

## Flood Protection Lervee

by
Eugenio Castaño

Technical Reports on
Natural Resource Systems

The University of Arizona
Tucson, Arizona 85721

# MODEL UNCERTAINTY IN THE DESIGN OF A FLOOD PROTECTION LEVEE 

by
Eugenio Castano

Reports on Natural Resource Systems Report No. 28

June 1976

Collaborative effort between the following Departments: Hydrology and Water Resources
Systems and Industrial Engineering

## PREFACE

This report constitutes the Master of Science thesis of the same title completed by the author in May, 1976, and accepted by the Department of Systems and Industrial Engineering.

The work upon which this thesis is based was supported in part by funds provided by the National Science Foundation grants: ENG-74-20462 entitled "Sensitivity of Decisions in Resource Engineering to Assumptions of Multivariate Models," and GF-38183 entitled, "Cooperative Research on Decision Making Under Uncertainty in Hydrologic and other Resource Systems."

Lucien Duckstein Jointly with the Departments of Systems and Industrial Engineering and Hydrology and Water Resources

Donald R. Davis Department of Hydrology and Water Resources

## ACKNOWLEDGMENTS

The author is deeply grateful to many people for their cooperation during the development of this work. Special thanks must be given to Professor Frenec Szidarovszky for his permanent encouragement and his many valuable suggestions; to Professors Lucien Duckstein and Donald R. Davis for their advice and positive criticism.

Much appreciation is extended to Mr. William Metler for sharing his experience with the confluence levee problem and for his program "LEVEE" that was the basis for the development of the program "LEV3."

The work upon which this thesis is based was supported in part by funds provided by the grants No. ENG 74-20462 entitled "Sensitivity of Decisions in Resource Engineering to Assumptions of Multivariate Models" as well as by the grant No. GF 381833 entitled "Cooperative Research on Decision Making Under Uncertainty in Hydrologic and other Resource Systems." Both grants are from the National Science Foundation.

TABLE OF CONTENTS
Page
LIST OF ILLUSTRATIONS ..... viii
LIST OF TABLES ..... X
ABSTRACT ..... xi
CHAPTER

1. INTRODUCTION ..... 1
1.1 Summary ..... 1
2. 2 Floods and Flood Control ..... 3
3. 3 Uncertainties in Engineering Design ..... 6
1.3.1 Natural Uncertainty ..... 7
1.3.2 Parameter Uncertainty ..... 8
1.3.3 Model Uncertainty ..... 9
1.3.4 Economic Uncertainty ..... 11
1.4 Scope ..... 13
4. PROBLEM DEFINITION AND DESCRIPTION OFTHE MODEL15
2.1 Hydraulic and Structural Considerations ..... 16
5. 2 Hydrologic Considerations ..... 18
6. 3 Economic Considerations ..... 20
2.4 A Model for Optimum Development ..... 21
2.4.1 Construction Cost ..... 22
2.4.2 Flood Losses ..... 23
2.4.3 Making the Goal Function Suitable for Numerical Evaluation ..... 24
2.4.4 Use of the Conditional pdf's in the Evaluation of the Expected Damages ..... 26
2.4.5 Final Form of the Goal Function ..... 28
2.5 The Complete Model ..... 28
7. SELECTION OF THE HYDROLOGIC SUBMODEL ..... 30
3.1 Generalities ..... 30
3.1.1 Model Selection from Sample Information ..... 31
3.2 Distributions to be Used ..... 33

TABLE OF CONTENTS--Continued

## Page

3.2.1 Lognormal Distribution ..... 33
3.2.2 Exponential Distribution ..... 34
3.2.3 Gamma Distribution ..... 34
3.2.4 Beta Distribution ..... 35
3.2.5 Normal Distribution ..... 35
3.2.6 Truncated Normal Distribution ..... 36
3.2.7 Bivariate Gamma Distribution ..... 37
3.2.8 Truncated Bivariate Normal Distribution ..... 38
3.2.9 Truncated Bivariate Lognormal
Distribution ..... 39
3.3 Criteria for Model Selection ..... 40
3.3.1 Goodness of Fit ..... 41
3.3.2 Most Likely Value of the Goodness of Fit Statistic ..... 43
3.3.3 A Composite Model Based on Sample Likelihoods ..... 43
3.3.4 Use of Sample Likelihoods for Model Selection ..... 47
4. NU̇MERICAL RESULTS ..... 48
4.1 Implementation of the Hydraulic and Economic Submodels ..... 49
4.2 Implementation of the Hydrologic Submodel ..... 50
4.2.1 The Sample ..... 50
4.2.2 Candidate pdf's ..... 54
4.2.3 Goodness of Fit Tests ..... 55
4.2.4 Selection of the "Best" pdf ..... 61
4.2.5 Summary and Remarks on Model Selection ..... 64
4.3 Computational Considerations ..... 67
4.3.1 Integration Region for Expected Losses ..... 68
4.3.2 Convergence Problems in
Numerical Integration ..... 69
4.4 Determination of the Optimum Levee Profile ..... 70
4.4.1 pdf's to be Used ..... 72
4.4.2 Preliminary Results ..... 72
4.4.3 Optimum Design Alternatives ..... 78
4.4.4 Discussion of Results ..... 87

## TABLE OF CONTENTS--Continued

Page
5. DISCUSSION AND CONCLUSIONS ..... 93
5.1 Discussion ..... 93
5.1.1 Hydraulics ..... 93
5.1.2 Economics ..... 93
5.1.3 Methodology ..... 94
5.1.4 pdf Fittings ..... 95
5.1.5 Model Selection ..... 96
5.2 Conclusions ..... 96
APPENDIX A. DATA USED IN THE CASE STUDY ..... 100
APPENDIX B, PROGRAM LISTING ..... 106
NOTATION ..... 136
REFERENCES ..... 138

## LIST OF ILLUSTRATIONS

Figure Page
2.1 The Complete Model ..... 29
4.1 Typical Cost Functions ..... 51
4.2 Typical Damage Functions ..... 52
4.3 The Sample ..... 53
4.4 Relative Ranking of Candidate pdf's ..... 65
4.5 Regions in the Space of ( $x, y$ ) for a Design Alternative ( $g_{m}, g_{t}$ ) ..... 68
4.6 Basic and Actual Integration Regions ..... 71
4.7 Marginals $f_{x}(x \mid E l)$ ..... 73
4.8 Marginals $f_{y}(y \mid E l)$ ..... 74
4.9 Marginals $f_{x}(x \mid E 2)$ ..... 75
4.10 Marginals $f_{y}(y \mid E 2)$ ..... 76
4.11 TYC Surface for Models Type I ..... 79
4.12 TYC Surface for Models Type II ..... 80
4.13 TYC Surface for Models Type III ..... 81
4.14 TYC Surface for Models Type IV ..... 82
4,15 TYC Surface for Composite Model ..... 83
4.16 Variation of TYC for $g_{t}$ Constant at Its Optimum Value for Each Model ..... 84
4.17 Variation of TYC for $g_{m}$ Constant at Its Optimum Value for Each Model . . . . . . . . . ..... 85
4,18 Optimum Design Alternatives for All Models ..... 86

## LIST OF ILLUSTRATIONS--Continued

| Figur |  | Page |
| :---: | :---: | :---: |
| 4.19 | Variation of Damages Due to E 2 with Change in $g_{t}$ for 4 Models $f_{y}(y \mid E 2)$. . . . . | 88 |
| 4.20 | Construction Cost Surface (Million Forints) | 90 |
| 4.21 | Existing Levee and Optimum Levee for the Composite Model | 92 |

## LIST OF TABLES

TablePage4.1 Pdf's of Flood Magnitudes Given Event El . . . . ..... 56
4.2 Pdf's of Flood Magnitudes Given Event E2 . . . . ..... 57
4.3 Results of Univariate Chi-square Tests . . . , ..... 58
4.4 Results of Bivariate Chi-square Tests for Candidate pdf's to $\mathrm{f}(\mathrm{x}, \mathrm{y} \mid \mathrm{El})$. . . . . . . . . ..... 59
4.5 Results of Bivariate Chi-square Tests for Candidate pdf's to $f(x, y \mid E 2)$ ..... 60
4.6 Probability of the Chi-square Statistic for Candidate pdf's to $f(x, y \mid E l)$ ..... 62
4.7 Probability of the Chi-square Statistic for Candidate pdf's to $f(x, y \mid E 2)$. . . . . . . . ..... 62
4.8 Weights for Candidate pdf's Based on Sample Likelihoods . . . . . . . . . . . . ..... 63
4.9 Optimum Design Alternatives ..... 78

## ABSTRACT

The model choice problem in Hydrology is illustrated by means of the optimum levee design for flat rivers along a confluence reach. Special attention is given to the selection of a probability distribution for the joint flood stages.

The optimality criterion used is the minimization of construction plus expected flood damage costs. The main assumption in the mathematical model is that the levee profile is uniquely determined as a function of the levee heights at the extremes of the reach; thus the problem is reduced to the determination of the optimum pair of extreme levee heights.

The selection of a probability distribution of flood stages, from a set of distributions estimated from the partial duration series, is performed using either one of two selection procedures: likelihood of the Chi-square statistic and sample likelihoods. A composite distribution, taking into account the model uncertainty, is also derived.

The methodology presented is applied to the remodeling of the levee on the west bank of the Zagyva River, in Hungary. A sensitivity analysis is performed, using the best ranking distributions according to the two model choice xi
procedures. The composite distribution appears to offer a reasonable choice.

## CHAPTER 1

## INTRODUCTION

The problem of determining the optimum levee design for flat rivers along a confluence reach is presented here, giving special attention to the selection of the hydrologic model.

### 1.1 Summary

After an introductory discussion of flood protection methods and the uncertainties in engineering design, the emphasis of the study is focused on the design of levees in flat rivers and the selection of a bivariate pdf to synthesize the hydrologic aspects of the system.

Chapter 2 is devoted to the analysis of the hydraulic, hydrologic, and economic aspects of flood protection levees in flat rivers, under random backwater effects. This analysis leads to the statement of the basic assumptions, and to the definition of the three major components of the mathematical model for the determination of the optimum levee profile. The hydraulic component will give the water surface corresponding to any valid pair of water stages at both ends of the levee reach; the hydrologic component gives the probability of any valid pair of water stages; finally, the economic component gives the
total yearly cost (TYC) for a levee design alternative, as the sum of the construction cost and the expected flood damages for that alternative. The interaction between the major components is defined by the goal function. The exploration of the computational aspects of the goal function leads to the mathematical expressions that will finally be used in the implementation of the model.

Chapter 3 deals with the selection of a probability distribution for the hydrologic submodel. It begins with an introduction to the model choice problem and its importance, especially when extreme events are considered. The pdf's to be used in Chapter 4 are presented next, giving their functional form and the equations for parameter estimation. In the last part of this chapter, two model selection criteria are presented: likelihood of the Chi-square statistic and sample likelihoods. The use of goodness of fit tests for a preliminary reduction of the candidate set is also considered. This chapter is closed with the derivation of a composite model consisting in the sum of several probability distributions, weighted by their sample likelihood.

In Chapter 4, the mathematical model defined in Chapter 2, and the model selection procedures discussed in Chapter 3 are used to determine the optimal levee profile for the remodeling of the existing levee on the west bank of the Zagyva River in Hungary.

The transformation of raw data into functional parameters for the hydraulic and economic submodels is briefly explained. The implementation of the hydrologic submodel, by the application of the model choice procedures, is explained in detail. Next, the computational aspects of the evaluation of expected losses are studied, and a further simplification is introduced to the model. In the final sections of Chapter 4, a sensitivity analysis is performed, using the best ranking pdf's from the model selection part. The optimum decisions and the total yearly cost surfaces are presented for five different pdf's. The results lead to the conclusion that, since the model uncertainty can not be completely resolved, the composite model, that takes into account the model uncertainty, is perhaps the most appropriate, especially since an expected value approach has been adopted in the definition of the goal function.

## 1. 2 Floods and Flood Control

Rivers, a very valuable natural resource for humanity, also produce each year tremendous material damage and loss of human lives due to floods. Every nation faces, in relation to its rivers, two simultaneous goals: first, to derive the mbst benefit from its rivers and, second, to minimize the damages and losses caused by floods. The second goal is what is called flood protection. Yevjevich (1974) summarizes the following methods that can be used,
alone or in combination, to protect a community or any valuable land against flood damage.

1. Reduction of the flood peaks by the use of reservoirs. The idea is to keep the river flow within safe limits by storing the excess water in a reservoir upstream of the protected area.
2. Confinement of the flow within predetermined channels. It consists in building a barrier parallel to the river channel and between the river and the protected area. Two types of barriers are in use: levees, which are earth dikes built with materials available at the site, and flood walls, which are masonry or reinforced concrete structures built parallel to the river. Levees are widely used because of their low cost. However, since they require a large base width, due to their flat side slopes, the cost of the land occupied by the levee may become too high in populated areas. In such case, the use of flood walls could be a more economical alternative. To have an idea of the size of these structures, we cite Van Ornum (1914) who presents a typical levee cross-section in the Mississippi River. Such levee is 6.10 m high and 36.58 m wide at the base, with a volume of 125.28 $\mathrm{m}^{3} / \mathrm{m}$. The largest embankment on the Tisza River (Hungary) has a volume of $10.273 \mathrm{~m}^{3} / \mathrm{m}$, the same
author reports. The levee reach along the Zagyva River (Hungary) has a length of 60.4 km (Bogardi, Duckstein, and Szidarovszky, 1975).
3. Reduction of the flood stage by increased speed. The flow speed can be increased by cleaning and straightening the channel.
4. Diversion of flood waters. This method, used by the Egyptians in the Nile River, consists in flooding a large, shallow valley to reduce the flood peak downstream.
5. Reduction of the runoff by land management. Reforestation of the basin provides some storage capable of regulating small runoffs. However, it is not effective in case of heavy precipitation.
6. Zoning. It consists in legally restricting the use of the flood plain.
7. Evacuation. It proves to be an effective solution when a warning system is available and the value of the land does not justify a more expensive form of protection. It consists in the evacuation of persons, livestock, and commodities from the threatened area when there is a certain risk of flood.
8. Null alternatives. That is, do not take any preventive measure.

Generally, a combination of several of the above methods is used. In the present work, only the design of levees will be considered.

## 1. 3 Uncertainties in Engineering Design

Any engineering design is the result of a best compromise among physical, economical, social, and legal factors. In the design of a flood protection system, for example, the input may include: physical constraints (terrain, strength of materials, flood regime), social requirements (recreational, ecological, risk to human life), economical constraints (costs, damages, budgetary restrictions, cash flow, effect on navigation), legal regulations (federal, state, and city codes), etc. The output consists in the optimum (under the constraints imposed) combination of flood protection methods and the engineering specifications for each component of the system.

In the input for the design process, the value of many of the parameters is not certain. Some of them will change along the lifespan of the project; others will remain constant but are only known approximately. Several authors, such as Wood, Rodriguez-Iturbe, and Schaacke (1974) and Benjamin and Cornell (1970), have classified uncertainty into: natural, parameter, and model uncertainty. Other authors go further and also consider: economic, technological, and strategical uncertainty (Bogardi, 1975;

Duckstein et al., 1975; Duckstein and Simpson, 1975). A brief review of these uncertainties follows.

### 1.3.1 Natural Uncertainty

It is associated with the stochastic nature of complex phenomena. It is present, not only in natural processes, but also in economic, social, and, in general, all kinds of phenomena that are too complex to be predicted deterministically (Bogardi, 1975).

To cope with natural uncertainty, it is necessary to treat the uncertain parameters as random variables, and a probability distribution (pdf) must be assigned to the uncertain design parameters. At this point, the following question comes to mind: Should all the uncertain variables be randomized? Benjamin and Cornell (1970, p. l) say: "If the degree of variability is small and if the consequences of any variability are not significant, the engineer may choose to ignore it (the uncertainty) by simply assuming that the variable will be equal to the best available estimate." Only those variables with a wide variability and with significant consequences associated with such variations should be randomized. In flood protection, for example, the maximum flood peak is a highly uncertain quantity. Since the degree of protection is very much based in such value, improper estimation of the maximum flood peak leads to a considerable extra cost in terms of damages
or over-design. To take into account the natural uncertainty in the flood peak, the optimum levee height (H*) can be chosen such that minimizes the expected net cost of the project.

$$
\begin{equation*}
Z\left(H^{*}\right)=\min _{H}\left[\int_{h} L(H, h) f(h) d h+K(H)\right], \tag{1.1}
\end{equation*}
$$

where
Z (H) $=$ net cost,
$L(H, h)=$ losses for $a$ flood of height $h$ with levee height $H$,
$K(H)=$ construction cost of a levee of height $H$, and $f(h) d h=p d f$ of $h$, the flood level.

### 1.3.2 Parameter Uncertainty

To account for the natural uncertainty, the designer adopts a probabilistic model to synthesize the behavior of the uncertain variable. Then he has to estimate the parameters of the probabilistic model from a finite sample. This leads to uncertainty in the distribution parameters, because of the randomness of the sampling process, in which different samples will lead to different parameter estimates for the same phenomenon. This new source of uncertainty is called parameter or sample uncertainty. In Bayes Decision Theory ( $B D T$ ), both the natural and the sample uncertainty are taken into account by using the expected value of $f(\underline{h})$ given the sample, where $f(\underline{h})$ is the pdf of the uncertain variables. In the previous
example, the optimum levee height taking into account the natural and the sample uncertainty will be $H^{*}$ such that:

$$
\begin{equation*}
Z(H *)=\min _{H}\left[\int_{h} L(H, h) \tilde{f}(h) d h+K(H)\right], \tag{1.2}
\end{equation*}
$$

where
$\tilde{f}(h) d h=\int f(h \mid \underline{a}) g(\underline{a}) d \underline{a}$,
$\underline{a}=$ parameter vector of $f(h)$,
$g(\underline{a})=$ prior pdf of $\underline{\text { a }}$.
All other parameters and variables were defined earlier in this section.

### 1.3.3 Model Uncertainty

A theoretical model is "a quantitative description of a physical phenomenon" (Smallwood, 1968, p. 333). Most models are unable to exactly mimic the phenomenon under study. Such mismatch between real and modeled behavior constitutes in itself a source of uncertainty.

Some phenomena are simple enough to be represented by a deterministic model, i.e., a more or less complicated functional relationship between inputs and outputs. However, most natural and social phenomena are too complex to be synthesized in a deterministic manner and a probabilistic approach must be used instead. Essentially, a probabilistic model gives the probability of occurrence for each possible outcome of the real phenomenon.

The choice of a probabilistic model can be based on theoretical considerations or on empirical observation. For
example, i.t is known that a process resulting from the sum of many independent causes is related to the normal distribution, and that a lognormal model is more appropriate when a multiplicative effect instead of an additive one is present. When the "internal mechanism" of the process is not known, model choice must be based on empirical observation, i.e., on the historical records. Very often the sample size available does not permit a precise discrimination among several pdf's that could explain a phenomenon. A typical case is that of hydrological extreme events. Here, the sample is abundant in realizations close to the center of the pdf, but very little is known about the tails of the distribution--the extreme events. In such case, the modeller may have several pdf's that fit the data equally well (or bad). The lack of sample information to choose the "true" model is an additional source of uncertainty associated with probabilistic models.
"In recent years, considerable progress has been made on the development of statistical procedures for discrimination among alternative models" (Wood et al., 1974, p. 231). However, most of the research has been oriented toward the area of econometrics where the model choice problem is most often in terms of which variables are to be included in a regression model. Nevertheless, several authors, Gaver and Geisel (1974); Atkinson (1970); Dumonceaux, Antle, and Haas (1972), among others, have
dealt with the model choice among different families of distributions. The classical or non-Bayesian procedures are usually oriented toward statistical testing of hypotheses. The drawbacks of the classical procedures have been pointed out by Wood et al. (1974). We mention: (1) the need to compare the models by pairs, (2) the low power of the tests for small sample sizes, and (3) classical procedures do not allow the economic assessment of the model uncertainty. The following statement by Gaver and Geisel (1974, p. 65) adds one more point to the list:

If a model has a positive probability (of being the true model), it contributes to our knowledge of future observations and there is no reason to neglect this contribution. Procedures that select one model are thus seen as approximations undertaken for simplicity of view or ease of computation.

Bayesian procedures have been applied to model choice in econometrics by Gaver and Geisel (1974). Smallwood (1968) presents a Bayesian framework for simultaneous consideration of several alternate models. Wood et al. (1974) apply Bayesian procedures to model selection in a flood protection problem. Further details on model selection will be given later in the present thesis.

### 1.3.4 Economic Uncertainty

Economic uncertainty is associated with our lack of knowledge about the value of economic parameters such as costs, losses, and benefits. Several authors consider the economic uncertainty as a part of the model uncertainty. We
presume this is so because they consider the model as the whole mathematical entity used to represent the problem. However, in the case of flood control, for example, the model has at least three very definite components: the hydrologic component, the hydraulic component, and the economic component; not to mention other components such as social, political, ecological, etc. Each one of these components is subject to the three basic uncertainties: natural, sample, and model uncertainty. Strictly speaking, then, when we refer to economic uncertainty, we should imply the natural, sample, or model uncertainty in the economic component of a larger model. We did not explore the realms of economics and econometrics to see what has been done to cope with uncertainty in economic models. In water resources applications, it is apparent that the three basic uncertainties have not been fully accounted for in the economic component of the models.

ECUP, Economic Uncertainty Programming (Szidarovszky et al., 1976, in press) is a procedure that accounts for the economic uncertainity (more precisely, the natural uncertainty in the economics). ECIJP is based on the fact that the loss and cost functions can be estimated only at discrete points in the decision space (the height of a levee for example). It assumes that these discrete points in the loss and cost functions are distributed as dependent joint normal variates, where the means are the estimated
points and the covariance matrix reflects the uncertainty in such estimates. Monte Carlo simulation is used to compute the empirical pdf of the optimum design. The use of the pdf of the optimum design, instead of a unique optimum, enables the decision maker to consider the uncertainty involved in the selection of the final design value. ECUP also permits the assessment of the economic uncertainty in monetary units, i.e., the expected opportunity loss and the value of perfect information may be easily computed. The hydrologic component in the mathematical formulation to compute the optimum design for every set of simulated economic values can take into account the natural and sample uncertainty in the hydrologic variables by any of the methods mentioned here in the previous section.

This concludes a brief review of the uncertainties in engineering design. The present work will focus on uncertainty in the pdf that synthesizes the hydrologic aspects of the problem. The parameter and the economic uncertainties will not be considered; the natural uncertainty is implicitly considered by assuming randomness in the hydrologic aspects of the problem.

## $1.4 \quad$ Scope

In the present work, an attempt will be made to resolve the hydrologic model uncertainty for a realistic situation. Namely, the optimum levee profile under
stochastic flow regime is determined for the levee reach at the Zagyva River along 60.4 kilometers before its confluence with the Tisza River, in Hungary.

## CHAPTER 2

## PROBLEM DEFINITION AND DESCRIPTION OF THE MODEL

The problem under consideration is the determination of the optimum levee profile for flood protection along a confluence reach under random backwater conditions. The relevant characteristics of the problem are now considered.

The location of the levee near a confluence means that the water surface curves along the reach will depend on the flows in the tributary and in the main river, along with the hydraulic characteristics of the channel and the flood plain. The latter can be assumed to be known and constant; the former are essentially of a random nature and must be treated as such in the decision process. The optimization aspect of the problem implies the search for the alternative that maximizes net benefits or, equivalently, that minimizes construction costs plus losses. Since only the quantifiable aspects of the problem are considered here, this is an optimum from the economic viewpoint only. It will have to be modified prior to its implementation, in order to satisfy social and political requirements that have not been included in the model.

### 2.1 Hydraulic and Structural <br> Considerations

The fact that the levee is near a confluence complicates somewhat the computation of surface curves due to the backwater effect. In this case, the water profile not only depends on the flow in the tributary but also on the flow in the main river. Assuming steady flow in both rivers, the water profile is defined by the differential equation (Kuiper, 1965):

$$
\begin{equation*}
\frac{d h}{d x}+\frac{\bar{u} d \bar{u}}{g d x}=\frac{Q^{2}}{K^{2}(h, x)} \tag{2.1}
\end{equation*}
$$

where x is the horizontal distance from the mouth of the river, $h$ is the water stage at $x, \bar{u}$ is the average velocity through the cross section at $x, g$ is the acceleration of gravity, $Q$ is the discharge of the tributary, and $K(h, x)$ is the conveyance of the cross section at $x$.

The solution of this equation with initial values $\left(h_{m}, h_{t}\right)$ at the extremes of the reach uniquely determines the water surface curve corresponding to the pair ( $h_{m}, h_{t}$ ). Depending on the difference $h_{t}-h_{m}$, the water surface curve can be concave upward ( $M_{I}$ or "damping" curve) if ( $h_{t}-h_{m}$ ) is less than Dh; or concave downward (MII or "draw down" curve) if $\left(h_{t}-h_{m}\right)$ is greater than or equal to $D h$, where Dh is the altitude difference between the gauges where $h_{m}$ and $h_{t}$ are measured. The "draw down" curves are associated
with floods in the tributary, and the "damping" curves with floods in the main river.

The solution of Equation (2.1) is a boundary value problem. However, since the discharge Q can be expressed as a function of the water stage at the upper end of the curve $\left(h_{t}\right)$, it can be transformed into an initial value problem. When the geometry of the channel and the conveyance function, $K(h, x)$, are both particularly simple, the equation can be solved analytically by conventional methods such as power series (Szidarovszky, 1974), Piccard's iteration, or quasilinearization (Bellman and Kalaba, 1965). In a natural channel, however, that simplicity is almost never enountered, and numerical methods must be used. General methods available include: Runge Kutta methods, linear multistep, predictor corrector methods (Ralston, 1965), and difference method (also known as step method) (Szidarovszky, 1974). There are also specialized graphical methods such as the Ezra and the Escoffier methods (Henderson, 1966).

For a computer solution of the equation, graphical methods are ruled out for obvious reasons. The first three of the general methods above mentioned require the equation to be expressed in explicit form, and the right hand side of the (explicit) equation must be evaluated a large number of times. The step method, consisting in the numerical solution of the finite difference form of the differential equation, is computationally simpler than the other methods
and the points (x) where the function is evaluated can be restricted to predetermined cross sections. It appears then that the step method is the best suited for the solution of Equation (2.1).

Overtopping occurs when the flood level is higher at some point than the existing levee. It is the most important mode of failure and the only one to be considered here. Other modes of levee failure are: water saturation and loss of soil stability, boils and hydraulic soil failure, and wave action (Bogardi and Zoltán, 1968). In addition to flood stages, the alternate modes of failure depend on other flood parameters such as: flood exposure, duration of the flood wave, and wind. A more detailed model could take into account these and other aspects of the flood protection problem.

## 2. 2 Hydrologic Considerations

The hydraulic component of the model will produce the water profile corresponding to any pair of stages ( $h_{m}, h_{t}$ ). If the most adverse pair to occur during the life span of the project were known, the optimum levee profile could be easily determined. However, due to the random nature of the water stages, the most adverse pair cannot be predicted. It is generally uneconomical to build a levee capable of standing extremely large (and unlikely) floods. Instead, a smaller structure is constructed and a certain
risk of failure is thereby accepted. The degree of risk involved is a function of the designed capacity of the levee and the probability of floods that could exceed such capacity.

The ignorance about future floods is by no means total. Historical records provide valuable information about the flood regime and are the basis for the estimation of probability distributions (pdf's) of flood frequencies and flood discharges. The most common form of flow records are periodic stages (annual, daily, etc.) and partial duration series. Instead of periodic records of flow stages, the partial duration series only show stages above a certain base level b. Such records are, consequently, event based in nature and automatically condition the type of probabilistic model that can be obtained, as event-based models.

In the case of a confluence, it is natural to expect some degree of correlation between the discharges (and also the stages) of the two rivers. Consequently, simultaneous records are required in order to estimate the joint pdf of flood stages. Bivariate partial duration series consist of all pairs $\left(h_{m}, h_{t}\right)$ such that either or both stages exceed the basic level.s. The sample space of the bivariate partial duration series can be defined as follows:

$$
\begin{equation*}
s=\left\{\left(h_{m}, h_{t}\right): \cdot h_{m}>b_{m}, \cdot O R \cdot h_{t}>b_{t}\right\} \tag{2.2}
\end{equation*}
$$

and it can be divided into the mutually exclusive and collectively exhaustive events:

$$
\begin{align*}
& E 1=\left\{\left(h_{m}, h_{t}\right): h_{m}>b_{m}, h_{t}<b_{t}\right\},  \tag{2.3}\\
& E 2=\left\{\left(h_{m}, h_{t}\right): h_{m}>b_{m}, h_{t}>b_{t}\right\},  \tag{2.4}\\
& E 3=\left\{\left(h_{m}, h_{t}\right): h_{m}<b_{m}, h_{t}>b_{t}\right\}, \tag{2.5}
\end{align*}
$$

where $b_{m}$ and $b_{t}$ are the basic flow levels in the main river and the tributary, respectively. The conditional pdf's of joint flood stages: $f_{1}\left(h_{m}, h_{t} \mid E l\right), f_{2}\left(h_{m}, h_{t} \mid E 2\right)$, and $f_{3}\left(h_{m}, h_{t} \mid E 3\right)$, can be estimated from the subsamples corresponding to events E1, E2, and E3, respectively. Those pdf's completely synthesize the hydrology of the problem for the purposes of the present study.

### 2.3 Economic Considerations

The two essential components of any economic enterprise, namely, investment and benefits, appear here as construction (and maintenance) costs and reduction of flood damages. An economically optimum alternative is sought such that maximizes the total net benefit or, equivalently, minimizes total net cost.

It is assumed here that construction cost, as a function of the levee height, and flood damages, as a function of the water stage, are available for every cross section along the reach. Because of its dependence upon random water stages, the total damage for a given levee
alternative must also be treated as a random variable. Different levee alternatives will be compared in terms of their total expected cost, which is defined as follows (where the underline denotes vectors):

$$
\begin{equation*}
T E C=C(\underline{G})+\int_{D} L(\underline{G}, \underline{H}) f(\underline{H}) d \underline{H} \tag{2.6}
\end{equation*}
$$

where:
$\underline{G}$ is the profile of the levee,
$\underline{H}$ is the (random) water profile,
$\mathrm{f}(\mathrm{H})$ is the pdf of the water profiles,
$L(\underline{G}, \underline{H})$ is the damages caused by a flood $\underline{H}$ when the levee profile is $\underline{G}$,
$C(\underline{G})$ is the yearly construction cost of the levee, and
D is the domain of all possible values of $\underline{H}$.
Since the expected damages are expressed on a yearly basis, the construction cost, which is a "one time" expense, must also be reduced to a yearly basis. This is done by multiplying the initial cost by the discount factor associated with the life of the project and a chosen interest rate.

### 2.4 A Model for Optimum Development

As a conclusion from the economics of the problem, the goal function can be expressed as:

$$
\begin{equation*}
\min \left\{C(\underline{G})+\int L(\underline{G}, \underline{H}) £(\underline{H}) d \underline{H}\right\} . \tag{2.7}
\end{equation*}
$$

All the elements of the goal function have been defined in Equation (2.6).

The difficulty in obtaining an analytical expression for $\underline{H}$, the solution of Equation (2.1), dictates the discretization of the water and levee profiles. $\underline{H}$ and $\underline{G}$ will then represent the $n$-vectors $\left[g_{1}, g_{2}, \ldots, g_{n}\right]$ and $\left[h_{1}, h_{2}, \ldots\right.$, $h_{n}$ ], respectively, where $g_{i}$ or $h_{i}$ are the levee or water profile heights at the i-th cross section located at a distance $x_{i}$ from the confluence.

The goal function adequately synthesizes all the properties of the model. The decision variables $g_{1}, \ldots, g_{n}$ are the levee heights at the n cross sections; the hydraulic variables and relations are imbedded in the total loss function $L(\underline{G}, \underline{H})$; the hydrologic component is present in the pdf of water profiles; finally, the economics appear in the cost and loss functions and in the form of the goal function itself. The different components of the model will now be analyzed in detail and finally assembled together into a complete model.

### 2.4.1 Construction Cost

Construction cost functions are available for each cross section. The total building cost for a levee alternative with profile $G$ is given by:

$$
\begin{equation*}
C(\underline{G})=\sum_{i=1}^{n} c_{i}\left(g_{i}\right) \tag{2.8}
\end{equation*}
$$

where $c_{i}\left(g_{i}\right)$ is the construction cost function for cross section $i$.

### 2.4.2 Flood Losses

Flood damage functions are also available for each cross section. The function $l_{i}\left(g_{i}, h_{i}\right)$ gives the damages when the levee is overtopped at cross section i. Obviously,

$$
\begin{align*}
l_{i}\left(g_{i}, h_{i}\right) & \geq 0 \text { for } h_{i}>g_{i} \\
& =0 \text { for } h_{i} \leq g_{i} \tag{2.9}
\end{align*}
$$

It is important to stress the fact that the function $l_{i}$ not only refers to damages in the vicinity of the failure but to the total area reachable by the flood waters from the point of failure.

Bogardi et al. (1975) have defined the failure mechanism of a levee reach in flat rives as follows. Based on the assumption that the levee follows its corresponding water profile (in other words, the water surface curve with stages $g_{l}$ and $g_{n}$ at its extremes), the failure point is determined as the first overtopped cross section from upstream for a flood with a draw down curve; and the first one from downstream for a flood with a damping curve.

The first assumption is based in the fact that if there is a section of the levee that is below this "optimum" profile, the levee will fail, precisely at that section, for some pair of stages $\left(h_{m}, h_{t}\right)$ such that $h_{m}<g_{m}$ or $h_{t}<g_{t}$ or both. Hence the levee is not providing as much protection as it could. On the other hand, if there is a section that is higher than the "optimum" profile, it represents an
unnecessary expense because it does not increase the total reliability of the structure.

Based on the above assumptions, the total loss function adopts the form:

$$
\begin{equation*}
L(\underline{G}, \underline{H})=\sum_{k=1}^{n} I_{k}\left(g_{k}, h_{k}\right) t_{k}(\underline{G}, \underline{H}) \tag{2.10}
\end{equation*}
$$

where $l_{k}\left(g_{k}, h_{k}\right)$ is the flood damage function for cross section $k, t_{k}(\underline{G}, \underline{H})$ is an indicator function such that: $t_{k}(\underline{G}, \underline{H})=1$ if: $h_{t}-h_{m}<D h$, and:

$$
h_{1} \leq g_{1}, h_{2} \leq g_{2}, \ldots, h_{k-1} \leq g_{k-1}, h_{k}>g_{k} ;
$$

OR: $\quad h_{t}-h_{m} \geq D h$, and:
$h_{n} \leq g_{n}, h_{n-1} \leq g_{n-1}, \ldots, h_{k+1} \leq g_{k+1}$

$$
\begin{equation*}
\mathrm{h}_{\mathrm{k}}>\mathrm{g}_{\mathrm{k}} \tag{2.11}
\end{equation*}
$$

$=0$ otherwise.
2.4.3 Making the Goal Function Suitable for Numerical Evaluation

Replacing the cost and loss terms in the objective function (2.7), the following expression is obtained:

$$
\begin{equation*}
\min _{\underline{G}}\left\{\sum_{k=1}^{n} c_{k}\left(g_{k}\right)+\int_{\underline{H}}\left[\sum_{k=1}^{n} l_{k}\left(g_{k}, h_{k}\right) t_{k}(\underline{G}, \underline{H})\right] f(\underline{H}) d \underline{H}\right\} \tag{2.12}
\end{equation*}
$$

Based on the assumption that, for heights at the extremes $\left(g_{1}, g_{n}\right)$, the optimum levee profile follows the water surface curve with the same extremes, a substantial simplification to the decision problem can be introduced.

Namely, the reduction of the set of decision variables from the n-element set: $\left[g_{1}, g_{2}, \ldots, g_{n}\right]$ to the two-element set: [ $g_{2} g_{n}$ ]. Taking also into consideration that the ordinates $h_{1}, h_{2}, \ldots, h_{n}$ of the water surface curve are functions of the extreme stages $\left(h_{1}, h_{n}\right)$, the goal function can be expressed in the form:

$$
\begin{align*}
& \min _{\left(g_{1}, g_{n}\right)}\left\{\sum_{k=1}^{n} c_{k}\left(g_{k}\left(g_{1}, g_{n}\right)\right)+\int_{h_{n}} \int_{h_{l}}\left[\sum_{k=1}^{n} \cdot\right.\right. \\
& \quad l_{k}\left[g_{k}\left(g_{1}, g_{n}\right), h_{k}\left(h_{1}, h_{n}\right)\right] t_{k}\left[\underline{G}\left(g_{1}, g_{n}\right), \underline{H}\left(h_{1}, h_{n}\right]\right] \\
&  \tag{2.13}\\
& \left.\quad f\left(h_{1}, h_{n}\right) d h_{l} d h_{n}\right\} .
\end{align*}
$$

In the numerical evaluation of expression (2.13), the gereration of water profiles by the step method, for each point in the double numerical integration of the expected losses, may render the problem computationally unfeasible if the number of cross sections is moderately large. Even for a computer solution, it is necessary to introduce additional simplifications.

Szidarovszky, Duckstein, and Bogardi (1975) have proposed a simplification consisting in the evaluation of the $\mathrm{h}_{\mathrm{k}}$ 's by the linear regression model:

$$
\begin{equation*}
h_{k}=a_{k} h_{t}+b_{k}+c_{k} h_{m} \tag{2.14}
\end{equation*}
$$

where $h_{m}$ and $h_{t}$ are equivalent to $h_{l}$ and $h_{n}$, and the coefficients $a_{k}, b_{k}$, and $c_{k}$ are estimated by least squares from a large number of profiles generated by the step
method. This simplification will be adopted here for reasons of computational speed. The vector $\underline{h}$ will subsequently denote the pair ( $h_{m}, h_{t}$ ).
2.4.4 Use of the Conditional pdf's in the Evaluation of the Expected Damages

It has been mentioned earlier that the partial duration series can be used to estimate the conditional pdf's: $f_{1}\left(h_{m}, h_{t} \mid E l\right), f_{2}\left(h_{m}, h_{t} \mid E 2\right)$, and $f_{3}\left(h_{m}, h_{t} \mid E 3\right)$. Their role in the evaluation of expected damages will now be investigated. Note: In the remainder of this section, the notation $e_{1}, e_{2}, e_{3}$ will be used to denote El,E2,E3 in order to avoid confusion with the expected value E[ ].

Let $N_{i}$ be the number of events $e_{i}$ that occur in one year, and $P_{i}\left(N_{i}=n_{i}\right), i=1,2,3$, the probability that event $e_{i}$ occurs $n_{i}$ times in one year. Also let $S_{i}$ be the set of pairs ( $h_{m}, h_{t}$ ) corresponding to event $e_{i}$. The following assumptions are made:

1. The number of yearly occurrences of the events are independent; i.e., $P\left(N_{1}, N_{2}, N_{3}\right)=P_{1}\left(N_{1}\right) \cdot P_{2}\left(N_{2}\right) \cdot$ $\mathrm{P}_{3}\left(\mathrm{~N}_{3}\right)$.
2. The losses are time independent functions of the water stages $\underline{h}=\left(h_{m}, h_{t}\right)$ and do not depend on the number of events per year or the time between floods.
3. The flood stages $\underline{h}=\left(h_{m}, h_{t}\right)$ are assumed to be independent sample elements from the same family, for each type of event.

Given that there are $n_{1}$ events $e_{1}$ in a year, the losses associated with the $n_{1}$-tuple of water stage pairs $\left(\underline{h}_{1}, \underline{h}_{2}, \ldots, \underline{h}_{n_{1}}\right)$ are:

$$
\begin{equation*}
\mathrm{YL}=\sum_{i=1}^{\mathrm{n}_{1}} \mathrm{~L}\left(\underline{h}_{i}\right) . \tag{2.15}
\end{equation*}
$$

The conditional expected losses are:

$$
\begin{aligned}
E\left[Y L \mid N_{1}=\mathrm{n}_{1}\right] & =\int_{S_{1}} \ldots \int_{S_{1}}\left[L\left(\underline{h}_{1}\right)+\left(L\left(\underline{h}_{2}\right)+\ldots+L\left(\underline{h}_{n_{1}}\right)\right]\right. \\
& \mathrm{f}\left[\underline{h}_{1}, \underline{h}_{2}, \ldots, \underline{h}_{n_{1}}\right] d\left[\underline{h}_{1}, \underline{h}_{2}, \ldots, \underline{h}_{n_{1}}\right] .
\end{aligned}
$$

By assumption 3,

$$
\begin{align*}
& E\left[Y L \mid N_{1}=n_{1}\right]=\int \ldots \int n_{1} L[\underline{h}]\left[f_{1}(\underline{h}) d \underline{h}\right]{ }^{n} 1 \\
& \quad=n_{1}\left[\int_{S_{1}} L(\underline{h}) f_{1}(\underline{h}) d \underline{h}\right]\left[\int_{S_{1}} \ldots \int_{S_{1}}\left[f_{1}(\underline{h}) d \underline{h}\right]{ }^{n_{1}-1}\right] \\
& =n_{1} E_{1}[L], \text { where } E_{1}[L] \\
& \quad=\int_{S_{1}} L(\underline{h}) f_{1}(\underline{h}) d \underline{h} \tag{2.16}
\end{align*}
$$

This result can be extended to all three events, so that: $E\left[Y L \mid N_{i}=n_{i}\right]=n_{i} E_{i}[L]$. For $n_{i}$ events $e_{i}(i=1,2,3)$ in one year, the losses are:

$$
\left(T Y L \mid N_{1}=n_{1}, N_{2}=n_{2}, N_{3}=n_{3}\right)=\sum_{i=1}^{3} n_{i} E_{i}[L]
$$

The total expected yearly losses are:

$$
E[T Y L]=\sum_{n_{1}=0}^{\infty} \sum_{n_{2}=0}^{\infty} \sum_{n_{3}=0}^{\infty}\left[\sum_{i=1}^{3} n_{i} E_{i}[L]\right] P\left(n_{1}, n_{2}, n_{3}\right) .
$$

By assumption 1,

$$
\begin{aligned}
E[T Y L]= & \left.\sum_{n_{1}=0}^{\infty} \sum_{n_{2}=0}^{\infty} \sum_{n_{3}=0}^{\infty} \sum_{i=1}^{\left[\sum_{i}\right.} n_{i}[L]\right] P_{1}\left(n_{1}\right) \cdot \\
& P_{2}\left(n_{2}\right) \cdot P_{3}\left(n_{3}\right) .
\end{aligned}
$$

Finally:

$$
\begin{equation*}
E[T Y L]=E_{1}[L] E\left[N_{1}\right]+E_{2}[L] E\left[N_{2}\right]+E_{3}[L] E\left[N_{3}\right] \tag{2.17}
\end{equation*}
$$

In conclusion, the expected damages can be computed by means of the conditional pdf's and there is not even need to determine the distributions of the yearly frequency of the flood events.
2.4.5 Final Form of the Goal Function

After incorporating the results of the preceding section, the final form of the goal function is:

$$
\begin{gather*}
\min _{\left(g_{m}, g_{t}\right)}\left\{\sum_{k=1}^{n} c_{k}\left[g_{k}\left(g_{m}, g_{t}\right)\right]+\sum_{i=1}^{3} E\left[N_{i}\right] \int_{S_{i}} L\left[\underline{G}\left(g_{m}, g_{t}\right),\right.\right. \\
\left.\left.\underline{H}\left(h_{m}, h_{t}\right)\right] f_{i}\left(h_{m}, h_{t}\right) d h_{m} d h_{t}\right\} \tag{2.18}
\end{gather*}
$$

### 2.5 The Complete Model

The three major components of the model, namely hydraulic, hydrologic, and economic, have thus been analyzed in detail. They will be assembled, along with an integration and a minimization routine, into the complete model. The assemblage is shown in Figure 2.l.


## CHAPTER 3

## SELECTION OF THE HYDROLOGIC SUBMODEL

In the present chapter, the problem of choosing the appropriate pdf for the hydrologic component of the model will be considered. The pdf of flood magnitudes is, obviously, a decisive component of the total model, since the final results are expected to be heavily dependent upon the type of distribution used to compute the expected losses.

### 3.1 Generalities

Since extreme events are considered, and samples contain small number of realizations of them, the model uncertainty is large. Two pdf's may fit a sample relatively well, but their tails may differ substantially. Since the losses associated with higher flood levels are large, the discrepancy in the tails is amplified manyfold as it is multiplied by the losses.

The choice of a pdf to represent a natural phenomenon can be made on a causal basis when, as a result of his knowledge about the internal mechanism of the phenomenon, the modeller can conclude that it follows a certain probabilisitic model. There are also situations where the model selection is based on convenience and
computational tractability; or even may have been regulated by design standards. Finally, the choice may be based only on the information contained in the sample. This is the approach to be used here, and it will be discussed in greater detail in the following paragraphs.
3.1.1 Model Selection from Sample Information

Given a sample of flood events, the problem is to determine the probability density function (pdf) that best models the phenomenon where the sample was taken from. This pdf will be used to predict the relevant aspects of the process in order to take some decision based on such predictions.

The simplest approach is to take the sample relative frequencies as the pdf itself. This is equivalent to assume that the future. will be like the past. Even for very large samples (uncommon in hydrologic problems), this approach has at least three serious drawbacks: (1) it ignores changes in the flow regime due to developments in the basin and in the river itself (for example urbanization, deforestation, and dredging); (2) it also ignores long term natural trends such as long term climatic changes; and (3) discrete records can produce serious distortions in the model, especially toward the tails. For example, if in the "true" model the probability of a flood between 50,000 and $60,000 \mathrm{cu} \mathrm{ft/sec}$ is . O01, one occurrence of such a flood in a 100 year sample
would make this probability equal to . 01 , that is, ten times larger.

A more rational approach is to assume that the "true" model belongs to one of the well known families of distributions, such as the normal, lognormal, gamma, etc. Once a set of candidate families has been decided upon, the parameters for one paf of each family can be obtained by estimation methods such as maximum likelihood or the method of moments. Then, the selection of the "best" pdf can be based, for example, on how well it fits the sample.

In relation to model choice, Wood et al. (1974, p. 27) raise an important point: "Most hydrologic processes are so complex that no model yet devised may be the true model or that no hydrologic events follow one particular model." Consequently, it could be reasonably expected that a combination of models, for example a weighted sum of individual pdf's from different families, would better "explain" the hydrologic process, than does a unique well known pdf.

In the present work, pdf's from different families will be fitted to the historical records, and a composite model, consisting in the sum of the individual pdf's weighted by their sample likelihood, will also be constructed. Some bivariate distributions (with non-zero correlation coefficient) will be used; but, since the sample correlation coefficient is low, independence will be
assumed in other models by taking the product of their marginal pdf's. The candidate distributions to be used, along with the estimators for their parameters, will now be presented.

### 3.2 Distributions to be Used

In the case study of the Zagyva River, considered in Chapter 4, several candidate distributions are selected in basis to the shape of the marginal histogram. This section presents the functional form of those distributions and their parameter estimators. The known base stages, $\mathrm{b}_{\mathrm{m}}$ and $b_{t}$, of the partial duration series (see section 2.2) will be used as the "shift" or location parameter for distributions such as the lognormal or the exponential. The variables $x$ and $y$ will be used subsequently instead of $h_{m}$ and $h_{t}$; also, E[.] and DT[.] stand for expected value and standard deviation, respectively.
3.2.1 Lognormal Distribution

The three parameter lognormal pdf (Johnson and Kotz, 1970a, p. l12) has the form:

$$
\begin{gather*}
f(x ; \theta, \xi, \sigma)=(\sqrt{2 \pi}(x-\theta) \sigma)^{-1} \exp \left(-.5[\log (x-\theta)-\xi]^{2} / \sigma^{2}\right) \\
x>\theta \tag{3.1}
\end{gather*}
$$

where $\theta$ is the location parameter,

$$
\begin{align*}
\xi & =\mathrm{E}[\log (x-\theta)],  \tag{3.2}\\
\sigma & =\mathrm{DT}[\log (x-\theta)] . \tag{3.3}
\end{align*}
$$

For $\theta$ known, the maximum likelihood estimators (MLE) of $\xi$ and $\sigma$ are:

$$
\begin{align*}
& \hat{\xi}=n^{-1} \Sigma \log \left(x_{i}-\theta\right)  \tag{3.4}\\
& \hat{\sigma}=\left(n^{-1} \sum\left[\log \left(x_{i}-\theta\right)-\hat{\xi}\right]^{2}\right)^{1 / 2} \tag{3.5}
\end{align*}
$$

### 3.2.2 Exponential Distribution

The exponential distribution with two parameters (Johrson and Kotz, 1970a, p. 207) has the form:

$$
\begin{equation*}
f(x ; \lambda, \theta)=\lambda^{-1} \exp [-(x-\theta) / \lambda], \quad x>\theta, \quad \lambda>0 . \tag{3.6}
\end{equation*}
$$

Where $\theta$ is the location parameter, and $\lambda=E[x-\theta]$. For $\theta$ known, the MLE for $\lambda$ is:

$$
\begin{equation*}
\hat{\lambda}=\mathrm{n}^{-1} \Sigma\left(\mathrm{x}_{\mathrm{i}}-\theta\right) . \tag{3.7}
\end{equation*}
$$

### 3.2.3 Gamma Distribution

The gamma distribution (Johnson and Kotz, 1970a,
p. 166) has the form:

$$
\begin{align*}
& f(, \alpha, \beta, \gamma)=\frac{(x-\gamma)^{\alpha-1}}{\beta^{\alpha} \Gamma(\alpha)} \exp [-(x-\gamma) / \beta],  \tag{3.8}\\
& x>\gamma, \alpha>0, \quad \beta>0 .
\end{align*}
$$

Assuming that the parameter $\gamma$ is known, the MLE for $\alpha$ and $\beta$ are obtained by solving the equations:

$$
\begin{align*}
& n^{-1}\left[\sum \log \left(x_{i}-\gamma\right)\right]=\log \hat{\beta}+\psi(\hat{\alpha}),  \tag{3.9}\\
& \log (\bar{x}-\gamma)=\log (\hat{\alpha})+\log (\hat{\beta}), \tag{3.10}
\end{align*}
$$

where $\psi(\alpha)=\frac{d}{d \alpha}[\log (\Gamma[\alpha])] \quad$ (the "digamma" function).

### 3.2.4 Beta Distribution

Although the beta distribution does not have a tail, it will surely be useful as the marginal of $h_{t}$ in the conditional pdf: $f\left(h_{m}, h_{t} \mid E_{1}\right)$ ( $\subset f$. Chapter 2). The beta distribution (Johnson and Kotz, 1970b, p. 37) has the form:

$$
\begin{array}{r}
f(x ; a, b, p, q)=\left[B(p, q)(b-a)^{p+q-1}\right]^{-1}(x-a)^{p-1}(b-x)^{q-1} \\
\text { for: } a<x<b ; p>0 ; q>0 ; B(p, q)=\frac{\Gamma(p) \Gamma(q)}{\Gamma(p+q)} \tag{3.11}
\end{array}
$$

FOr $a$ and $b$ known, the MLE of $p$ and $q$ are obtained by solving the equations:

$$
\begin{align*}
& \psi(\mathrm{p})-\psi(\mathrm{p}+q)=\mathrm{n}^{-1} \sum \log \left(\frac{\mathrm{x}_{\mathrm{i}}-\mathrm{a}}{\mathrm{b-a}}\right)  \tag{3.12}\\
& \psi(\mathrm{q})-\psi(\mathrm{p}+q)=\mathrm{n}^{-1} \sum \log \left(\frac{\mathrm{~b}-\mathrm{x}_{i}}{\mathrm{b-a}}\right) \tag{3.13}
\end{align*}
$$

where $\psi($.$) represents the "psi" or "digama" function.$
3.2.5 Normal Distribution

The normal distribution (Johnson and Kotz, 1970a,
p. 40) has the form:

$$
\begin{equation*}
f(x ; \mu, \sigma)=[\sqrt{2 \pi} \sigma]^{-1} \exp \left[-.5(x-\mu)^{2} / \sigma^{2}\right] . \tag{3.14}
\end{equation*}
$$

The MLE for $\mu$ and $\sigma$ are:

$$
\begin{gather*}
\hat{\mu}=n^{-1} \sum\left(x_{i}\right)  \tag{3.15}\\
\hat{\sigma}=\left(n^{-1} \sum\left(x_{i}-\hat{\mu}\right)^{2}\right)^{1 / 2} \tag{3.16}
\end{gather*}
$$

3.2.6 Truncated Normal Distribution

Based on the fact that only flood magnitudes above
a base level are being considered, it seems reasonable to expect that the tail of a normal distribution would fit the sample fairly well. The truncated normal distribution is then a common normal pdf, but defined only for values of $x$ above certain value $x_{L}$ and multiplied by a normalizing constant. The distribution has the form:

$$
f\left(x: x_{L}, \mu, \sigma\right)=\left\{\begin{array}{l}
{[K \sqrt{2 \pi} \sigma]^{-1} \exp \left[-.5(x-\mu)^{2} \sigma^{-2}\right], x_{L}<x<\infty}  \tag{3.17}\\
0, \quad-\infty<x \leq x_{L}
\end{array}\right.
$$

where:

$$
K=\int_{z_{L}}^{\infty}(2 \pi)^{-.5} \exp \left(-.5 u^{2}\right) d u, z_{L}=\left(x_{L}-\mu\right) / \sigma .
$$

For known $\mathrm{x}_{\mathrm{L}}$, the following equations, relating the first to moments, $\mu_{1}^{\prime}$ and $\mu_{2}^{\prime}$, of the truncated distribution to the mean and variance of the complete distribution, can be obtained:

$$
\begin{align*}
& \mu_{1}^{\prime}=\left[f\left(x_{L}, \mu, \sigma\right) / F\left(x_{L}, \mu, \sigma\right)\right]+\mu  \tag{3.18}\\
& \mu_{2}^{\prime}=\left(\sigma x_{L}+\sigma \mu\right)\left[f\left(x_{L}, \mu, \sigma\right) / F\left(x_{L}, \mu, \sigma\right)+\sigma^{2}+\mu^{2}\right. \tag{3.19}
\end{align*}
$$

where

$$
\begin{gathered}
f\left(x_{L}, \mu, \sigma\right)=\exp \left[-.5\left(x_{L}-\mu\right)^{2} / \sigma^{2}\right] / \sqrt{2 \pi \sigma} \\
F\left(x_{L}, \mu, \sigma\right)=\int_{x_{L}}^{\infty} f(u, \mu, \sigma) d u .
\end{gathered}
$$

### 3.2.7 Bivariate Gamma Distribution

In the literature, several forms of three-parameter bivariate gamma distributions can be found (see, for example, Mardia, 1970; or Johnson and Kotz, 1972). Ghirtis (1967) presents a method for estimation of parameters of the five parameter form of the "Double Gamma" distribution introduced by David and Fix (1961). The five parameter double gamma distribution (Ghirtis, 1967) has the form:

$$
\begin{align*}
& f(x, y ; a, b, c, \lambda, \mu)= \\
& k^{-1} \exp \left[-\frac{x}{\lambda}-\frac{y}{\mu}\right] \int_{0}^{m} u^{a-1}(x-\lambda u)^{b-1}(y-\mu u)^{c-1} e^{u} d u \tag{3.20}
\end{align*}
$$

where:

$$
\begin{aligned}
& \mathrm{k}=\lambda^{\mathrm{b}} \mathrm{c}^{\mathrm{c}} \Gamma(\mathrm{a}) \Gamma(\mathrm{b}) \Gamma(\mathrm{c}) \\
& \mathrm{m}=\min [\mathrm{x} / \lambda, \mathrm{y} / \mu] \\
& \mathrm{a}, \mathrm{~b}, \mathrm{c}, \lambda, \mu>0 \\
& \mathrm{x}>0, \mathrm{y}>0
\end{aligned}
$$

The parameters $a, b, c, \lambda$, and $\mu$ can be computed by the equations (Ghirtis, 1967):

$$
\begin{align*}
\hat{\lambda} & =k_{20} / k_{10}  \tag{3.21}\\
\hat{\mu} & =k_{02} / k_{01}  \tag{3.22}\\
\hat{a} & =k_{11} k_{10} k_{01} / k_{20} k_{02}  \tag{3.23}\\
\hat{b} & =k_{10}^{2} / k_{20}-\hat{a}  \tag{3.24}\\
\hat{c} & =k_{01}^{2} / k_{02}-\hat{a} \tag{3.25}
\end{align*}
$$

Where $k_{i, j}$ is the $i, j-t h$ cummulant of the population and can be estimated by the Fisher's bivariate $k$ statistics:

$$
\begin{aligned}
& \mathrm{k}_{10}=\mathrm{s}_{10} / \mathrm{n} \\
& \mathrm{k}_{20}=(\mathrm{n}-1)^{-1}\left(\mathrm{~s}_{20}-\mathrm{s}_{10}^{2} / \mathrm{n}\right) \\
& \mathrm{k}_{11}=(\mathrm{n}-1)^{-1}\left(\mathrm{~s}_{11}-\mathrm{s}_{10} \mathrm{~S}_{01} / n\right), \mathrm{s}_{i j}=\Sigma \mathrm{x}^{\mathrm{i}} \mathrm{y}^{j} .
\end{aligned}
$$

3.2.8 Truncated Bivariate Normal Distribution

The truncated bivariate normal distribution, for $y$ truncated at $Y_{L}$, has the form:

$$
\begin{align*}
& \mathrm{f}\left(\mathrm{x}, \mathrm{y} ; \mu_{\mathrm{x}}, \sigma_{\mathrm{y}}, \sigma_{\mathrm{x}}, \sigma_{\mathrm{Y}}, \rho, \mathrm{Y}_{\mathrm{L}}\right)= \\
& \exp \left[-.5\left(\mathrm{z}^{2}-2 \rho \mathrm{zw}+\mathrm{w}^{2}\right) /\left(1-\rho^{2}\right)\right] /\left(2 \pi k \sigma_{\mathrm{x}} \sigma_{\mathrm{y}} \sqrt{1-\rho^{2}}\right) \tag{3.26}
\end{align*}
$$

where

$$
\begin{aligned}
& k=\int_{y_{L}}^{\infty} \exp \left[-.5\left(z^{2}-2 \rho z w+w^{2}\right) /\left(1-\rho^{2}\right)\right] d y /\left(2 \pi \sigma_{x} \sigma_{y} \sqrt{1-\rho^{2}}\right) \\
& z=\left(x-\mu_{x}\right) / \sigma_{x} \\
& w=\left(y-\mu_{y}\right) / \sigma_{y} .
\end{aligned}
$$

Des Raj (1952) presents the following equations to compute the MLE of the parameters $\mu_{x}, \mu_{y}, \sigma_{x}, \sigma_{y}$, and $\rho$, when the truncation point, $Y_{L^{\prime}}$, is known.

$$
\begin{align*}
& v_{02}=\sigma_{y}^{2}\left(1-k^{\prime}\left(z_{1}-k^{\prime}\right)\right)  \tag{3.27}\\
& v_{01}=\sigma_{y}\left(z_{1}-k^{\prime}\right)  \tag{3.28}\\
& \mu_{y}=y_{L}-k^{\prime} \sigma_{y}  \tag{3.29}\\
& \mu_{x}+\left(z_{1}\right) \rho \sigma_{x}-v_{10}=0
\end{align*}
$$

$$
\begin{align*}
& \sigma_{y}\left(z_{1}-k^{\prime}\right) \mu_{x}+\sigma_{y} \rho \sigma_{x}-v_{l l}=0  \tag{3.31}\\
& \sigma_{x}^{2}=v_{20}-\mu_{x}\left(\mu_{x}+2 \rho \sigma_{x} z_{1}\right)-\left(\rho \sigma_{x}\right)^{2} k^{\prime} z_{1} \tag{3.32}
\end{align*}
$$

where $v_{i j}$ is the ij-th sample moment about the truncation,

$$
\begin{aligned}
& k^{\prime}=\left(y_{L}-\mu_{Y}\right) / \sigma_{y} \\
& z_{l}=\phi\left(k^{\prime}\right) /\left[1-\int_{k^{\prime}}^{\infty} \phi(t) d t\right] \\
& \phi(t)=(2 \pi)^{-.5} \exp \left[-.5 t^{2}\right] .
\end{aligned}
$$

3.2.9 Truncated Bivariate Lognormal Distribution

The truncated bivariate lognormal distribution for $y$ truncated at $y_{L}$ has the form:

$$
\begin{align*}
& \mathrm{f}\left(\mathrm{x}, \mathrm{y} ; \theta_{\mathrm{x}}, \theta_{\mathrm{y}}, \xi_{\mathrm{x}}, \xi_{\mathrm{y}}, \sigma_{\mathrm{x}}, \sigma_{\mathrm{y}}, \rho, \mathrm{y}_{\mathrm{L}}\right)= \\
& \exp \left[-.5\left(\mathrm{z}^{2}-2 \rho \mathrm{zw}+\mathrm{w}^{2}\right) /\left(1-\rho^{2}\right)\right] /\left[\left(\mathrm{x}-\theta_{\mathrm{x}}\right)\left(\mathrm{y}-\theta_{\mathrm{y}}\right) \sigma_{\mathrm{x}} \sigma_{\mathrm{y}} 2 \pi k \sqrt{\left.1-\rho^{2}\right]}\right. \\
& \quad \theta_{\mathrm{x}}<\mathrm{x}<\infty, \quad \theta_{\mathrm{y}}<\mathrm{y}_{\mathrm{L}}<\mathrm{y}<\infty, \tag{3.33}
\end{align*}
$$

where

$$
\begin{aligned}
& \mathrm{z}=\left(\log \left(\mathrm{x}-\theta_{\mathrm{x}}\right)-\xi_{\mathrm{x}}\right) / \sigma_{\mathrm{x}} \\
& \mathrm{w}=\left(\log \left(\mathrm{y}-\theta_{\mathrm{y}}\right)-\xi_{\mathrm{y}}\right) / \sigma_{\mathrm{y}} \\
& \xi_{\mathrm{x}}=E\left[\log \left(\mathrm{x}-\theta_{\mathrm{x}}\right)\right] \\
& \xi_{\mathrm{y}}=E\left[\log \left(\mathrm{y}-\theta_{\mathrm{y}}\right)\right] \\
& \sigma_{\mathrm{x}}=\sigma\left[\log \left(\mathrm{x}-\theta_{\mathrm{x}}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
& \sigma_{y}=\sigma\left[\log \left(y-\theta_{y}\right)\right] \\
& k=\int_{z_{L}}^{\infty} \frac{\exp \left(-u^{2} / 2\right) d u}{\sqrt{2 \pi}}, \text { with } z_{L}=\left(\log \left(y_{L}-\theta_{y}\right)-\xi_{y}\right) / \sigma_{y}
\end{aligned}
$$

For known $y_{L}$ and location parameters, $\theta_{x}$ and $\theta_{y}$, the parameters $\xi_{x}, \xi_{y}, \sigma_{x}, \sigma_{y}$, and $\rho$ can be computed as in the truncated bivariate normal after making the transformation;

$$
\begin{align*}
& x^{\prime}=\log \left(x-\theta_{x}\right)  \tag{3.34}\\
& y^{\prime}=\log \left(y-\theta_{y}\right) . \tag{3.35}
\end{align*}
$$

### 3.3 Criteria for Model Selection

The point has finally been reached where a set of distributions is available to choose from. Each pdf belongs to a different family and, when its parameters have been estimated by the method of maximum likelihood, it is the most likely source of the data in the sample, given that the model space is restricted to that particular family When confronted with the data, some of the candidates will show such poor fits that they could be discarded at once. More powerful tools will be required, however, to discriminate between those models that fit the sample equally well. A review of the most relevant methods for the purposes of the present work is presented in the following sections.

### 3.3.1 Goodness of Fit

Several model verification methods are based in "goodness of fit" criteria; that is, a comparison between the relative frequencies contained in the sample (histogram) and those predicted by the hypothesized model. A quick visual comparison of the histogram and the (discretized) pdf may point out enough discrepancy to reject the model. Except for obvious or extreme cases, the subjective judgment involved makes this method inappropriate.

Statistical goodness of fit tests constitute a more objective and precise way to measure discrepancies between the model and the sample. The Chi square and the KolmogorovSmirnov (KS) are the most widely used goodness of fit tests. (A very good presentation on the application of these tests can be found in Benjamin and Cornell, 1970, Chapter 4.) Actually, these tests are not designed for model choice, but to test statistical hypotheses of the form: "The random variable $x$ is distributed $f_{0}\left(x ; \theta_{0}\right)$ "; where $f_{o}$ is a particular pdf, and $\theta_{0}$ is a point in its parameter space.

The Chi square goodness of fit test is based on the statistic:

$$
\begin{equation*}
D=\sum_{i=1}^{k} \frac{\left(N_{i}-n P_{i}\right)^{2}}{n P_{i}} \tag{3.36}
\end{equation*}
$$

where

$$
\begin{aligned}
& N_{i}=\text { the number of sample points in interval } I_{j}, \\
& P_{i}=\int_{x \varepsilon I_{j}} f_{o}\left(x ; \theta_{0}\right) d x, \text { and }
\end{aligned}
$$

$\mathrm{n}=$ the sample size.
For: $k=$ number of intervals, $p=$ number of parameters estimated from the sample, the statistic $D$ is approximately distributed Chi square with $k$ - $p$ - l degrees of freedom. The Kolmogorov-Smirnov goodness of fit test is based on the statistic:

$$
\begin{equation*}
D=\operatorname{Max}\left|F_{n}(x)-F(x)\right| \tag{3.37}
\end{equation*}
$$

where $\mathrm{F}_{\mathrm{n}}(\mathrm{x})$ is the sample cumulative distribution function, and $F(x)$ is the hypothesized CDF (cumulative distribution function). The KS goodness of fit test has the advantage that it does not require the definition of a priori intervals, as in the Chi square test. On the other hand, it should only be used for continuous univariate distributions, and when the distribution has been obtained independently of the sample. (In other words, the sample can not be used to estimate the parameters of the distribution.) When the parameters have been estimated from the sample, the critical (rejection) value should be smaller (Benjamin and Cornell, 1970, p. 488).

In the Chi square and the Kolmogorov-Smirnov goodness of fit tests, the null and the alternative hypotheses take the form:

$$
\begin{aligned}
& H_{0}: x \text { is distributed } f_{0}\left(x ; \theta_{0}\right) \\
& H_{1}: x \text { is not distributed } f_{0}\left(x ; \theta_{0}\right),
\end{aligned}
$$

and the critical region is: $D>c$. The computation of the
type $I$ error, $P\left(D \geq c \mid H_{0}\right)$ is a simple matter. However, since the alternative hypothesis is in a very general form, the type II error, $P\left(D<c \mid H_{1}\right)$, can not be computed.

Although the goodness of fit tests are not intended for discrimination between models, they can be used for a "preliminary screening" of the set of candidate pdf's to discard those that are rejected at a given significance level (say, 5\%).
3.3.2 Most Likely Value of the Goodness of Fit Statistic

Benjamin and Cornell (1970) suggest the use of the Chi square statistic as a tool for model choice by selecting the model for which the likelihood of the observed value of the corresponding closeness-of-fit statistic is largest. Such value is more related to the mode of the Chi square pdf than to the miniumum value of the statistic. This criterion is based on the fact that the minimum value of the Chi square statistic is not the most likely outcome (except for the distribution with two degrees of freedom, where it is zero). In order to compare models using this criterion, it is necessary to have the same number of degrees of freedom for all models.
3.3.3 A Composite Model Based on Sample Likelihoods

Wood et al. (1974) have used the sample likelihoods to account for model uncertainty within a Bayesian
framework. They formulated the following composite model:

$$
\begin{equation*}
\hat{f}(q)=\sum \frac{K_{i}}{K_{*}} p^{\prime}\left(\theta_{i}=1\right) \quad \tilde{f}_{i}(q), \tag{3.38}
\end{equation*}
$$

where
$\tilde{f}_{i}(q)=$ the "Bayesian Distribution" given model $i$,
$\theta_{i}\left\{=1\right.$ if $f_{i}$ is the true model,
$i^{\{ }=0$ otherwise,
$P^{\prime}\left(\theta_{i}=1\right)$ is the prior probability of $f_{i}$ being the true model,
$K_{i}$ is the marginal likelihood function of the observations for model i.

For $\underline{A}=$ Model parameters,

$$
\begin{aligned}
& K_{i}=\int f\left(q \mid \underline{A}_{i}, M_{i}\right) f\left(\underline{A}_{i} \mid M_{i}\right) d \underline{A}_{i} \\
& K_{*}=\Sigma K_{i} \cdot P^{\prime}\left(\theta_{i}=l\right) .
\end{aligned}
$$

A Bayesian approach has not been attempted here because of the complexity of the Bayesian distribution in the bivariate case. (The Bayesian distribution takes into account the parameter uncertainty.) Nevertheless, a similar composite model can be developed without taking into account the parameter uncertainty.

Let:

$$
\begin{equation*}
f(x \mid \underline{\theta})=\sum_{k=1}^{m} \theta_{k} f_{k}(x) \tag{3.39}
\end{equation*}
$$

so that the composite model has the form:

$$
f(x)=\int f(x \mid \underline{\theta}) f(\underline{\theta}) d \underline{\theta},
$$

where $\underline{\theta}$ is an $m$ vector such that:

$$
\begin{aligned}
\theta_{\mathrm{k}}\{ & =1 \text { if } \mathrm{f}_{\mathrm{k}}(\mathrm{x}) \text { is the true model, } \\
& =0 \text { otherwise }
\end{aligned}
$$

and

$$
\sum_{k=1}^{m} \theta_{k}=1
$$

Notice that this definition reduces the possible values of $\theta$ to: 100...0, 010...0, 001...0, ..., 000...l. The likelihood function given a sample $\underline{\text { x }}$ is:

$$
\begin{align*}
L(\underline{x} \mid \underline{\theta}) & =\prod_{i=1}^{n}\left[\sum_{k=1}^{m} \theta_{k} f_{k}\left(x_{i}\right)\right] \\
& =\sum_{k} \theta_{k}\left[\Pi f_{i}\left(x_{i}\right)\right] \\
& =\sum_{k} \theta_{k} L_{k}(\underline{x}) \tag{3.40}
\end{align*}
$$

Notice that there are no crossproducts because of the restrictions imposed on $\underline{\theta}$. Also, $L_{k}(\underline{x})$ is the sample likelihood for model $k$.

Let the prior pdf of $\underline{\theta}$ be:

$$
\begin{equation*}
f^{\prime}(\underline{\theta})=\sum_{i=1}^{m} \theta_{i} p^{\prime}\left(\theta_{i}=1\right), \tag{.3.41}
\end{equation*}
$$

where $\mathrm{p}^{\prime}(\theta=1)$ is the prior probability of model i. Given a sample $\underline{x}=\left(x_{1}, x_{2}, \ldots, x_{n}\right)$, the prior pdf of $\underline{\theta}$ can be updated using Bayes Theorem:

$$
\begin{aligned}
f^{\prime \prime}(\underline{\theta} \mid \underline{x}) & =\frac{f(\underline{x} \mid \underline{\theta}) \cdot f^{\prime}(\underline{\theta})}{f(\underline{x})} \\
& =\frac{L(\underline{x} \mid \underline{\theta}) \cdot f^{\prime}(\underline{\theta})}{\int f(\underline{x} \mid \underline{\theta}) f^{\prime}(\underline{\theta}!\underline{\theta}}
\end{aligned}
$$

$$
\begin{align*}
& =\frac{\left[\Sigma \theta_{i} L_{i}(\underline{x})\right] \cdot\left[\Sigma \theta_{i} P^{\prime}\left(\theta_{i}=l\right)\right]}{\int\left[\Sigma \theta_{i} L_{i}(\underline{x})\right] \cdot\left[\sum \theta_{i} P^{\prime}\left(\theta_{i}=1\right)\right] d \underline{\theta}} \\
& =\frac{\sum \theta_{i} L_{i}(\underline{x}) \cdot P^{\prime}\left(\theta_{i}=l\right)}{\Sigma L_{i}(\underline{x}) P^{\prime}\left(\theta_{i}=1\right)} \\
& =\Sigma \theta_{i} P^{\prime \prime}\left(\theta_{i}=l\right) \tag{3.42}
\end{align*}
$$

where

$$
P^{\prime \prime}\left(\theta_{i}=1\right)=\frac{L_{i}(\underline{x}) \cdot P^{\prime}\left(\theta_{i}=1\right)}{\sum L_{i}(\underline{x}) \cdot P^{\prime}\left(\theta_{i}=1\right)}
$$

are the posterior model probabilities.
The composite model can now be updated by replacing the posterior pdf of $\theta$ :

$$
\begin{align*}
\hat{\mathrm{f}}(x) & =\int \hat{\mathrm{f}}(x \mid \underline{\theta}) \mathrm{f}^{\prime \prime}(\underline{\theta}) \mathrm{d} \underline{\theta} \\
& =\int\left[\Sigma \theta_{i} f_{i}(x)\right] \cdot\left[\sum \theta_{i} P^{\prime \prime}\left(\theta_{i}=1\right)\right] d \underline{\theta} \\
& =\Sigma f_{i}(x) P^{\prime \prime}\left(\theta_{i}=1\right) . \tag{3.43}
\end{align*}
$$

So the composite model is a linear combination of the candidate models, weighted by their posterior probabilities. In the case of equal prior model probabilities $P^{\prime}\left(\theta_{i}=1\right)=P^{\prime}$, (reflecting perhaps a state of total ignorance about the true model) the posterior model probabilities become:

$$
\begin{equation*}
P^{\prime \prime}\left(\theta_{i}=1\right)=\frac{L_{i}(\underline{x})}{\sum L_{i}(\underline{x})} \tag{3.44}
\end{equation*}
$$

that is, proportional to the sample likelihood given that each model is the true model.
3.3.4 Use of Sample Likelihoods
for Model Selection
In addition to the composite model, the performance of the posterior model probabilities $P^{\prime \prime}\left(\theta_{i}=1\right)$ will be studied here, as a criterion for ranking the competing models for selection. At this point, the choice is not among families of distributions but, rather, among the pdf's with the maximum sample likelihood from each family. Hence, the paf so chosen has the maximum sample likelihood among all the possible pdf's in all the families considered. This does not intend to be a proof that it is the best model; but at least an intuitive argument in its favor.

## CHAPTER 4

## NUMERICAL RESULTS

The general model defined in Chapter 2 and the model selection procedures described in Chapter 3 will now be appJied to the case of the optimum levee profile for the Zagyva River, an important tributary of the Tisza River, both located in Hungarian territory. The low lands to the west of the Zagyva River are protected by a 60.4 km levee reach between the Jasztelek gaging station and the mouth in the Tisza River, at Szolnok. The existing levee, shown in Figure 4.21 (p. 92), does not follow a water surface. The implementation of the model will take into account that when the existing levee is higher than the proposed one, the existing height should be used. The average slope between the two ends of the reach is $.01 \%$.

Wetted cross sectional areas and hydraulic radii have been determined for 49 cross sections along the reach. Construction costs and flood losses for different levee heights are also available for each cross section (Szidarovszky et al., 1975; Bogardi et al., 1975; Szidarovszky and Yakowitz, 1976).

### 4.1 Implementation of the Hydraulic and Economic Submodels

The hydraulic information to be used in the generation of water surface profiles has been synthesized by Szidarovszky et al. (1975) in a linear regression model that gives the water level at each cross section as a function of the water levels at the ends of the reach (see Equation [2.14]). Two different sets of regression coefficients are used depending on the type of water surface curve: concave upward or "damping," when $\left(h_{t} h_{m}\right)<6.5 m$; concave downward or "draw down" when $\left(h_{t}-h_{m}\right) \geq 6.5 m$; where 6.5 m is the alitude difference between the two gaging stations at both ends of the reach. Tables A.l and A. 2 show a listing of the coefficients $a_{k}$, $b_{k}$, and $c_{k}$ for damping and draw down surface curves.

The economic submodel has been implemented with the construction costs and flood damage data used by Bogardi et al. (1975). Two modifications, however, have been introduced: (l) instead of piecewise linear functions, quadratic functions have been fitted to the data, and (2) the damage functions have been extrapolated beyond existing data by fitting them to a square root function. The damage functions have then the form:

$$
f(h)=\left\{\begin{array}{l}
a h^{2}+b h+c, h \leq-\frac{b}{2 a}-d, a<0, \\
(h-e)^{\cdot 5}+g, h>-\frac{b}{2 a}-d
\end{array}\right.
$$

where $h$ is the levee height, and $a, b, c, d$ are known constants (obtained from the damage data), and,

$$
\begin{aligned}
& e=-\frac{b}{2 a}-\left[4 a\left(-\frac{b}{2 a}-d\right)+2 b\right]^{-.5} \\
& g=f\left(-\frac{b}{2 a}-d\right)-\left(-\frac{b}{2 a}-d-e\right) .
\end{aligned}
$$

Figures 4.1 and 4.2 show typical cost and damage functions. The cost and damage coefficients are listed in Tables A. 3 and A. 4 respectively.

### 4.2 Tmp1ementation of the Hydrologic Submodel

The hydrologic submodel consists in the joint probability density function of flood magnitudes. A set of candidate pdf's will be determined from the historic records and the methods discussed in Chapter 3 will be applied for the selection of one pdf. A composite model, combining some of the candidate pdf's will also be constructed.

### 4.2.1 The Sample

The joint partial duration series for 36 years is available. Figure 4.3 shows a plot of the 68 sam points and Table A. 5 presents a listing of the sample. It is important to realize that the sample only contains pairs $(x, y)$ of water stages at both ends of the reach such that $x>X_{L}$ and/or $y>Y_{L}$, where $X_{L}=85.95 m$ and $Y_{L}=91.04 m$ are the base levels. In other words, a simultaneous reading was

Figure 4.1. Typical Cost Functions


Figure 4.2. Typical Damage Functions

taken only if the water surface was above the base level in the tributary and/or in the main river.

This sample space is, in fact, a subset of the set of all possible values of ( $\mathrm{x}, \mathrm{y}$ ), and it could be considered as an "L-shaped" tail of a bivariate distribution. An unsuccessful attempt was made to estimate the underlying ("complete") distribution from such "L-shaped" truncated sample. As an alternative, it was shown in Chapter 2 how the sample space can be divided into two or more events, so that the joint magnitude pdf's, conditional upon the occurrence of such events, can be computed. It was, then, decided to divide the sample space into two subsets (see Figure 4.3): event El, corresponding to flood in the main river only; and event $E 2$, corresponding to floods in the tributary. The sample contains 21 occurrences of event El and 47 occurrences of event E2. Hence, the expected number of events per year can be estimated as:

$$
\begin{align*}
& E\left[N_{1}\right]=21 / 36=.58333,  \tag{4.1}\\
& E\left[N_{2}\right]=47 / 36=1.3056 . \tag{4.2}
\end{align*}
$$

### 4.2.2 Candidate pdf's

After consideration of the marginal sample histograms, it was decided to fit a number of pdf's to the sample at hand. Since the sample correlation coefficients are small (less than .l0) both bivariate pdf's and products of independent marginal pdf's will be used. The equations
for estimation of parameters have been presented in Chapter 3. It will be noticed that the two sets of candidates are not identical. The candidate pdf's were selected in basis to the shape of the marginal histograms. The pdf's to be used and their estimated parameters are presented in Tables 4.1 and 4.2.
4.2.3 Goodness of Fit Tests

Using an $\alpha$ (probability of type $I$ error) of $5 \%$, univariate Chi-square tests were run for the marginal pdf's. The results of these tests are shown in Table 4.3. Based on the results of these tests, the set of candidate pdf's will be reduced to those not rejected at the 5\% level. It is important to mention that the Kolmogorov-Smirnov test did rot show enough discriminatory power in this particular situation. Using the $K-S$ test, none of the candidate marginals is rejected at the $5 \%$ lev.

Using an $\alpha$ of $5 \%$, bivariate Chi-square tests were run for the products of the non-rejected marginals, and for the bivariate distributions. Tables 4.4 and 4.5 show the results of these tests and also the abbreviations to be used subsequently in this text. It can be concluded from Tables 4.3 through 4.5 that this case study has too few degrees of freedom to make any valid inference; nevertheless, the methodology is illustrated.
Table 4.l. Pdf's of Flood Magnitudes Given Event El



Table 4.3. Results of Univariate Chi-square Tests

| pdf | D.F. | stat. | Reject | pdf | D.F. | stat. | Reject |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  |  |  |
| Lognormal | 3 | 13.77 | Yes | Lognormal | 3 | 3.03 | No |
| Exponen'ial | 4 | 7.98 | No | Exponential | 5 | 1.14 | No |
| Gamma | 2 | 7.96 | Yes | Gamma | 3 | 2.30 | No |
| Beta | 3 | 4.80 | No | Beta | 4 | 1.92 | No |
|  |  |  |  |  |  |  |  |
| Lognormal Gamma Normal | 9 | $\begin{aligned} & 34.50 \\ & 20.19 \\ & 13.20 \end{aligned}$ | Yes Yes No | Lognormal | 5 | 8.10 | No |
|  |  |  |  | Trnc. Logn. | 5 | 5.56 | No |
|  |  |  |  | Gamma | 4 | 7.27 | No |
|  |  |  |  | Trnc. Normal | 5 | 6.54 | No |



Table 4.5. Results of Rivariate Chi-square Tests for
Candidate pdf's to $f(x, y \mid E 2)-$ Abbreviated
names also given.

4.2.4 Selection of the "Best" pdf

There were initially, 18 candidate pdf's for $f(x, y \mid E 1)$ and 14 pdf's for $f(x, y \mid E 2)$. By the use of goodness of fit tests, these sets have been reduced to 10 and 6 candidates, respectively. Two different criteria will be now used to select the "best" pdf, namely, sample likelihoods and likelihood of the Chi-square statistic. Finally, a composite model as defined in Chapter 3 will be constructed.

### 4.2.4.1 Likelihood of the Bivariate Chi-Square

 Sta stic. Bivariate Chi-square statistics were recomputed for the non-rejected pdf's in order to have the same nunber of degrees of freedom in each group: the Chi-squace statistics for the candidates of $f(x, Y \mid E I)$ have three degrees of freedom, and those of $f(x, y \mid E 2)$ have eight degrees of freedom. Tables 4.6 and 4.7 present the Chisquare statistic for each candidate and its corresponding value of the ordinate in the Chi-square pdf with three or eight degrees of freedom, for $f(x, y \mid E 1)$ and $f(x, y \mid E 2)$, respectively. According to these tables, the models beta•beta for $f(x, y \mid E l)$ and normal truncated lognormal for $f(x, y \mid E 2)$ should be selected as the "best" models.4.2.4.2 Sample Likelihoods. Weights for each model were computed according to the definition given in Chapter 3 (Equation [3.44]) and are presented in Table 4.8.
Table 4.

| $p d f$ | E.L | E.E | E. G | E.B | B.Is | B. E | B. G | B.B | BVL | DBG |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $x^{2}$ stat. | 6.482 | 5.793 | 4.721 | 3.9 .18 | 5.300 | 4.637 | 3.389 | 2.872 | 7.011 | 4.714 |
| Probability | . 0397 | . 0530 | . 0818 | . 1113 | . 0649 | . 0846 | . 1349 | . 1608 | . 0317 | . 0820 |


| pdf | N. L | N. TL | N.G | N.TN | BVN | BVL |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $x^{2}$ stat. | 12.79 | 7.979 | 9.625 | 8.941 | 8.736 | 14.293 |
| Probability | . 0364 | . 0979 | . 0755 | . 0852 | . 0880 | . 0240 |

Table 4.8. Weights for Candidate pdf's Based on Sample Likelihoods

| $f(x, y \mid E l)$ |  | $f(x, y \mid E 2)$ |  |
| :---: | :---: | :---: | :---: |
| $p d f$ | weicht | pdf | weight |
| E.L | . 0167 | N.L | . 1433 |
| E.E | . 0025 | N.TL | . 1296 |
| E. G | . 0123 | N. G | . 4586 |
| E. B | . 0084 | N.TN | . 0173 |
| B. L | . 3721 | BVN | . 2496 |
| B. E | . 0536 | BVL | . 0018 |
| B. G | . 2657 |  |  |
| B. B | . 1769 |  |  |
| BVL | . 0709 |  |  |
| DBG | . 02.10 |  |  |
| Normalizing factor is: |  | Normalizing factor is: |  |
| $\Sigma L_{k}=4.37551 \mathrm{E}-14$ |  | $\Sigma L_{k}=9.8024624 \mathrm{E}-59$ |  |

According to Table 4.8, the models beta•lognormal for $f(x, y \mid E l)$ and normal•gamma for $f(x, y \mid E 2)$ should be selected as the "best" models.

$$
\text { 4.2 4. } 3 \text { A Composite Model. A composite model can }
$$

now be defined according to expression (3.43). Disregarding the models with weight less than .05 , and readjusting the weights so that they add up to 1.00 the following composite models are obtained:

$$
\begin{aligned}
& f_{C}(x, y \mid E I)=.42 f_{B \cdot L}+.30 f_{B \cdot G}+.20 f_{B \cdot B}+.08 f_{B V L} \\
& f_{C}(x, y \mid E 2)=.15 f_{N \cdot L}+.47 f_{N \cdot G}+.25 f_{B V N}+.13 f_{N \cdot T L}
\end{aligned}
$$

where the subindices are the abbreviations used to name the candidate models (see Tables 4.4 and 4.5).
4.2.5 Summary and Remarks on Model Selection

Two sets of pdis's (for $f(x, y \mid E l)$ and $f(x, y \mid E 2)$ were studied in order to determine the most appropriate model, given the sample at hand. After eliminating those pdf's that were rejected at the $5 \%$ level with the Chi-square test, two criteria were used to rank the remaining candidate models: the likelihood of the Chi-square statistics and the sample likelihoods. Figure 4.4 presents the relative ranking of the pdf's for the two choice criteria. The rat igs shown were obtained in the following manner: for each set of pdf's and for each choice criterion, a rating


Figure 4.4. Relative Ranking of Candidate pdf's
between 0 and 1 was obtained for each element of the set by dividing its indi $x$ (Chi-square probability or sample likelihood weight) by the greatest index in the set, for the criterion considered. Table 4.4 could also be viewed as a multicriteria decision problem. It would be interesting to apply some of the existing methods (such as ELECTRE, Benayoun, Roy, and Sussmann, 1966) to make a selection under the two criteria simultaneously.

It can be concluded from Figure 4.4 that the two choice criteria considerably disagree. Two factors seem to have contributed to this result. First, the small sample size used, especially for event El. A five-parameter bivariat distribution requires at least 7 cells to be defined in the sample space for the application of the Chisquare test. For a sample of $2 l$ points, the requirement of having at least. 5 points per cell cannot be et. This situation makes the results of the Chi-square tests rather unreliable. Again, the example is for illustrative purposes only. A second factor that may bias the results of the Chi-square test is the arbitrariness in the definition of the cells. It could perfectly happen that a given arrangement of cells might favor a particular pdef or group of pdit's and disfavor others. It was also realized that the sample likelihoods are extremely sensitive to errors in the normalizing constants (commonly used in the truncated models). A $1 \%$ error in such constant, making the computed
probabilities l\% larger, will double the sample likelihood in a sample of 70 points. This will erroneously increase the degree of importance of a model.

It also appears from the result that the sample likelihoods have a better ability to discriminate between models. This is well illustrated in Table 4.6 , where for $f(x, y \mid E)$ only threc models out of ten have a weight greater than . 10 . This ability to reduce the choice set in a very definite manner is a desirable feature in a model selection procedure.

When models are ranked to make a selection, it is possible that two or more models rank very close together. When the sample likelihoods are ured, the weights so obtained are actually the posterior probabilities of each model being the true one. In suri case, a composite model (as defined in Equation [3.1.8]) could be used, rather than making a more or less arbitrary choice of a unique pdf.

## 4. 3 Computational Consicorations

Before going into the acual numerical results, it is necessary to give some attention to the computer implementation of the model. A listinc of the program is presented in Appendix $B$. The implementation of the model is rolatively simple; however, it is necessary to consider here the computational aspects of the evaluation of expected
losses, which proved to be a critical factor in terms of execution time.
4.3.1 Integration Region for Expected Losses

Basically, the integration region for the evaluation of expected losses is the sample space itself. However, since the water level at the main river (x) cannot be greater than the water level at the tributary (y), the additional constraint $x \leq y$ must be introduced. The basic integration region is then:

$$
\begin{align*}
& S_{1}: \quad x \leq y, \quad x \geq 85.95, \quad 88.10 \leq y \leq 91.04  \tag{4.3}\\
& S_{2}: \quad x<y, \quad x \geq 77.9, \quad y \geq 91.04 \tag{4.4}
\end{align*}
$$

The basic integration region is shown in Figure 4.5 .


Figure 4.5. Regions in the Space of ( $x, y$ ) for a Design Alternative $\left(g_{m}, g_{t}\right)$
4.3.2 Convergence Problems in Numerical Integration

Using Silap: on's numerical integration, the evaluation of expected damages:'

$$
\begin{equation*}
\int_{(x, y) \varepsilon S_{1} U S_{2}} L\left(g_{m}, g_{t}, x, y\right) f(x, y) d x d y \tag{4.5}
\end{equation*}
$$

where:

$$
\begin{aligned}
& L\left(g_{m}, g_{t}, x, y\right)=\text { danages for water stag } s(x, y) \text { with } \\
& \quad \text { levee alternative }\left(g_{m}, g_{t}\right) \text {, } \\
& f(x, y)=\text { pdf of water stages, }
\end{aligned}
$$

presents convergence problems because of the existence of a region where the damages are zero, and a jump to non-zero damages across the boundary of such region. Consequently, the properties of the integration region were studied in detull and are discussed in the following paragraphs.

Given a levee design alternative ( $g_{m}, g_{t}$ ), three regions can be defined in the space of water stages ( $x, y$ ). The three regions are shown in Figure 4.5. Water stage pairs in region 1 do not cause any damages because the levee provides enough protection, and no overtopping n occur. Pairs in region 3 will always produce damages, and the losses will be associated with the failure of either end (but not both) of the levee reach. Water stage pairs in region 2 may or may not produce failure of the levee, and it will not be at the extremes of the reach but somewhere in between. Such failures are called here "internal failures."

A detailed analysis showed that region 2 is a very narrow strip (with average width of .0lm) located along the segments $A B, B C$ (in Figure 4.5). Experience with the model domonstrated the high cost, in terms of computer time, of including region 2 in the evaluation of expected losses. Consequently, it was decided to disregard partially region 2 by computing expected losses over the area to the right and above the sogments $A B$ and $B C$. This simplification introduc d substantial savings in execution time without much loss in accuracy.

In the final implementation, the basic integration ar a was divided into 4 subregions in order to handle the discontinuity of the loss function. Figure 4.6 shows these regions fcr a typical levee design alternative ( $g_{m}, g_{t}$ ). Notice that region 1 corresponds to event El, and regions 2,3 , and 4 correspond to event E2.

As will be explained in Appendix $B$, the program may be run either to determine the total yearly cost (TYC) of individual design alternatives, or to automatically find the optimum alternative.

### 4.4 Determination of the optimum

The problem of model selection has not completely been solved as far as determining a unique best pdf. Nevertheless, the number of candidate pdf's has been substantially reduced. In addition, the composite model seems

to be the most reasonable choice, because it "integrates out" the model uncertainty. The next stage is a sensitivity analysis, to determine up to what extent the most likely models affect the final decision.
4.4.1 pdf's to be Used

The complete model will be run for several combinations of pdf's (f(x,y|E1), $f(x, y \mid E 2))$, and the sensitivity of the optimun decision to the pdf's will be studied. The results of such analysis will give the designer an additional piece of information for the final selection.

In addition to the best ranking pdf's in each group, the model will also be run for other candidate models to illustrate the effect of the hydrologic submodel uper the final result. The pdf's to be used are:

$$
\begin{aligned}
& f(x, y \mid E I): \quad B \cdot B, E \cdot B, B \cdot E, B \cdot L, \\
& f(x, y\{E 2): \quad N . T L, B V N, N \cdot L, N . G .
\end{aligned}
$$

The shape of the marginal distributions, shown in Figures 4.7 through 4.10 , was also taken into account to select the additional models.

### 4.4.2 Preliminary Results

The model was initially run for a few design alternatives using the 16 possible pairs of pdf's $(f(x, y \mid E I), f(x, y \mid E 2))$. The results of this preliminary run showed:





1. The optimum design alternative lies in the region $g_{m}>90$.
2. The variation of the TYC with the change of models for $E l, f(x, y \mid E l)$, decreases as $g_{m}$ increases, and is zero for $g_{m}>91$. This is a consequence of the properties of the marginal pdf's $f_{x}(x \mid E l)$ and the shape of the integration region for the expected losses.

It can be verified in Figure 4.6 that the integration arca of event El (area 1 in figure) decreases as $g_{m}$ increases and is zero for $g_{m}>91.04$. Two more points must also be taken into account: (1) the expected number of yearly occurrences of event E 2 is twice as large as that of event El; and (2) the probability of water stages in the main river, $x$, being greater than 89 m is very small. See Figure 4.7.

As a consequence of the preliminary results, the 16 model pairs were grouped into the following 4 types, according to $f(x, y \mid E 2)$ :

$$
\begin{array}{ll}
\text { Type I: } & f(x, y \mid E 2)=N \cdot G, \\
\text { Type II: } & f(x, y \mid E 2)=B V N, \\
\text { Type III: } & f(x, y \mid E 2)=N \cdot L \\
\text { Type IV: } & f(x, y \mid E 2)=N \cdot T L .
\end{array}
$$

4.4.3 Optimum Design Alternatives

The optimum design alternative was determined for the above 4 types and for the composite model. Figures 4.11 chrough 4.15 show the contours of the TYC surfaces for the different alternative models. Figures 4.16 and 4.17 show the variation of the TYC with the levee height at one end of the reach while the other end is kept constant at its optimum value for each model type.

The optimum levee design alternatives are shown graphically in Figure 4.18 and are listed in Table 4.9.

Table 4.9. Optimum Design Alternatives

| TYpe | f(X,Y\|E2) | $g_{\mathrm{m}}$ | $g_{\mathrm{t}}$ | TYC | Construction <br> cost | Expected <br> losses |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| I | N.G | 91.4 | 94.6 | 15.24 | 132.39 | 3.32 |
| II | BVN | 91.6 | 94.3 | 15.36 | 136.63 | 3.64 |
| III | N.L | 91.4 | 96.6 | 26.03 | 209.41 | 27.33 |
| IV | N.TL | 91.4 | 94.4 | 14.37 | 126.37 | 3.00 |
| Composite |  | 91.5 | 95.0 | 18.46 | 145.26 | 5.38 |



Figure 4.ll. TYC Surface for Models Type I --- TYC in million Forints.


Figure 4.12. TYC Surface for Models Type II -- TYC in million Forints.


Figure 4.13. TYC Surface for Models Type III -- TYC in million Forints.


Figure 4.14. TYC Surface for Models Type IV -- TYC in million Forints.


Figure 4.15. TYC Surface for Composite Model -- TYC in million Forints.


Figure 4.16. Variation of TYC for $g_{t}$ Constant at Its Optimum Value for Each Model

Figure 4.17. Variation of TYC for $g_{m}$ Constant at Its Optimum Value for Each Model


Figure 4.18. Optimum Design Alternatives for All Models

### 4.4.4 Discussion of Results

The first fact noticed in Figure 4.18 is that the optimum for the lognormal model is significantly deviated from the rest of the models. Figures 4.11 through 4.15 also show that the TYC for the N.L model is consistently higher than for the other models.

This deviation can only be explained in terms of the marginal distributions $f_{Y}(y \mid E 2)$ (see Figure 4.10). The lognormal distribution, in this case, has the longest tail among the pdf's considered. Taking, for example, the gamma distribution, which appears to be close to the lognormal in Figure 4.10 , it can be verified that the probability of $y$ being grater than 95.0 is approximately .0006 for the gamma and .0118 for the lognormal. That is, about twenty times larger. When these small quantities are multiplied by the damages, generally above 1000 cost units, the differences in the expected losses are within the range of variation observed.

At a first glance, it might be expected that the differences would disappear as the levee height at the tributary $\left(g_{t}\right)$ tends to 91.04; that is, when the area under all distributions $f_{y}(y \mid E 2)$ is 1.0 . However, this does not occur, because for levee heights $g_{t}<93.7$ (the existing levee height at the tributary) the losses due to event E2 become constant. Figure 4.19 illustrates a typical case


Figure 4.19. Variation of Damages Due to E2 with Change in $g_{t}$ for 4 Models $f_{y}(y \mid E 2)$
of the variation of flood damages due to event $E 2$, with the increase of $g_{t}$, for the 4 marginals considered.

The results obtained for the lognormal models illustrate very well the effects of the tail of a distribution upon decisions based on extreme events. For the model choice criteria used here, the models with lognormal marginal in $y$ ranked poorly enough not to be selected as the "best" model; and, even in the composite model, its influence is greatly reduced by a low weight. If a more simplistic model choice procedure had been used, it is probable that the lognormal marginal had been chosen because its parameters are easy to estimate and the evaluation of the function itself is easy, too. On the other hand, results such as obtained for the lognormal not always should be regarded as undesirable. For example, if a minimax decision procedure were used, this result would be worth much attention.

It is also worth noticing the steep descent in the TYC with the increase of $g_{m}$. This suggests (under the assumptions of the model) that the existing levee does not provide enough protection toward the end at the main river, and that any improvement must be first done toward that end of the levee reach.

The construction cost surface (shown in Figure 4.20) in conjunction with the TYC surface, can be used to determine the most cost effective stages of development in


Figure 4.20. Construction Cost Surface (Million Forints)
case the project is not accomplished all at once. The trajectory $a, b, \ldots, h$ in Figure 4.15 (the composite model) shows the development program with successive investments of 20 million Forints.

Figure 4.21 shows the optimum levee profile for the composite model, along with the existing levee profile. Notice that the vertical scale is extremely exaggerated. One unit length in the vertical scale represents 1.0 m whereas in the horizontal represents about 1200 m .

Clearly this result does not imply that the existing levee has been poorly designed. Assumptions have been made along the way, in part with respect to the loss functions that might have exaggerated the necessity for reinforcement of the levee. Other difficulties in the procedure are discussed in detail in the next chapter.


Figure 4.21. Existing Levee and Optimum Levee for the Composite Model

## CHAPTER 5

## DISCUSSION AND CONCLUSIONS

### 5.1 Discussion

Some points have been raised along the development of this work. They will now be discussed in relation to each submodel; however, since the hydrologic submodel occupies most of this thesis, it will be divided here as: methodology, pdf fitting, and model choice.

### 5.1.1 Hydraulics

The gaging station on the Tisza River is located downstream of its confluence with the Zagyva River. The fact that the readings for the main river are directly affected by the tributary was not taken into consideration in the implementation of the model for the Zagyva River. The effects of this assumption upon the final results were not explored in the present work.

### 5.1.2 Economics

The accuracy of the economic information has not been questioned here, especially in relation to the flood losses. Quadratic loss functions are commonly used in theoretical works but this does not imply that they are of common occurrence in the real world. The loss functions
in this example probably exaggerated the flood damages. For example, the assumption that the damages become approximately constant beyond a certain flood level, and its implementation by the use of a square root function is reasonable. However, the implicit assumption made here, that this "leveling" of the losses occurs after the quadratic function reaches its maximum is not necessarily true.

### 5.1.3 Methodology

The use of water stages instead of discharges presented some problems for the distribution fitting and may have affected the final decision. The transformation operated upon the random discharges by the shape of the channel and the flood plain adversely affects the resulting behavior of the random stages. The cross sectional area increases more than linearly as the water rises. Hence, even if the pdf of the discharges has a long tail, the pdf of the stages will have a shorter tail (if any). The fitting of exponential type distributions to water stages may, consequently, erroneously increase the expected damages. It seems that a better approach would be, when possible, to obtain the distribution of the discharges, and then use the stage/discharge relationship to determine the water levels.

The properties of the integration area for the evaluation of expected damages, discussed in Section 4.3, open the possibility of a further simplification of the model, that surely will improve the manageability of the whole problem. It is based on the assumption that the levee follows a water profile, and it consists in the consideration of levee failures at the extremes of the reach only. This makes the expected losses dependent only upon the damage functions of the two extreme cross sections and eliminates the need to generate water profiles for each possible pair of water stages.

### 5.1.4 pdf Fittings

One major difficulty at the beginning of this thesis was trying to estimate a single bivariate pdf from an "Lshaped" tail. This proved to be a difficult estimation problem; and only very simple cases (linear truncations, univariate cases, etc.) were found in the literature consulted. The division of the sample space and the use of conditional distributions is, however, a valid approach too, and it completely satisfies the needs of the problem treated in this thesis.

The location parameters of some distributions were assumed to be known, in order to simplify the parameter estimation in the example of the Zagyva River. The use of the base levels, $x_{L}=85.95$ and $y_{L}=91.04$, as the location
parameters for the lognormal, exponential, gamma, and beta distributions seems rather reasonable. Other parameters, such as the location parameter for the truncated lognormal distribution and the "b" parameter for the beta distribution, were fixed in a more arbitrary fashion, after a subjective judgment of the sample.

### 5.1.5 Model Selection

Although the model selection procedures used here were relatively successful in reducing the choice set, the discrepancy between the two methods suggests the need for additional study (for example, using simulation) of the performance and accuracy of such methods. It also suggests the need for extreme caution and awareness of the possibie pitfalls in their application.

The deviation of the optimal decision and the inflated total yearly costs obtained for the lognormal model in the case of the Zagyva River clearly illustrate the importance of model selection for decision making based on extreme hydrologic events.

### 5.2 Conclusions

The following points can be concluded from the present work:

1. When expected damages of extreme events are sought, the problem of estimating a bivariate pdf from an "L-shaped" truncated sample may be overcome by
proper subdivision of the sample space and the use of conditional pdf's.
2. Using standard goodness of fit tests, the distributions of different families can be fitted to the joint flood magnitudes in the Tisza and the Zagyva Rivers.
3. None of the marginal distributions considered can be rejected at the $5 \%$ confidence level by the Kolmogorov-Smirnov test.
4. The two model selection procedures presented here can be successfully applied to reduce the set of candidate pdf's. However, the ordering of the distributions is different for the two criteria.
5. Model selection by sample likelihoods is very sensitive to numerical errors in the normalizing constant of a distribution.
6. Under the likelihood of the Chi-square statistic criterion, the best ranking distribution of water stages are: the Beta.Beta for the conditional on event El; and the Normal.Truncated-Lognormal for the conditional on event E2. Under the sample likelihoods criterion, the best ranking distributions are: the Beta.Lognormal for the conditional on event El; and the Normal.Gamma for the conditional on event E2. The composite model is a weighted sum of the Beta.Lognormal, Beta.Gamma,

Beta.Beta, and Bivariate Lognormal distributions for the joint pdf of water stages conditional upon event El; and a weighted sum of the Normal.Lognormal, Normal.Gamma, Bivariate Normal, and Normal.TruncatedLognormal distributions, for the joint pdf of flood magnitudes conditional upon event E2.
7. For decisions based on expected values, the composite model represents a very good alternative, because it takes into account the model uncertainty. In the case of flood protection for the Zagyva River, the decision rached with the composite model lies within the range of variation of those obtained with the component models.
8. The discontinuity in the integration for the evaluation of the expected losses plays a critical role in the speed and accuracy of the results, and it should receive special attention in similar problems.
9. Using the composite model, the optimum levee design has a height (over the sea level) of 95.0 m at the tributary, and 91.5 m at the main river. Its initial construction cost has been computed as 145.26 million Forints and has an expected total yearly cost of 18.46 million Forints. (One Forint is approximately $\$ 0.05$ U.S.)
10. According to the model formulated here, the total yearly cost of the levee system for the Zagyva River
is sensitive to the type of distribution used for the joint probability of flood stages conditional upon event E 2 , and is not sensitive to the distribution of flood stages given event El.
11. Under the economic assumptions of the case study, the results of all the models considered show high investment returns for the improvement of the levee toward the downstream end of the reach.

## APPENDIX A

DATA USED IN THE CASE STUDY

Table A.l. Regression Coefficients, Draw Down


Table A. 2 Regression Coefficients, Damping


Table A. 3. Construction Cost Coefficients


Table A.4. Damage Coefficients

| K | A (K) | B(K) | C(k) |
| :---: | :---: | :---: | :---: |
| 1 | -16.000 | 2944.000 | -132605.000 |
| 2 | -5.600 | 103.5.760 | -45066.224 |
| 3 | -4.800 | 890.160 | - 38436.968 |
| 4 | -2.400 | 453.040 | -18529.696 |
| 5 | -2.000 | 377.400 | -14948.720 |
| 6 | - 2.000 | 376.200 | - 14835.680 |
| 7 | -2.000 | 373.400 | -14573.320 |
| 8 | -2.000 | 371.400 | -14387.120 |
| 9 | -2.000 | 373.000 | - 14536.000 |
| 10 | -2.000 | 370.200 | - 14275.880 |
| 11 | - 2.000 -2.000 | 389.400 386.700 | -15943.970 -15633.680 |
| 13 | -2.000 | 386.200 | - 15533.680 |
| 14 | -2.000 | 386.600 | - 15442.320 |
| 15. | -2.000 | 386.600 | -10442.320 |
| 16 | - 2.000 | 386.600 | -16442. 220 |
| 17 | -2.000 | 386.600 | -15442.320 |
| 18 19 | -2.000 -2.000 | 388.600 388.600 | - 16636.120 |
| 20 | -2.000 | 395.000 | - 17263.000 |
| 21 | -2.000 | 395.800 | -17342.080 |
| 22 | -2.000 | 396.600 | -17421.320 |
| 23 | -12.000 | 2234.400 | -101849.320 |
| 24 | -2.000 | 399.400 | -17699.920 |
| 25 | -2.000 | 400.200 | -17779.880 |
| 26 | - 2.000 | 401.000 | - 17860.000 |
| 27 28 | - 2.000 -2.000 | 403.000 403.000 | -18061:000 |
| 28 29 | - 2.000 | 403.000 403.400 | -18061.000 |
| 30 | -2.000 | 403.800 | -18531.680 |
| 31 | -2.000 | 404.600 | -1.8612.520 |
| 32 | -?.000 | 405.400 | -18693.52C |
| 33 | -2.000 | 405.400 | -18693.520 |
| 34 | -2.000 | 406.200 | -18774.680 |
| 35 | -2.000 | 407.000 | -18856.000 |
| 36 | -2.000 | 407.000 | -18856.000 |
| 37 | -2.000 | 408.600 | -19019.120 |
| 38 | - 2.000 | 408.600 | - 19019.120 |
| 39 | -2.000 | 409.000 | -19060.000 |
| 40 | -2.000 | 409.400 | -19100.920 |
| 41 | -2.000 | 411.000 | -19265.000 |
| 42 | -2.000 | 411.000 | -19265.000 |
| 43 | - 2.000 | 413.000 | - 19471.000 |
| 44 | - 2.000 | 415.400 | -19719.520 |
| 46 | -2.000 | 415.400 | -19719.520 |
| 47 | -2.000 | 415.000 | -19678.000 |
| 48 | -2.000 | 416.200 | -19802.680 |
| 49 | -6.222 | 1228.044 | -58838.582 |

Table A.5. Sample of Flood Magnitudes


APPENDIX B

PROGRAM LISTING

PROGRAM LEV3.
THIS IS THE MAIN PROGRAM FOR THE LEVEE DESIGN PROQLEM. LEV3 READS THE INPUT DATA, SETS UP TABLES TO BE USED BY OTHER ROUTINES, CONTROLS PROCESING DF
INDIVIDUAL ALTERNATIVES OR THE MINIMIZATION PROCESS
IT ALSO STOPS EXECUTION AT THE END OF DATA.
CALLS. SLGEN.SLVE
CALLED BY. NONE
IMPORTANT VARIABLES AND COMMON AREAS.
IPARAM/ = CONTAINS PARAMETERS OF DISTRIBUTIONS TO BE USED.
/LMTSI/ = CONTAINS BASIC INTFGRATION LIMITS.
/BLDM/ = CONTAINS COEFFICIENTS OF ADJOINT DAMAGE
/SCHAR/ = CONTAINS SEARCH AREA FOR MINIMIZATION.
/CONSTI/ = CONTAINS CONSTANTS FOR NUMERICAL INTEGRA
BETA(I,J,K) ${ }^{\text {TION ROUTINE. }}=\mathrm{J}-\mathrm{TH}$ CORRE
BETA(I,J,K)=$\quad J-T H$ CORRELATION COEFFICIENT FOR THE
OF CURVE ( $K=1$ DRAW DOWN, $K=2$ DAMPING SURFACE CURVE).
CDLTH $=\triangle A L T I T U D E$ DIFFERENCE BETWEEN GAUGING STATIDN
CDP = STED SIZE FOR MINIMIZATIDN.
CDPMN = MENIMUM CDP
CER $=$ TOLERANCE FQR INTEGRATION R

CIDM(I) CROSSECTION.
C1DM(I), CI2DM(I) = COEFFICIENTS FDR THE ADJOINT
DAMAGE FUNCTION FOR THE I-TH CROSSECTION.
DLTH = COLTH
DMCF $(I, J)=\underset{\text { FUNCTION }}{\text { JOFFICIENT FOR THE I-TH DAMAGE }}$
$D P=C D P$
DPMN = CDPMN
DSGN(I) $=$ LEVEE HEIGHT $A T$ CROSSECTION I FOR A CURRENT OESIGN ALTERNATIVE.
EXLVE(I) $=$ LEVEE HEIGHT $\triangle T$ CROSSECTION I FOR EXISTING LEVEE.
G1 = LEVEE HEIGHT AT THF MAIN RIVER.
G49 = LEVEE HE[GHT $\triangle T$ THE TRIRUTARY END.
IBUGI = CDNTROL FOR INTEGRATION ROUTINE: $K=1$ PRINTS
IDSPL $=$ INTERMFDIATERESULTS, K = 2 DO NOT.
IDSPL $=$ SAMPLE NUMBER, FOR OATA IDENTIFICATION.
ISRCH $=$ CONTRQL\& ISRCH = 0 FOR INDIVIDUAL ALTERNATI
ITRC $=$ CONTROL; ITRC=2 PRINTS INTERMEDIATE RESULTS,
ITYD = DISTRIBUTION CDDE NUMRER FOR EVENT EZ
ITYU = DISTRIBUTION CODE NUMBER FOR EVENT EI
KIXITO $=$ CONTROL USED TO LIMIT THE TOTAL NUMBER OF
$K M A X I=M A X I M U M$ NUMBER DF ITERATIDNS IV EACH
KMAXI $=$ MAXIMUM NUMBER DF ITERATIONS IN EACH
DIMENSION FOR MULSNP. SEE MULSMP.
KNTXSC = NUMBER OF CROSSECTIONS.
NTXSC =KNTXSC.
TYC = TOTAL YEARLY COST.
WPRFL(I) = WATER LEVEL AT CRCSSECTION I.
$X I N T=B A S E$ LEVEL FOR $X$.
XL,XS = INTEGRATION LIMITS.
XMIDM(I) = POINTGEYOND WHICH ADJDINT DAMAGE YINT = BASE LEVEL FOR Y.
YLOYS = INTEGRATION LIMITS.
INPUT/OUTPUT.

| $C$ $C$ $C$ $C$ $C$ $C$ $C$ $C$ $C$ $C$ $C$ $C$ $C$ $C$ $C$ $C=$ | INPUT SEQUENCE. <br> - PROGRAM CONSTANTS, 1 CARD, SEE STATEMENT 102 - 1 <br> - PROGRAM CONTROLS. ITRC=U/2, IBUG=1/0. ISRCH=1/0. <br> - SEARCH AREA. SEE STATEMENT 114 - 1 . <br> - DAMAGE COEFFICIENTS. <br> - CONSTFUCTION CDEFFICIENTS <br> - REGRESSION COEFFICIENTS. FIRST DRAW DOWN, THEN OAMP ING. <br> - INTEGRATION LIMITS. <br> - PARAMETERS OF DISTRIBUTION. FIRST PDF FOR EI. THEN, PDF FOR E2. TO END DATA $A N D$ STOP PROGRAM, PUT TWO BLANK CARDS INSTEAD. <br> - DESIGN ALTERNATIVE OR STARTING DOINT FOR OPTIMIZATION. (LAST ALTERNATIVE, PIUT GI .LT. O. THIS FORCES A BRANCH TO READ A NEW DISTRIBUTION). |
| :---: | :---: |
|  | PROGRAM LEV3(INPUT. OUTPUT) <br> DIMENSION IDPDFU(19), IDPDFD(14) <br> CDMMON BETA $50,3,7), \operatorname{DMCF}(50,3), \operatorname{CTCF}(50,3), \operatorname{EXLVE}(50)$ <br> 1,WPRFL (50), DSGN(50), NTXSC,DLTH, ITRC,IRUG1 <br> COMMDN/PARAM/PU(10), PD(10), ITYU,ITYD,BAR <br> COMMON/LMTSI/XL,YL,XS,YS,XINT, YINT <br> COMMON /BLDM/C1OM(50), XM10M(50),C12DM(50) <br> COMMON/SCHAR/XSCHL, YSCHL, XSCHS, YSCHS <br> COMMON/CONSTI/CER,KMAXI,KIXITO <br> DATAIDDDFU/"LG","BVL","L•B", "EGG", "E, L", "E, B", "DBG", <br>  2"B.G","B.B","G.F"/ <br> DATA IDPDFD/"NTL","BVN","NG","GTL","STN","G.G","BVL", 1"LTN","L.G","N.L","G.L","L.L","LTL","NTN"/ |
| C | READS PROGRAM CONSTANTS |
| 102 | $\begin{aligned} & \text { READ } 102, \text { CDLTH, CDPMN, CDP, CER, XINT, YINT,KIXITO,KMAXI, } \\ & \text { 1KNTXSC } \\ & \text { FORMAT(6F10.0,3I5) } \end{aligned}$ |
|  | $\begin{aligned} & \text { NTXSC }=\text { KNTXSC } \\ & \text { DLTH }=\text { CDLTH } \\ & \text { DPMN }=\text { CDPMN } \end{aligned}$ |
| C | READS PROGRAM CONTROLS |
| 111 | READ 111,ITRC,IBUG1, ISRCH, BAR FORMAT(3I1,F1C.0) |
| C | READS SEARCH AREA |
| 114 | $\begin{aligned} & \text { READ } 114, \quad \times S C H L, Y S C H L, X S C H S, Y S C H S \\ & \text { FORMAT } 4 \text {, } 4 \text {, } 1 \text { ) } \end{aligned}$ |
| C | READ DAMAGE COEFFICIENTS |
| 100 | $\begin{aligned} & \text { DO } 10 \text { I }=1, N T X S C \\ & R E A D 100,(\text { DMCF I, }), J=1,3) \\ & \text { FORMAT } 3 E 2 C .11, F 10,0) \\ & \text { IF(ITRC } 2304,04,10 \end{aligned}$ |
| $\begin{aligned} & 04 \\ & 251 \\ & 10 \end{aligned}$ | PRINT 251, (DMCF (I, J), J=1,3) FORMAT ("DMCF", 3E15.7) CONTINUE |
| C | COMPUTE PARAMETERS OF SECOND DAMAGE EQN. TO BE USED WHEN H(I) .GT. XMIDM(I) FOR $I=1,49=X S E C T I O N$ NUMRER. |
|  | DO 63 I $=1, N T X S C$ |

304 IF(ISRCH) 306,306,307

```
\(\stackrel{C}{C}\)
    ISRCH=O MEANS INDIVIDUAL ALTERNATIVES. PROCEEDS TO
    EVALUATE TYC OF ALTERNATIVE JUST READ.
306 CALL SLVE(G1,G4O,TYC,IWHT)
    GOTO 302
307
C MINIMIZATION PRICEDURE
PZMN, PXMN, PYMN, ARE USED TO SAVE THE LOWEST POINT
    COMPUTED IN AN ITERATION, SONEXT ITFRATION STARTS
    AT SUCH POINT.
    \(P Z M N=1 \cdot E 20\)
\(D P=C D P\)
    ITRBL \(=0\)
C COMPUTES TYC AT THE VERTICES OF A TRIANGLF \(\quad\) O, 1,2
309
352
    \(P X O=G 1\)
\(P Y O=G 49\)
\(P C Y=1\).
\(P C X=1\)
    CALL SLGEN(PXO, PYO, PCX,PCY,DP,PZO,KLOST)
    IF (KLOST) \(351,352,352\)
    CONTINUE
    IF (PZO PZMN) 430,432,432
    CXMNTINUE
    \(P X M N=P X O\)
    PYMN \(=P Y O\)
PZMN \(=P Z O\)
    PZMN = PYO
    \(P \times I=P X O+D P\)
    PY1 \(=\) PYO
    \(P C X=1\).
    \(P C Y=0\).
    CALL SLGEN(PX1, PY1.PCX,PCY,DP, PZ1,KLOST)
    IF(KLDST) \(351,353,353\)
    CONTINUE
    IF (PZ1 - PZMN) 4i0,412,412
    \(P X M N=P X I\)
    PYMN \(=P Y 1\)
PZMN \(=P Z 1\)
    PZMN \(=\) CONTINU
    PXZ \({ }^{\mathrm{CONPO}} \mathrm{PXO}\)
    \(P Y 2=P Y 0+D P\)
    \(P C X=0\).
\(P C Y=1\).
    CALL SĹGEN(PX2, PYZ, DCX,PCY,DP,PZ2,KLOST)
    IF(KLOST) \(351,354,354\)
        CONTINUE
    IF (PZ2 - PZMN ) 420,422,422
    CONTINUE
    \(P \times M N=P \times 2\)
    PYMN \(=\) PY2
PZMN \(=P Z 2\)
    \(P Z M N=P Z 2\)
        CONTINUE
    COMPUTES DIRECTION OF THE MAX SLOPE IN PLANE
    PASSING THRU POINTS \(0,1,2\)
    \(P \times 10=P X I-P X 0\)
    PYIO \(=P Y I-P Y O\)
```




SUBROUTINE CONSTR(CCDST)
COMMON BETA (50, 3,2), DMCF 50,3$)$, CTCF $(50,3)$, EXLVE(50)
1, WPRFL(50), DSGN(50), NTXSC,DLTH, ITRC,IBUGI

IF(DSGN(I) GT. EXLVE(I)) GD TO 15
DSGN(I) = EXLVE(I)

| XCGST |
| :--- |
| GOTO |
| 18 | $\mathrm{O}^{\circ}$

$Z_{Z}^{G O}=\operatorname{TDS}{ }^{18}(I)$
$15 \quad \mathrm{XCOSTGN}=\operatorname{CTCF}(I, 1) * Z * Z+\operatorname{CTCF}(I, 2) * Z+\operatorname{CTCF}(I, 3)$
18 COVTINUE
IF (ITRC - 2)77,77,78
77 PRINT 201,I,DSGN(I), XCOST
201 FORMAT("OSGN", I2.2E15.7)
$\begin{array}{ll}78 & \text { COOST }=\text { COSST }+X C O S T\end{array}$
CONTINUE
RETURN
RETURN
END

SUBROUTINE DAMAGE(HI,H4G, VLDMG)
COMMON RETA $50,3,2)$, DMCF $(50,3)$, CTCF $(50,3)$, EXLVE(50)
1,WPRFL(50), DSGN(50),NTXSC.DLTH,ITRC,IBUGI
COMMON /BLOM/CIDM(50), XM1DM(50),C120M(50)

```
    COMMON/BKTFL/KTFL(49)
    G1= DSGN(1)
    G49 = DSGN(NTXSC)
        FINDS FIRST FAILING X-SECTION LOOKING FROM DOWN UP
    KT(='1
        NEXT 8 STATEMENTS. SEE IF FAILING X-SECTIDN AT THE
        EXTREMMES. IF SO, SKIPS PROFILE GENERATION.
    IF(H49 - - G49)91,91,82
    GO TO 27
    IF(H1-G1)91,91,92
    Z = = H1 %
        FINDS WATER PROFILF CORRFSPONOTNG TD (H1,H49).
        PROFILE RETURNED FROM HYDRA IN WPRFL.
    CALL HYDRA(H1,H49)
    DO 10 KI = 1,NTXSC
    I}=\frac{K}{
M
C NO OVERTOPPING, VLDMG = 0
    VLDMG = 0.
    RETURN
C OVERTOPPING AT XSECT I, CALC VLDMG
25 Z = WORRFL(I)
    KTFL(I)= KTFL(I) + 1
    IF(Z.LE. XMIDM(I)) GO TO 31
    IF ? GREATER THAN XMIDM(I), SQUARE RGOT DAMAGE
    VLDMG = SORT(Z - CIDM(I)) + CI2DM(I)
    CONTINUE
    VLDMG = DMCF(I,I)*Z*Z + DMCF(I,2)*Z + DMCF(I,3)
    RETURN
    END
```



```
            SUBROUTINE HYDRA(H1gH49)
            COMMDN BETA (5C,3,2), DMCF 50,3 ), CTCF 50.3\()\), EXLVE(50)
            1,WPRFL (50), DSGN(50), NTXSC, DLTH, ITRC, IBUGI
C DETERMINE TYPE DF PROFILE TO BE GENERATED
```



```
    DO 10 I \(=1\), NTXSC
    WPRFL(I) = BETA(I, 1,KT) + BETA(I, 2,KT)*HI +
    BETACI
CONTINUE
10 CONTINUE
C INSURE MONCTONICITY
    DO \(20 \mathrm{I}=2, \mathrm{NT} \times S \mathrm{C}\)
    IF(WPRFL(I)-WPRFL(I-1)) 16,17,17
16 WPRFL(I) =WPRFL(I-1.)
17 IF 20 (WPRFL(I) -GT•WPRFL(NTXSC)) WPRFL(I) =WPRFL(NTXSC)
    IF
CONTINULE
RETURN
    RETURN
    END
```



SUBROUTINE SLGEN(PX,PY,PCX,PCY,DP,PZ,KLOST)
COMMON ISCHAR/XL,YL,XS,YS
$K \operatorname{LOST}=0$
$\begin{aligned} X & =P X \\ Y & =P Y\end{aligned}$
IF(X:LT: XL : OR: X :GT: XS) GO TO 335
$\operatorname{IF}(X: G E: Y) G O$ TO 335
CALL SLVE (X,Y,TYC,IWHT)
IF (IWHT - 1)10,20,20
$P X=$
$P Y=$
$P Z=$
$P Z=T Y C$
$R E T I J R N$
10
$X=X+P C X * D P$
$Y=Y+P C Y * D P$
$G D T D 05$
KLOST $=-1$
RETURN
END


```
SURROUTINE SLVE. GIVEN A LEVEF DESIGN ALTFRNATIVE ( \(X, Y\) ), SLVE GENE RATES LEVEE PROFILE, COMPUTES CONSTRUCTION COST, EX PECTED LOSSES AND TOTAL YEARLY COST (TYC) VARIABLES.
BAR NDT USED
CCOST CCNSTRUCTION COST (INITIAL)
COSTD CCNSTRUCTION COST DISCOJNTED (YEARLY)
COSTD CCNSTRUCTION COST DISCQUNTED (YEARLY)
DSGN(I) LEVEE HEIGHT AT CROSSECTION I FOR PROPOSED
LEVEE (X,Y) Y
ER \(=\) CER SEE MULSMD
EXLVE FXISTINGLEVEE
IBUG \(=\) IFUGI SEE MULSMP
ITYD, ITYU DISTRIBUTION CODE NUMBERS SEE LEVB
\(=0\) INTEGRAL DIDNOT CINNVERGE
IXIT = IXITO SEELEV3, MULSMP
IXIT14 \(K\) USED TO PRINT IXIT OF EACH REGICN.
KXITO = IXITO
KTFL(I) USED TO PRINT THF NUMBFF OF TIMES
CROSSECTION I FAILEO IN THE PRICESS
KXIT INCICATOR PRINTED. \(=0\) IF ALL REGIONS CONVERGE
MAXI SEE MULSMP
ND SEE MULSNP
PD, PU PARAMETERS OF DISTRIBUTJINS IN USE. PFOR EZ
AND EI RESPECTIVELY)
TYC TOTAL YEARLY COST
VALUE SEE MULSMP
VLI... 3 VALUEFNR REGIONS 1...3. 4 AND 5 UNUSED.
WPRFL WATER PRDFILF.
XINT,YINT SEE LEV3
XLZ,XSZYLZYS2 ACTUAL INTEGRATINN LIMITS
INPUT INFO. ( \(X, Y\) ) LEVEE ALTERNATIVE, FROM LEV3 OR S
SLGEN, VJA ARGUMENTS
CONSTII CONSTANTS FOR MULSMP
WPRFL FRIMM HYDRA, VIA CQMMON
CCOST FRDM CONSTR
DSGN. MODIFIED RY CONSTR, VTA COMMON
VLI... 3 FRGM MULSMP, VIA ARGUMENTS (VALUE). ARE
```



```
    SUBROUTINE SLVE(X,Y,TYC,IWHT)
    FXTERNAL ALIM1,ALIM3,BLIMI.BLIM2,FINTI,FINT2,FINT3
    DIMENSION IXITI4(5)
    COMMON BETA (50,3,2), DMCF(50,3), CTCF(50,3), EXLVF(50)
    1,WPRFL(50), OSGN(50),NTXSC,DLTH,ITRC,IRUG1
    COMMDN/PARAM/PU(1O),PD(1O), ITYU,ITYD,BAR
    COMMON/LMTS1/XLI,YLI,XS1,YSI,XINT,YINT
    COMMON/LMTS2/YL2,YL2,XS2,YS2
    COMMON/EKTFL/KTFL(4O)
    COMMON/CONSTI/CER,KMAXI,KIXITO
    DATA IFLAG/C/,ITSK,KBLAN/"***"," "/
C FIRST TIME CALLED SETS UP PAPAMETERS FOR MULSMP
    IF(IFLAG)1C,10,20
10 IFLAG= =1
    ND=2
    IXITO = KIXITO
    MAXI=KMAXI
    ER = CER
    IBUG = IBUGI
C UPPER INTEGRATIJN LIMIT FIJR BETAPDF IN X IN EI
    XSBETA = 8E.499
20
    CONTINUE
    GENERATES WATER PROFILE CORRESPONDING TE PAIR (X,Y)
    CALL HYORA(X,Y)
C MOVES PROFILE TO DESIGN.
30 DOGSOII = 1,NTXSC
C
    OBTAINS FINAL DESIGN AND CDNSTRUCTION COST
    CALL CONSTR (CCOST)
    CALL SECOND(TA)
        OBTAINS EXPECTED LOSSES
        VLI=0.
    IXIT14(4)=1
C INTEGRATION OVER REGION I. EVENT I, DISC ON XL
    XL2 = x
    XS2 = XS1
```

YL2 $=Y L 11$
c
$c$

CHECK IF BETA OISTR. IS USED IN X VARIABLEFOR EL
 UPPERLIMIT IS XSBETA.
continue
VL1 $=10$

$\times 52=$ XSBETA
CONTINUFO
CALL

continue.
integration quer region 2. event 2, disc on xl

YLL $=$ YINT +.01
$\begin{aligned} & \text { YS2 } \\ & \text { IXIT }\end{aligned}={ }^{\text {Y }}$ IXITO
CALL
1MULSMP(ND,ALIM1,BLIM1,MAXI,ER,FINT2,VL2,IXIT,IBUG)
IXIT14(2) = IXIT
integration over regions 3,4.
$\mathrm{XL2}$ CONTINUE
integration dVer region 3 . Disc on yl

YS2 $=Y \leq 1$.
IXIT = IXITO

integration over region 4
XL2 $=$ EXLVE(1)
YL2 $=$ YL
YS
XS
YS
YS
YS
YS2 $=$ Y YSIT $=$ IXITO
CALL
1MULSMP(ND,ALIM1,BLIM1,MAXI,ER,FINT2,VL4,IXIT,IBUG)
IXIT14(4) = IXIT
KXIT $=\operatorname{IXIT14(1)*IXIT14(2)*IXIT14(3)*IXIT14(4)~}$
VALUE $=.583333333333 * V L 1+1 \cdot 30555555556 *(V L 2+V L 3+V L 4)$

CALL SECONO(TB)
$\cos T D^{T B}=.09 * C \cos T$

NOTE. . 09 IS DISCOUNT FACTOR.
totex = value + costd
TYC = TOTEX
IO = ITSK
IWHT $\left.\operatorname{IW}=\mathrm{T}^{0}-1\right) 41,42,41$
42
IO $=K B L A N$
IWHT = 1
41
201
PRINT 201, $X, Y,(K T F L(J), J=1,49)$
FORMAT ("OGI=", F7:3," G4a=",F7.3," FAIL=",49I1)
PRINT 202,CCOST, COSTD,VALUE,TOTEX,TT
FORMAT(" CC=", E12•S," YC=", E12•6," EL=", E12.6," TYC=",
1
E12.6," TIME=", F7.3," SEC")
PRINT 203,VLI, VL2,VL3,VL4,VL5,IXIT14,IO
203 FORMAT("VI=", E12.6," V2=", E12.6," V3=",F12.6," V4 =", 1 RET


```
    O DQIL L=I,I
    K(J) =1
    A(J) = ALIM(J,XNEW)
    1 XNEW(J)=A(J)
    2 I=1
    FA(1)=FINT(XNEW)
    B(1)=BLIM(I,XNEW)
    XNEW(1)=R(1)
    FB(I)=FINT(XNEW)
C
    1000 XH(I)=B(I)-A(I)
    IF (.25 * XH(I) ,EQ&OO)
    XIR(I)=0.5*XH(I)
    XHA(I)=XIR(I)/3.
    XJ(I)=XIR(I) %(FA(I)+FB(I))
    INDEX(I)=0
    XNEW(I)=A(I)+XIR(I)
C}1006 IF (I .FO. 1) GOTO 14
    15 I=I-I TO
    14 FNEWX(1)=FINT (XNEN)
C1001 IF (INDFX(I) NGGTO 0) GO TO 1003
    KOUNT = KRUNT, + I +
    YI(I) = XHA(I) *(FR(I) +FA(I) +4.*FNEWX(I))
    1004 XJ(I) =0.25*(XJ(I)+3.*XI(I))
    INDEX(I)=INDEX(I)+1
    KOUNT = KOUNT + I
    IF(KOUNT, GT. MXCNT) GO TO T776
    1010 XH(I)=0.5*XH(I)
    IF (:5 * XH(I), EQQ O;) GO TO }
    XNEW(I)=0.5*XH(I)+A(I)
    1005 IF (XNEW(I) .LT. B(I)) GD TO 1000
```



```
    77 IF(I BUG &EQ., 1)PRINT 2OOO,I,XIP
    XE(I)=ABS (E*XIP(I);
    IF (ABS(XIP(I) - XI(I)) &LE. XE(I)) GO TO 1009
    1008 XI(I)=XIP(I)
    GO TD 1004
    1003 S(I)=FNEWX(I)+S(I)
    \NEW(I)=XNEW(I)+XH(I)
C
c 4XIP(I) = 0.
    1009 IF (I .EQ.N) GO TO 16
    17 II=I+1
        J=K(I1)
    T0 (11,12,13),J
C
    11 FA(I1)=XIP(I)
    K(II)=2
    B(II) = BLIM(II,XNEW)
    GOTO TO 3
C
    12 FB(II)=XIP(I)
```

C

13 FNEWX(II)=XIP(I) $I=I 1$
16 VALUE $=X I P(N)$
1014 RETURN
$C$
$C$
$C$
1011 PRINT 1012, I, INDEX(I)
1012 FORMAT (4HOTHE, I3, $34 H$ INTEGRAL HAS NOT CONVERGED AFTER
C $\quad$, I 3,11 HITERATII
$\mathrm{c} \quad \mathrm{VALUE}=\mathrm{XIP}(N)$
${ }_{\text {GO }}{ }^{\text {IX TO }}={ }^{2}{ }^{2} 14$
C

61 10X2HB(I2,2H)=E20.8)
VALUE $=0$.

C



FUNCTION FINTI(Z)
DIMENSION Z(2)
COMMON /PARAM/PU(10), PD(10), ITYU,ITYD
$x=2(1)=$
$y=2(1)$

CALL DAMAGE (X,Y,VDMG)
GO TO $1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18.19)$
1,ITYU

```
            OLNTINUE
\(=\) OLGN(X,PU(1),PU(2),PU(3))
```

    \(\begin{aligned} T 1 & =U L G N(X, P U(1), P U(2), P U(3)) \\ T 2 & =U G A M(Y, P U(4), P U(5), P U(B), P U(7))\end{aligned}\)
    \(T 2=U G A M(Y, P U(4), P U(5), P\)
    $F I N T 1$
$G O V D M G T 1 * T 2 / D U(10)$
GO TO 90
02 BVLDG
CONTINUE
FINT1 = $\operatorname{BVLGN(X,Y,PU(1),PU(2),DU(3),PU(4),PU(5),PU(6),~}$
1 PU(T))/PU(10)*VDMG
OTO 99
OUTN ${ }^{R}$
T1 $=$ CONTINUE $U L G N(X, P U(1), P U(2), P U(3))$
T2 $=U B E T A(Y, P U(4), P U(5), P(J(6), P U(7), P U(8))$
FINT1 $=$ VDMG*TI*TZ/PU(10)
$\begin{array}{ll}C & \text { COLTEG } \\ 04 & \text { CONTINE }\end{array}$
$T 1=U E X P(X, P U(1), P U(2))$
T2 $=$ UGAM (Y,PU(3):PU(4),PU(5).PU(6))
FINT1 $=$ VOMG*TI*T2/PU(10)
GOTD 99
$C$
05
COSTENUE
$T 1=U E X P(X, P U(1), O U(2))$
$T 2=U G G N(Y, P U(3), P U(4), P U(5))$
FINT1 = VDHE*T1*T2/PU(10)
GOTO 99
CONTE日BE
$\begin{aligned} T 1 & =U E X P(X, P U(1), P U(2)), P U(5), P U(6), P U(7))\end{aligned}$
$T 2=U B E T A(Y, P U(3) P P U(4)$,
$F I N T 1=V D H G * T 1 * T 2 / P U(10)$
GOTO TO DOURG
$C$
07

| C |
| :--- |
| O |

    FINTI \(=\underset{R V G M}{\operatorname{CONT}}(x, Y, P U(1), P U(2), P U(3), P U(4), P U(5), P U(6)\).
    1 PU(7),PU(8))/PU(10)*VDMG
    GO TO 08 G.L
        CONTINUE
    $=\operatorname{UGAM}(X, P U(1), P U(2), P U(3), D U(4))$
$T 1=U G A M(X, P U(1), P U(2), P U(3) ;$
T2 $=U L G N(Y \cdot P U(5) ; P U(6) ; P$
$F I N T 1=V D G * T 1 * T 2 / P U(10)$
GO TO ${ }_{0} 99$.
CONTINUE
$T 1=U G A M(X, P U(1), P U(2), P U(3), P U(4))$
$T 2=U B E T A(Y, P U(5), P U(6), P U(7), P U(8), P U(0))$
FINT1 = VOME*TI*T2/PIJ(10)
GO TO 99
$\begin{array}{ll}\text { C } & \\ 10 & \text { CONTNNUE }\end{array}$
$T 1=U N O R(X, P U(1), P U(2))$
$T 2=U N O R(Y, P U(3), P U(4))$
FINTI = VDMG*TI*TZ/PU(1O)
$\begin{array}{ll}\text { C GO TO } \\ \text { C } & \text { IIGGG } \\ 11 & \text { CONTINUF }\end{array}$
$T 1=U G A M(X, P U(1), P U(2), P U(3): P U(4))$
$T 2=U G A M(Y, P U(5), P U(6), P U(7), P U(3))$

```
    FINTI = VOME*T1*T2/PU(10)
    GO TO I2, L.L
    CONTINUE
        T1 = ULGN(X,PU(1),PU(2),PU(3))
        T2 = ULGN(Y,PU(4),PU(5),PU(6))
        FINT1 = VDMG*T1*TZ/PU(10)
        GO TO 99, L.E
        CONTINUE
        T1 = ULGN(X,PU(1),PU(2),PU(3))
        T2 = UEXP(Y,PU(4),DU(5))
        FINT1 = VDMG*T1*T2/PU(10)
        GOTO 99
        CONTINEEE
    T1 = UEXP(X,PU(1),OU(2))
    FINTI = VDMG*TI*T2/PU(10)
    GO TO 99
C CONTNNUL
    T1 = UBETA(x,PU(1),PU(2),PU(3),PU(4),PU(5))
    T2 = ULGN(Y,PU(t),PU(7),PU(8))
    FINT1 = VDMG*T1*T2/PU(10)
    GO TO 99. (16 B.E
    CONTINUE
    T1 = UBETA(X,PU(1),PU(2),PU(3),PU(4),PU(5))
    T2 = UEXP(Y:PU(6),PU(7))
    FINT1 = VDNG*TI*T?/PU(10)
    CONTBNG
    T1 = URETA(X,PU(1),PU(2),PU(3),PU(4),DU(5))
    T2 = UGAM(Y,PU(6),PU(7),PU(8),DU(9))
    FINT1 = VDMG*TI*T2/PU(IO)
    GO TO 99
        13 R.B
        T1 = UBETA(x,PU(1),PU(2),PU(3),PU(4),PU(5))
        T2 = URETA(Y,PU(6),PU(7),PU(8),PU(9),PU(10))
        FINT1 = VDNE*T1*T2
        GO TO 199.E.E
        CONTINUE
        T1=UGAM(X,PU(1),PU(2),PU(3),PU(4))
        T2 = UEXP(Y,DU(5),PU(G))
        FINT1 = VOMG*T1*T2/PU(10)
        GOTO 99
    9 9
    RETURN
```



FUNCTION FINT？（Z）
DIMENSION Z（2）
COMMON／PARAM／PU（10），PD（IO），ITYU，ITYD
C USESPD，ITYD
$X=Z(1)$
$Y=Z(2)$
CALL DAMAGE（X，Y，VDMG）
GO TO（01，02，03，04，05，06，07，08，09，10，11，12，13，14），ITYD 01 NTL
CONTINUE
$T 1=\operatorname{UNDR}(X, P D(1), P D(2))$
T2 $=$ ULGN（Y，PD（3），PD（4），PD（5））
FINT？＝VDNGねTまれT2／PD（IO）
02 BVNOR
CONTINUE
FINT2＝VDMC＊BVNML（X，Y，PD（1），PD（2），PD（3），PD（4），PC（5））

03 N．G
$T 1=U N D R(X, P D(1), P D(2))$
$T 2=\operatorname{UGAM}(Y, P D(3), P D(4), P D(5), P D(6))$
FINT2 $=$ GDNG＊T1＊T2／PD（10）
CONTINL GUE
$T 1=U G A M(X, P D(1), P D(2), P D(3), P D(4))$
$T 2=U L G N(Y, P D(5), P D(6), P D(7))$
FINT2 $=$ YDMG＊T1＊T2／PD（10）
GO T0 99
05 G．TN
$\mathrm{C}_{\mathrm{C}}^{\mathrm{C}}$
CDivTINUE
$T 1=U G A M(X, P D(1), P D(2), P D(3), P D(4))$
T2 $\equiv \operatorname{UNOR(Y,PD(5),PD(6))~}$
FINT？$=$ VDMG＊T1＊T2／PD（10）
GO 1090
${ }_{\mathrm{C}}^{\mathrm{C}} \mathrm{C}$
06 G．G
$T 1=U G A M(X, P D(1), P D(2), P D(3), D D(4))$
$T 2=U G A M(Y, P D(5), P D(6), P D(7), P D(8))$ FINT2 $=$ VDNE＊T1＊T2／PD（10）
GOTO 90

## $\begin{array}{ll}\text { C } & \text { O7 BVLOG } \\ 07 & \text { CONTINUF }\end{array}$

FINT2 $=B V L G N(x, Y, P D(1), P D(2), P D(3), P D(4), P D(5), P D(6)$,
1 PD（7））／PD（10）＊VOMG
$\begin{array}{ll}\text { C } & 08 \text { LITN } \\ 08 & \text { CONTINUE }\end{array}$
$T 1=U L G N(X, P D(1), P D(?), P D(3))$
T2＝JNOR（Y，PD（4），DD（5））
FINT2 $=$ VDMG＊T1＊T21PD（10）

| $\begin{aligned} & \mathrm{C} \\ & 09 \end{aligned}$ | ```GO T0 99 03 L.G CDNTINUE``` |
| :---: | :---: |
|  |  |
|  | FINT? = VONG* TI*T2/PD(10) |
| ${ }^{C} 10$ | GO 10.09 |
|  | CONTINUF |
|  | $T 1=U N O R(X, P D(1), P D(2))$ |
|  | T2 = ULGN(Y. DO (3):DO(4), PD(5)) |
|  | FINT2 $=$ VDMG*TI*T2/PD(10) |
|  | G0 T0 99 |
| ${ }^{\text {C }} 11$ | $11 \mathrm{G} \cdot \mathrm{L}$ |
|  | CONTINUE |
|  | $T 1=U G A M(X, O D(1), P D(2), P D(3) \cap P D(4))$ |
|  | T2 $=$ ULGN(Y, OD(5), PD(5), PD(7)) |
|  | FINT2 = VDMG*T $*$ (2/PD(10) |
|  | GO TO 90 |
| C12 | 12 L . L |
|  | continue |
|  | $T 1=U L G N(Y, P D(1), P D(2), P D(3))$ |
|  | $T 2=U L G N(Y, P D(4) \cdot P D(5), P D(6))$ |
|  | FINT2 = VOMG*T1*T2/PD(10) |
|  | G0 T0 99 |
| C13 | 13 L.TL |
|  | CONTINUF |
|  | $T 1=U L G N(x, P D(1), O D(2) \cdot P D(3))$ |
|  | $T 2=U L G N(Y, P D(4), P D(5), P D(6))$ |
|  | FINT2 = VOMG*T1*T2/DD(10) |
|  | G0 T0 90 |
| C | 14 N.TN |
| 14 | Continue |
|  | T1 = UNQR (X, DD(1), PD(2)) |
|  | T2 $=\operatorname{UNGR(Y,DD(3),~PD(4))~}$ |
|  | FINT2 = VONG*T1\%T2/PD(10) |
| 99 | RETURN |
|  | END |



FUNCTIDN FINTB(Z)
DIMENSION Z(2)
C USES PD,ITYD
$X=$ USES PD, ITYD
$Y=Z(2)$
$Y=Z(1)$
$C A L L A M A G E(X, Y, V D M G)$
GO TOI O1, 02,03,04,05,06,07,08,09,10,11,12,13,14), ITYD
CINTINLE
$T 1=\operatorname{UNOR}(X, P D(1), P D(2))$
T2 $=$ ULGN(Y, PD (2), PD ( 4 ), PD(5)
FINT2 = VOMG*T1*T2/PD(10)



## 80 IF(ALIMI - YY) $81,81,82$ <br> 82 81 $A L I M 1=Y Y$ RETURN END



FUNCTION BLIMI (I, X)
DIMENSION $x(2)$
COMMCN /LMTS2/XLZ,YL2,XS2,YS2
IF (I - 1 ) C5,10,20
05 STGP 323
20 ELIMI = YS2
RETURN
10 BLIM1 $=X S 2$
IF (BLIM! - $x(2)) 30,30,32$
32 BLIMI $=x(2)$
30
RETURN
END


FUNCTIDN ALIMZ (I $x$ )
GIVFS LCWER LIMITS FOR INTEGRATION OVER REGION 3
ALIM3, BLIM3, FINT3. X(1) $=\mathrm{Y}, \mathrm{X}(2)=X$
DIMENSION $x(2)$
COMMDN/LMTS?/XL2,YL2,XS2,YS2
IF (I -1 ) $05,10,20$
05
20
STOP 332
ALIM3
RETURN
${ }^{D G}=M^{16}=$
$\left.\begin{array}{l}\text { EPSLIM } \\ X X Y \\ X \\ X \\ Y\end{array}\right) .01$
$Y Y=Y L ?+.001$
ISWCH = $=1$
CALL OAMAGE $X Y, Y Y$
23 CALLOAMAGE(XY,YY
25
$I S W C H=0$
$Y Y=Y Y+D C$

```
30 IF (ISWCH) \(32,32,33\)
33
    ALIM \(3=Y Y\)
```



```
    \(Y O=Y Y-D E\)
    \(\left.\left.{ }_{A L I M B S}^{\text {IF }}=Y P-Y O\right)-E P S L I M\right) 37,37,38\)
    GOTO 90
    \(Y Y=.5 *(Y P+Y O)\)
    CALL DAMAGE (XX,YY,VDMG)
    IF(VDMG)41,41,42
\(41 \quad\) YO \(=Y Y\)
    \(\begin{array}{ll}42 & G O \text { TO } 35 \\ & \text { GP }=Y Y \\ & G O T O\end{array}\)
    \(\begin{array}{l}80 \\ 82\end{array}\) IF (ALIMB \(\quad\) ALIM3 \(\left.=X X\right) 82,81,81\)
82
82
81
32
35
37
38
41
    IF (ALIM3 \(\overline{\text { M }} \bar{X} X X) 82,81,81\)
    RETURN
    END
```


FUNCTION BLIM3(I, $X$ )
DIMENSIDN $\times(2)$
COMMON/LMTS?IXLZ,YL?,XSZ,YSZ
IF (I - 1)05,10,20
$\begin{array}{ll}05 & \text { STOP } 333 \\ 20 & \text { BLIM3 }=X S 2\end{array}$
$20 \quad$ BLIM3 $=X S 2$
$\begin{aligned} & \text { RETURN } \\ & \text { BLIM3 }\end{aligned}=Y .52$
IF(BLIM3 - X (2) ) $32,30,30$
IF (BLIM3
RLIM3
$=12)$
$\begin{aligned} & \text { RLIM3 } \\ & \text { RETURN }\end{aligned}=X(2)$
END


RETURN
END


[^0]UBETA $=((X-A) * *(P-1)) *.((R-X) * *(Q-1).) / D E N$ RETURN
END


FUNCTION UNDR $(X, X M, D T)$
$U=(X-X M) / D T$
$U=-.5 * U * U$
IF(U +673.$) 10,10, ? 0$
UNOR $=0$.
20 RETURN
UNQR $=\operatorname{EXP}=(11) /(2.506628275 * D T)$


FUNCTIDN EVLGN(X,Y•XC,XML.DTXL•Pח,YC,YML, DTYL)

$X S H=X-X C$
YSH $=Y-Y C$
IF (XSH) 31,31,32
IF (YSH)31,31,33
PR INT 202, X,Y
FDRMAT("O *** FRDM LDGNGRMAL BV, $(X, Y)=1,2 E 15.7)$
STOP 51
$33 \quad \begin{aligned} X S H L & =A L C G(X S H) \\ Y S H L & =\triangle L O G(Y S H)\end{aligned}$
$21=(X S H L-X M L) / D T X L$
$Z 2=(Y S H L-Y M L) / D T Y L$
$0=-\quad 5 *(21 * Z 1,-2 * Z 1 * Z 2 * R 0+Z 2 * Z 2) / C 1$
IF(Q $+673.121,21,25$
25
$B V L G N=E X P(Q) * C 2 /(X S H * Y S H)$
RETURN
BVLGN $=0$ 。
RETURN
END


```
    FUNCTION BVGM.
    CDMPUTES DDUBLE GAMMA PDF.
    ARGUMENTS.
    (CF. FUNCTIONAL FORM IN CHAPTEP 3)
    CLAM LAMBCA
    CMIJ MU
    A,R,C SAME AS IN DFFINITIDN.
    COEF= I/(CLAM**A*MU**C*G(A)*G(B)*G(C)) WHERE G(.)
    IS THE GAMMA FUNCTION.
    XC,YC LOCATION PARAMETERS DF X AND Y.
    CALLS INTGM.
    CALLEDEY FINTI.FINT3.
```



```
    FUNCTION BVGM(XX,YY,CLAM,C:MU,A,B,C,CDEF,XC,YC)
    COMMON/EDBG/DUMY(E),ERRG,DX:HNG
    COMMON/BINT/XL,YH,AMI,BMI,CMI
    ERRG = .0?
    DXMNG =.0001
    AM1=A - = 10
    BM1 = 员 = 1. 
    X=XX - XC
    XL=X/CLAM
    YM = Y/CMU
    UM = YM
    IF(XL - YM)15,15,25
15 UM = = XLLINUE
IF(UM)89,8S,Q8
89 EVGM=0.
    RETURN
        CONTINUE
    D=-XL-YM
    IF(D+673.1R9.21, 21 
    IF(D + 673.)R9.21, 2l
    IF (IER)22.22, 22
22
    RETURN
    END
```



SUBROUTINE INTGM (XF,RES,IER)
CIMMDN/BDPG/CLAM,C YU,A,B,C,CREF,ERRG,DXMNG
COMMON/BINT/XL,YMOAM1, RMIOCMI
C
$F(Z)=(Z * * A M 1)+((X L-Z) * * O M 1) *((Y M-Z) * * C M 1) * E X P(7$. SHOULD BE SO $=F(0)+.F(X F) B U T B O T H A R E=0$.





```
    FUNCTIDN FINTZ(Z)
    DIMENSTON7(?)
    DATA XM1,DTY1/83.7047.2.64371
    DATA TY1,YM1,DTYi,DIO1/87.00,1.47267,.152888,.65991
    DATA XM2,DTX2,YM2,DTY2,POZ,P1C2/83.6095,2.73228,
    1 91.1749,0.8745,0.09438,0.521461
    DATA AY3,BY3&CY3.BGY3,P103/1.72813,0.387949,01.04.
    l OOG178002,OGG452'
    1 (1010/91.04,-0.71632.0.0.7106,
    DATA W1,W2,W3,W4/.255,.467,.146,.1321
    x = 7(1)
    YAULL(ZAMA
    CALL DAMAGE(X,Y,VDMG)
C BIVARIATE NDRMAL COMPONENT DISTRIRUTITN CDDE = O?
    F1 = BVNML(X,Y,XM2,DTYZ,YMZ,OTYZ,ROR)/P1O2
C NORMAL-GAMMA COMPONENT. DISTRIBUTION CIDE = 03
    T1 = UNOR (X,XM1,OTX1)
    T2 = UGAM(Y,AYZ,BY3,CY3,RGY3)
    F2 = T1*T2/fl03
C NORMAL-LCGNORMAL COMPONENT. DISTRTBUTION CODE = 10
```



```
    F3 = T1*T2/F1010
C NORMAL-TRUNCATED LOGNORMAL COMPONENT. CODF= CI
    TZ = ULGN(Y,TYi,YMI,DAYj)
    F4=T1*TE/PICI
    FINT2 = (W1*F1 +W2*F2 +W3*F3+W4*F4)*VDMG
    RETURN
    END
```



FUNCTION FINT3(Z)
DIMENSION Z(2)
DATA XM1. DTX1/23.7047,2.64371
DATA TYI, YMI, DTYi, P101/87.00,1.47267,.152888,.65991
DATA XM2, DTXZ,YM2,OTY2,RO2, P102183.6095,2.73228,
1 91.174 , C. $8745,0.09438,0.521461$
DATA AYZ.RY3. CY3.DGY3,P103/1.72813,0.387949,91.04,
$10.173002, C .04521$
DATA TY10,YM10,DTY10,P1010/91.04,-C.71632,0.87106,
10.9301221

```
    DATA.W1,W2,W3,W4/.255..467..146,.1321
    x = z(2)
    Y = Z(1)
    CALL DAMAGE(X,Y,VDMG)
C
        BIVARIATF NORMAL COMPONFNT DISTRIEUTIDN COOE = 02
        FI = BVNML(X,Y,XM2,DTX2,YM2,OTY2,RO2)/0102
    C NORMAL-GAMMA CDMPONENT. DISTRIBUTION COOF = 03
    T1 = UNIRP(Y,XM1, DTX1)
    T2 = UGAM(Y,AY3, BY3.CY3,RGY3)
    F2 = T1*T2/P103
    NDRMAL-LOGNORNAL COMPONENT. DISTRIBUTION CODE = 10
    T1 = UNOR(*O.O)= SANE AS AROVE
    T2 = ULGN(Y,TYY10,YM10,DTY1O)
    F3 =T1*T2/R1010
C NORMAL-TRUNCATED LOGNORMAL COMDONENT. CGDE= OI
    TI=UNNR(..) = SANEASS ABOVE
    T2=ULGN(Y:TY1,YM1,DTY1)
    FINT3 = (W1*F1 +W2*F2 +W3*F3+W4*F4)*VDMG
        RETURN
        END
```


## NOTATION

$\mathrm{b}_{\mathrm{m}} \quad$ Basic flow level for the main river
$b_{t} \quad$ Basic flow level for the tributary
C(G) Total construction cost of levee with profile $\underline{G}$
$c_{i}($.$) \quad Construction cost function for cross section i$
D Domain of water profiles
$E_{1}, E_{2}, E_{3}$ Flood events
f(H) pdf of water profiles
G Levee profile; the vector $\left(g_{1}, g_{2}, \ldots, g_{n}\right)$
$g_{i} \quad$ Levee height at cross section $i$
$g_{m} \quad$ Levee height at main river (same as $g_{1}$ )
$g_{t} \quad$ Levee height at tributary end (same as $g_{n}$ )
h A flood stage pair ( $h_{m}, h_{t}$ )
H Water profile; the vector ( $h_{1}, h_{2}, \ldots, h_{n}$ )
$h_{i} \quad$ Water level at cross section $i$
$h_{m} \quad$ Water level at main river (same as $h_{l}$ )
$h_{t} \quad$ Water level at tributary end (same as $h_{n}$ )
L(G,H) Losses caused by flood $\underline{H}$ with levee $\underline{G}$
$l_{i}(.,$.$) Damages function for cross section i$
S Sample space
TEC Total expected cost
$t_{k}(\underline{G}, \underline{H})$ Indicator function; see Equations (2.10) and (2.11)
TYC Total yearly cost (usually in million Forints)
x
Water level at the main river (also $h_{1}$ or $h_{m}$ )
$\begin{array}{ll}x_{L} & \left.\text { Base level for main river (also } b_{m}\right) \\ y & \left.\text { Water level at the tributary (also } h_{n} \text { or } h_{t}\right) \\ y_{L} & \left.\text { Base level for the tributary (also } b_{t}\right)\end{array}$

## REFERENCES

Atkinson, A. C., 1970, "A method for discriminating between models," Journal of the Royal Statistical Society, Series B, Vol. 32, No. 7, pp. 323-353.

Bellman, R., and R. Kalaba, 1965, Quasilinearization and non-linear value problems. American Elsevier Publishing Company, Inc., New York.

Benayoun, R., B. Roy, and B. Sussmann, 1966, Electre, une méthode pour guider le choix en présence de points de vue multiples, Note de Travail no. 49, Direction scientifique, S.E.M.A. (Metra International), Paris.

Benjamin, J. R., and C. A. Cornell, 1970, Probability, Statistics ard Decision for Civil Engineers. McGraw-Hill Book Co., Inc., New York.

Bogardi, I., 1975, "Uncertainty in water resources decision making," Proceedings, U.N. Interregicnal Sem. on River Basin Develop., Budapest, Hungary.

Bogardi, I., L. Duckstein, and F. Szidarovsky, 1975, "Hydrologic system reliability at confluence of rivers," Proceedings, Symp. on Math. Models in Hydrol. \& Water Resources Systems, Bratislava, Czech., No. 115, pp. 36-44.

Boyardi, István, and Máthé Zoltán, 1968, "Determination of the degree of protection offered by flood levees and the economic improvement thereof," National Water Authority, Department for Flood Control and River Training, Budapest, Hungary.

David, F. N., and Evelyn Fix, l96l, "Rank correlation and regression in a non-normal surface," Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, vol. l, Jerzy Neyman, editor, University of California Press, Berkeley, Los Angeles, pp. 177 ff.

Des Raj, 1952, "On estimating the parameters of a bivariate normal population from double or singly linearly truncated samples," Sankya, Vol. 12, pp. 277-290.

Duckstein, L., I. Bogardi, F. Szidarovszky, and D. Davis, 1975, "Sample uncertainty in flood levee design: Bayesian versus non-Bayesian methods," Water Resources Bulletin, Vol. 1l, No. 2, pp. 425-435.

Duckstein, L, , and E. Simpson, 1975, "Uncertainties in karstic water resources systems," Proceedings, Seminar in Karst Hydrology \& Water Res., Dubrovnik, Yugoslavia.

Dumonceaux, Robert, Charles E. Antle, and Gerald Haas, 1972, "Likelihood ratio test for discrimination between two models with unknown location and scale parameters," Technometrics, Vol. 15, No. 1, pp. 1927.

Gaver, Keneth M., and Martin S. Geisel, l974, "Discriminating among alternative models: Bayesian and non-Bayesian methods," Frontiers in Econometrics, Paul Zarembka, edito: Academic Press, New York, London, pp. 49-77.

Ghirtis, G. C., 1967, "Some problems of statistical inference relating to the Double-Gamma distribution," -Trabajos de Estadistica, Vol. 18, pp. 67-87.

Henderson, F., 1966, Open Channel Flow. The Macmillan Company, New York.

Johnson, Norman L., and Samuel Kotz, l970a, Distributions in Statistics: Continuous Univariate Distributions-l. Houghton Mifflin Company, Boston.

Johnson, Norman L., and Samuel Kotz, l970b, Distributions in Statistics: Continuous Univariate Distributions-2. Houghton Mifflin Company, Boston.

Johnson, N., and S. Kotz, 1972, Distributions in Statistics: Continuous Multivariate Distributions. John Wiley \& Sons, Inc., New York, London, Sidney, Toronto.

Kuiper, E., 1965, Water Resources Development. Butterworths, London.

Mardia, K. V., l970, "Families of bivariate distributions," Griffin's Statistical Monographs and Courses \#27, Griffin, London.

Ralston, A., 1965, A First Course in Numerical Analysis. McGraw-Hill, Inc., New York.

Smallwood, R., 1968, "A decision analysis of model selection," IEEE Trans. Systems Science and Cybernetics, Vol. SSC-4, No. 3, pp. 333-341.

Szidarovszky, Ferenc, 1974, "Introduction to numerical methods" (in Hungarian) Közgazdasagi ès Jogí Könyvkiadó, Budapest.

Szidarovszky, F., I. Bogardi, L. Duckstein, and D. Davis, 1976, "Economic uncertainties in.water resources project design," Water Resources Research, in press.

Szidarovszky, F., L. Duckstein, and I. Bogardi, 1975, "Levee system reliability along a confluence reach," ASCE, J. of the Engineering Mechanics Division, Vol. 101, No. EM5, pp. 609-622.

Szidarovszky, F., and S. Yakowitz, 1976, "Analysis of flooding for an open channel subject to random inflow and blockage," unpublished manuscript, Department of Systems Engineering, University of Arizona.

Van Ornum, J. L., 1914, The Regulation of Rivers. McGrawHill Book Company, Inc., New York.

Wood, Eric F., I. Rodriguez-Iturbe, and J. C. Schaake, Jr., 1974, "The methodology of Bayesian inference and decision making applied to extreme hydrologic events," Ralph M. Parsons Laboratorv for Water Resources and Hydrodynamics TR 178, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge.

Yevjevich, V., 1974, "Systematization of flood control measures," Journal of the Hydraulics Division, American Society of Civil Engineers, Vol. loo, No, HY1l, pp. 1537-1548.


[^0]:     FUNCTIDN UBETA. COMPUTFS UNIVARIATE BETA PDF. ARGIJMENTS.
    $A, B, P, Q, A S$ DFFINED IN FUNCTIONAL FORM OF BETA PDF. $D E N=G(P) \nLeftarrow G(Q) / G(P+Q) *(R-A) \nleftarrow *(P+Q-1)$, WHERE G(.) I I S THE GAMMA FUNCTITN. CALLED BY FINTI,FINT3.

    FUNCTION UBETA $(X, A, B, P, Q, D E N)$
    IF $(x-A) 10,10,20$
    $20 \operatorname{IF}(B-X) 10,10,30$
    10 URETA $=0$.
    RETURN

