failed [DWR, 2009]. Wind wave caused erosion is omitted. After all the variables are identified and quantified, Monte Carlo simulation is employed to find the Pf, i.e., $P_{foverall} = P_{fov} + P_{fer} - P_{fov} * P_{fer}$. Here, 10,000 realizations are selected. Again, three scenarios for flood storage capacity were compared. Here, levee overtopping and erosion failure are assumed to be independent. The final overall failure probability was calculated based on these two failure mechanisms. Table 5.6 presents the results. Figure 5.9 shows the plot of Monte Carlo results. From the results, the current return of period of

overall levee failure in Lower Mokelumne River is about 7 years. This is verified by EBMUD's bottomland flooded map [EBMUD, 2009]. Besides this map, the survey was performed to account flooding caused by levee failures in last 10 years along Upper reach of Mokelumne River. There are total 37 levee failures resulting in flooded areas in last 10 years. The flooded area's land use is mainly for agriculture purpose. The main causes are overtopping, seepage and broken levee. The levee is designed against 15 years flood. However, due to poor maintenance, the current return of period of overall levee failure is about 7 years.

With the flood storage is reduced from 200 TAF to 150 TAF, the overall return of period of levee failure decrease from 7 years to 4 years.

Table 5.6 Lower Mokelumne River Levee Failure Analysis Summary

Reservoir Re-operation	Overtopping failure	Erosion failure	Overall failure	Return of period (Year)
Storage = 150 TAF	0.15	0.09	0.2265	4
Storage = 180 TAF	0.12	0.08	0.1904	5
Storage = 200 TAF	0.1	0.05	0.1450	7

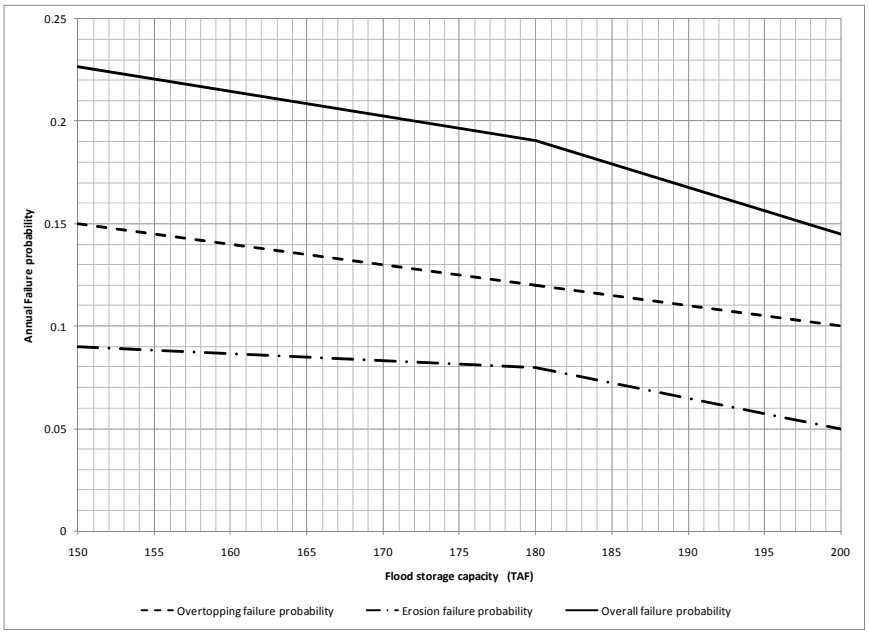


FIGURE 5.9 FAILURE PROBABILITIES FOR FLOOD STORAGE CAPACITY IN CAMANCHE/PARDEE RESERVOIRS

5.5 CONCLUSIONS AND LIMITATIONS

This chapter introduces a framework to estimate overall levee reliability due to overtopping and erosion, incorporating multiple uncertainties including hydrologic, hydraulic and geotechnical factors. Two main contributions include overall failure probability estimation and load-resistance interference risk analysis. Overtopping and erosion failure probability analysis are usually performed separately in water resources engineering and geotechnical engineering, respectively. This chapter presented a comprehensive approach combining these two failure mechanisms. Also load-resistance interference reliability analyses were introduced to perform overtopping between flood magnitude and levee capacity and erosion failure

between velocity and soil strength. Both analyses are performed by Monte Carlo simulation to determine overall levee failure probability.

Some shortcomings in this study include less consideration of other failure mechanisms, assumptions of independence and normality assumptions.

1) More failure mechanisms. The levee failure is complicated, involving many potential failure mechanisms. This approach only considers overtopping and erosion failures. Other failure mechanisms such as under seepage and through seepage are also common causes of levee breaches. More research is needed for seepage failure analysis.

2) Assumption of independence. This study assumes failure modes are independent and uncorrelated. This is not necessarily true, as some conditions increasing the probability of failure for one mode are likely increase the probability of failure by another. In reality, overtopping and erosion usually occur in similar conditions of high water. More research focus on correlation between different failures modes is needed.

3) Normality assumption. In this study, most of random variables are assumed as log-normally distributed including hydrologic variables, flow magnitude, flood duration, hydraulic variables, channel cross section area, levee cross section areas, Manning's n values, and geotechnical variables, soil shear stress, erodibility coefficient. In fact, these variables probability distributions vary from site to site. More work is needed on these distributions.

REFERENCES

Ang, A., Tang, W. (1984). "Probability Concepts in Engineering Planning and Design, Vol.II: Decision, Risk, and Reliability", John Wiley and Sons, New York.

Baecher, G., Christian, J. (2003) "Reliability and Statistics in Geotechnical Engineering". John Wiley and Sons, Inc: New York.

Berthouex, P., (1975). "Modeling concepts considering process performance variability, and uncertainty", Mathematical Modeling for Water Pollution Control Process, ed. By T.M.Keinath and M.P.Wanielista, Ann Arbor Science, Ann Arbor, MI, 405-439.

Bogardi, I. (1968). "Flood exposure recommended as a parameter for describing the fatigue loading on flood control structures", Hydrological Science Journal, Vol13(3): 14-24.

Bogardi,I ., L. Duckstein, and Szidarovszky, (1975). "On the reliaility of flood levees systems, paper presented at Second Conference on Application of Statistics and Probability in Soil Structural Engineering, Aache, West Germany, September 15-18, 1975..

Bras, R.L. and S.-O. Chan (1978) "Theoretical Models of Hydrologic Parameters Using Derived Distribution Techniques in Proc.", Vol. II, Int. Syrup. on Risk and Reliability in Water Resources, Ontario, Canada, 1978.

Briaud,J., Ting, F., Chen, Gudavalli, S., Perugu, S., Wei, G. (2001b) "Apparatus and methods for prediction of scour related information in soils", United States Patent, Patent No. US 6,260,409.

Briaud,J., Ting, F., Chen, H., Cao, Y., Han, S., Kwak, K. (2001b) "Erosion function apparatus for scour rate prediction", Journal of Geotechnical and Geoenviromental Engineering, ASCE, Vol. 127(2): 105-113.

Briaud, J., Chen, H., Li, Y., Nurtjahyo, P., Wang, J. (2003) "Complex pier scour and contraction scour in cohesive soils", National Cooperative Highway Research Program Report. January.

China Information Almanac (2009). Edit by National Development and Reform Council: Beijing.

Chow, W. (1959). "Open-channel hydraulics", McGraw-Hill Book Company: New York.

Christian, (2004). "Geotechnical Engineering Reliability: How Well Do We Know What We Are Doing", Journal of Geotechnical and Geoenviromental Engineering, ASCE, Vol 130 (10): 985-1003.

DeGroot, M., Schervish, M.,(2002). "Probability and Statistics", Addison Wesley: New Jersey.

Danish Hydraulic Institute (DHI) (2006). “Mike 21 Flow Model.” Mud Transport Module, Scientific Background. DHI, Denmark.

Department of Water Resource (DWR) (2009). “California Water Plan Update 2009”, Department of Water Resources, California, 2009.

Department of Water Resource (DWR) (2009). “DWR Levee Repair Program: Geotechnical subsurface investigations”, Department of Water Resources, California, 2009. http://www.water.ca.gov/levees/

Duncan, J., Houston, W. (1983) “Estimating failure properties for California levees” Journal of Geotechnical Engineering, Vol. 109(2): 260-268.

East Bay Municipal Utilities District (EBMUD) (2009). “HEC-RAS model for Mokelumne River”. EBMUD.

East Bay Municipal Utilities District (EBMUD) (2009). “Flooded bottomlands map in Lower Mokelumne River”. EBMUD.

Goldman, D. (1997). “Estimating expected annual damage for levee retrofits”, J. Water Resour. Plann. Manage., 123(2): 89-94.

Epsein, B. (1948). “Some application of the Mellin transform in statistics”, The Annuals of Mathematical Statistics, Vol. 19 (3): 370-379.

Fishman, G. (1996) “Monte Carlo: Concepts, Algorithms and Applications”. Springer series in operations

research. Springer-Verlag, New York, 1996.

Hanson, G. and Temple, D. (2002) “Performance of bare-earth and vegetated steep channels under long-duration flows”. Trans. ASAE 45(3): 695-701.

Harr, M. (1989) “Probabilistic estimates for multivariate analyses”, Applied Mathematical Modeling Vol.13: pp. 313-318.

Hanson, G., Simon, A. (2001) “Erodibility of cohesive streambeds in the Loess area of the Midwestern USA”. Hydrological Process, Vol.15 (1): 23-28.

Johnson, P,. (1996). “Uncertainty of hydraulic parameters”, J. of Hydraulic Engineering, Vol. 122, No.2, February 1996, pp. 112-114.

Kapur, K., Lamberson, L. (1977). “Reliability in engineering design”, Wiley: New York.

Kelly, Robert L. (1989). “Battling the Inland Sea: Floods, Public Policy, and the Sacramento Valley”, University of California Press.

Kuo, J., et al. (2007). “Risk analysis for dam overtopping – Feitsui Reservoir as a case study”, J. of Hydraulic Engineering, Vol. 133(8): 955-963.

Lee, H., and Mays, L. (1986). “Hydraulic uncertainties in flood levee capacity”, J. of Hydraulic Engineering, Vol. 112(10): 928-934.

Lund, J. R. (2009). “Probabilistic Design and Optimization”, Class ECI249, Department of Civil and Environmental Engineering, University of California, Davis.

National Research Council (2000). “Risk Analysis and Uncertainty in Flood Damage Reduction Studies”, National Academy Press, Washington. http://www.nap.edu/books/0309071364/html

Plate, E., Ihringer, J.(1986). “Failure probability of flood levees on a tidal river”, Stochastic and risk analysis in hydraulic engineering, B.C.Yen, ed., Water Resources Publications, Littleton, Colorado.

Plate, J.E., Duckstein, L. (1988). “Reliability-based desgn concepts in hydraulic engineering”, Water Resources Bulletin, Vol. 24(2): 235-245.

Robinson-Bryon, Inc (2006). “Floodplain resources characterization report: Consumnes & Mokelumne river floodplain integrated resources management plan”, Robinson-Bryan, Inc.

Seed, R. et al. (2006). “Investigation of the performance of the New Orleans flood protection systems in Hurricane Katrina on August 29, 2005”, NSF Independent Levee Investigation Team, Draft Final Report, Report No.UCB/CCRM-06/01, May 22. www.ce.berkeley.edu/~new_orleans/report.

Stedinger JR., Vogel RM., Foufoula-Georgiou E. (1993). “Frequency analysis of extreme events” In Handbook of Hydrology, Maidment D. (ed) from 18-1 to 19-65. McGraw Hill: New York.

Tung, Y.K., and Mays, L.W. (1980). “Risk analysis for hydraulic structure”, Journal of Hydraulics Engineering, ASCE, 106(5):893-913.

Tung, Y.K., Mays. L.W. (1981a). “Risk models for levee design”, Water Resource Research, AGU(4):833-842.

Tung, Y. Y., Mays, L. W. (1981b). “Optimal risk-based design of flood levee systems”, Water Resource Research, AGU, 17(4): 843-852.

Tung, Y., Yen, B. (1993). “Some recent progress in uncertainty analysis for hydraulic design”, Reliability and uncertainty analyses in hydraulic design, ASCE, New York.

Tung, Y.K., Yen, B.C. (2005). “Hydrosystems Engineering Uncertainty Analysis”, McGraw-Hill: New York.

Tung, Y.K., Yen, B.C., Melching, C.S. (2006). “Hydrosystems Engineering Reliability Assessment and Risk Analysis”, McGraw-Hill: New York.

U.S. Army Corps of Engineers (1981). "USACE Water control manual for Camanche dam and reservoir”, September, 1981, U.S. Army Corps of Engineers, Washington, D.C.

U.S. Army Corps of Engineers (1991). "Benefit determination involving existing levees”, Policy Guidance Letter No.26. U.S. Army Corps of Engineers, Washington, D.C.

U.S. Army Corps of Engineers (1996). “Risk-Based Analysis for Flood Damage Reduction Studies”, Manual No. 1110-2-1619, U.S. Army Corps of Engineers, Washington, D.C.

U.S. Army Corps of Engineers (1993). “Hydrologic Frequency Analysis”, EM 1110-2-1415, U.S. Army Corps of Engineers, Washington, D.C.

U.S. Army Corps of Engineers (1997). “Hydrologic Engineering Requirements for Reservoirs”, EM 1110-2-1420, U.S. Army Corps of Engineers, Washington, D.C.

U.S. Army Corps of Engineers (2006). “Risk-based analysis for flood damage reduction studies”, ER 11052–2-101, U.S. Army Corps of Engineers, Washington, D.C.

U.S. Army Corps of Engineers (2007). “RES-SIM Reservoir System Simulation: User's Manual”, U.S. Army Corps of Engineers, Washington, D.C.

United State Geological Service (USGS), (2009). “Quad Map in Central Valley, California”. http://www.usgs.org.

Van M., Vrjling, H., Van Gelder, P. (2007). “Probabilistic design and risk based approach in civil engineering”, 1st International Conference on Modern Design, Construction and Maintenance of Structures, 10-11 December 2007, Hanoi, Vietnam

Vootman, H.G., (2002). “Risk-based design of large-scale flood defense systems”, Ph.d - thesis, Delft University of Technology, 2002.

Vrijling, J.K., (2001). “Probabilistic design of water defense systems in The Netherlands”, Reliability Enginering and System Safety 74(2001): 337-344

Vrouwenvelder, A. (1987) “Probability design of flood defenses”, Report No.B-87-404. IBBC-TNO (Institute for Building Materials and Structures of the Netherland Organization for Applied Scientific Research), The Nethlands.

Wolfe, T.(2008). “Reliability of levee systems”, Reliability-based design in geotechnical engineering, Edited by K. Phoon, Taylor & Francis: New Jersey.

Wood, E.F.(1977). “An analysis of flood levee reliability”, Water Resource Research, 13(3): 665-671.

Yuceman et al., Tang, W. (1976) “Probability-based short term design of soil slopes”. *Canadian Geotechnical Journal* 13: 201–215.

Zhao, B., Mays, L. (1996). "Uncertainty and Risks Analysis for FEMA Alluvial Fan Method," Journal of Hydraulic Engineering, ASCE, Vol.122(6): 325-332.

Explanation of Variables and simplified terms

ASTM – American Society for Testing and Materials

CI – confidence interval

CDF – cumulative density function

DWR – Department of Water Resources

EBMUD- East Bay Municipal Utilities District

EFA – Erosion Function Apparatus

FOVE – First-Order Variance Estimation

MC – Monte Carlo Simulation

MFOSM – Mean-Value First-Order Second-Moment Method

Msl – mean sea level

PDF – probability of density function

PE – Point Estimate

psf – pound per square feet

RPE – Rosenblueth’s PE

TAF – Thousand ac-ft

USACE – U.S. Army of Corps of Engineer

USGS – United States of Geological Services

WSE – Water Surface Elevation

CHAPTER 6 SUMMARY AND CONCLUSIONS

Golden Ox is to memorize Da Yu and scare away floods. With it, Beijing City is long life.

- From Back of Bronze Ox, Summer Palace, Beijing, China

This dissertation examines levee failure analysis with reservoir re-operation in Lower Mokelumne River, North California. Based on theoretical analysis and simple case studies, the following conclusions are drawn.

1) Inflow analysis

In inflow analysis, this study presents a procedure of applying the bivariate normal distribution model with normal marginals to analyse multivariate flood events. The model is used to develop joint distributions of combinations of flood characteristics, namely flood peaks and volumes, and then flood volumes and durations. Based on this model, if the marginal distributions of two random variables can be represented by the lognormal distribution, one can readily obtain the joint probability distributions, the conditional distributions and the associated return periods of these variables. The parameters of the model can be estimated from the sample data on the basis of the single variable normal distribution.

The method is tested using flood observations from the Mokelumne River basin in Northern California. A good agreement is observed between the theoretical and observed distributions. The proposed method provides additional information unavailable from single variable flood frequency analysis, such as the joint return periods of the combinations of variables of interest (flood peak and volume, or flood volume and duration), and the conditional return periods of these variables. These joint results are more useful for probability studies of reservoir operations than single variable frequency estimates. For example, given a flood-event return period, it is possible to obtain various occurrence combinations of flood peaks and volumes, and vice versa. These various scenarios can be useful in the analysis and assessment of the risk associated with several hydrologic problems, such as spillway design and flood control in reservoir operation.

2) Design flood hydrograph

To develop design flood hydrographs for reservoir reoperation, three steps are presented: 1) Flood hydrograph separation and modification: 11 flood hydrographs were selected, separated and converted to dimensionless ones; 2) Hydrograph form fitting and selection: Beta, Gamma. Lognormal and Weibull distributions were selected and compared to fit as standardized hydrograph shapes based on goodness of fit criteria including RMSE and coefficients of determination. And 3) Development of design flood hydrographs: The design shape variables were determined from frequency analysis and finally, the design flood hydrographs including 10-, 20-, 50-, 100- and 200-year return periods were derived from the combinations of hydrograph shapes, flood volumes and durations. The Gamma PDF shape was the most suitable to fit Mokelumne River's flood hydrographs.

Some shortcomings in this study include the independence assumption, neglect of multi-peak hydrographs, and underestimation of flood peak values. The two shape parameters (mean and standard deviation) are considered as independent. Generally, a correlation may be between the parameters. Also, flood volume and duration are likely to be correlated as well. In such cases, bivariate joint probability distributions should be applied to represent the joint statistical properties of two correlated variables. All multi-peak hydrographs were removed from analysis as unsuitable for this analysis. In reality, multi-peaks floods are common in California. To solve this issue, base flow values may be increased. Finally, since the purpose of this study is to construct flood hydrographs for reservoir reoperation, the more essential parameters of a flood hydrograph are usually flood volume and duration. Flood peak values are somewhat ignored in the process of converting of original hydrographs to dimensionless hydrographs making them less suitable for estimating unimpaired hydrographs. More work is needed to adjust flood peak values on the designed flood hydrographs.

3) Regulated flow frequency analysis

To estimate regulated flow frequency for changing flood storage capacity, three main steps are presented including unregulated flow frequency analysis, unregulated/regulated flow transformation, and regulated flow calculation. The main contributions include modification of unregulated flow calculations to accommodate reservoir flood control and fitting unregulated and regulated flow to a probability distribution. The results will be used for levee failure analysis.

Some limitations are: 1) Extrapolation is a big issue for flood frequency analysis. To estimate low frequent events, one must extrapolate due to short of records. Viessman et al (1977) note that "as a general rule, frequency analysis should be avoided…in estimating frequencies of expected hydrologic events greater than twice the record length." In this study, only 86 years records are available. The generated regulated flow distribution will be used to determine the extreme event such as 100 year and 200 year. 2) This study assumes stationary. Historical records only provide the past records. With the climate change, the hydrology would change. There are many known causes for non-stationary ranging from the dynamics of the earth's motion to human cause changes in land use [Klemes, 1986]. This study assumes stationary. And 3) Simplicity of Inflow/outflow relationship curves: Scores of other factors exist in regulated versus unregulated flow relationship such as initial water surface elevation at the beginning of flood, outlet capacity, operation uncertainty. However, this study only sets flood storage capacity as a factor. This greatly simplifies the whole process.

Future study suggestions for regulated/unregulated flow analysis are: 1) For inflow frequency estimation, Partial Duration Series (PDS) should be considered to analyze flood pulse statistics since PDS has advantage over Annual Maximum Series (AMS). And 2) In regulated/unregulated flow relations, more factors including initial storage condition, downstream capacity changes and operation uncertainty could be considered and quantified.

4) Levee failure analysis

Finally to estimate overall levee reliability incorporating multiple uncertainties including hydrologic, hydraulic and geotechnical factors, two main contributions include overall reliability estimation and load-resistance interference risk analysis. Overtopping and erosion failure probability analysis are usually performed separately in water resources engineering and geotechnical engineering, respectively. This study estimates a comprehensive reliability combining these two failure mechanisms. Also load-resistance interference risk analyses were introduced to perform overtopping between flood magnitude and levee capacity and erosion failure between velocity and soil strength. Both analyses are performed by Monte Carlo simulation.

Some shortcomings include less consideration of other failure mechanisms, assumption of independence and normality assumptions. The levee failure is complicated, involving many potential failure mechanisms. This approach only considers overtopping and erosion failures. Other failure mechanisms such as under seepage and through seepage are also common. More research is needed for seepage risk analysis. This study also assumes failure modes are independent and uncorrelated. This is not necessarily true, as some of conditions increasing the probability of failure for one mode are likely to increase the probability of failure by another. In reality, overtopping and erosion often occur under similar conditions. More research focus on correlation between failures modes is needed. Finally, most of random variables are assumed to have log-normal distributions including hydrologic variables, flow magnitude, flood duration, hydraulic variables, channel cross section area, levee cross section areas, Manning's n values, and geotechnical variables, soil shear stress, erodibility coefficient. In fact, the variables probability distributions vary from site to site.

Floods have accompanied with human being in past thousands years and it will be with us for ever. However, as long as we know better flood inflow, i.e., drive force, we will have better resistant force, i.e., flood control system to deal with floods. Reservoirs and levees are essential components in flood control system and they should be utilized better.