AAiT

Addis Ababa University
Institute of Technology (AAiT)

Chapter Five

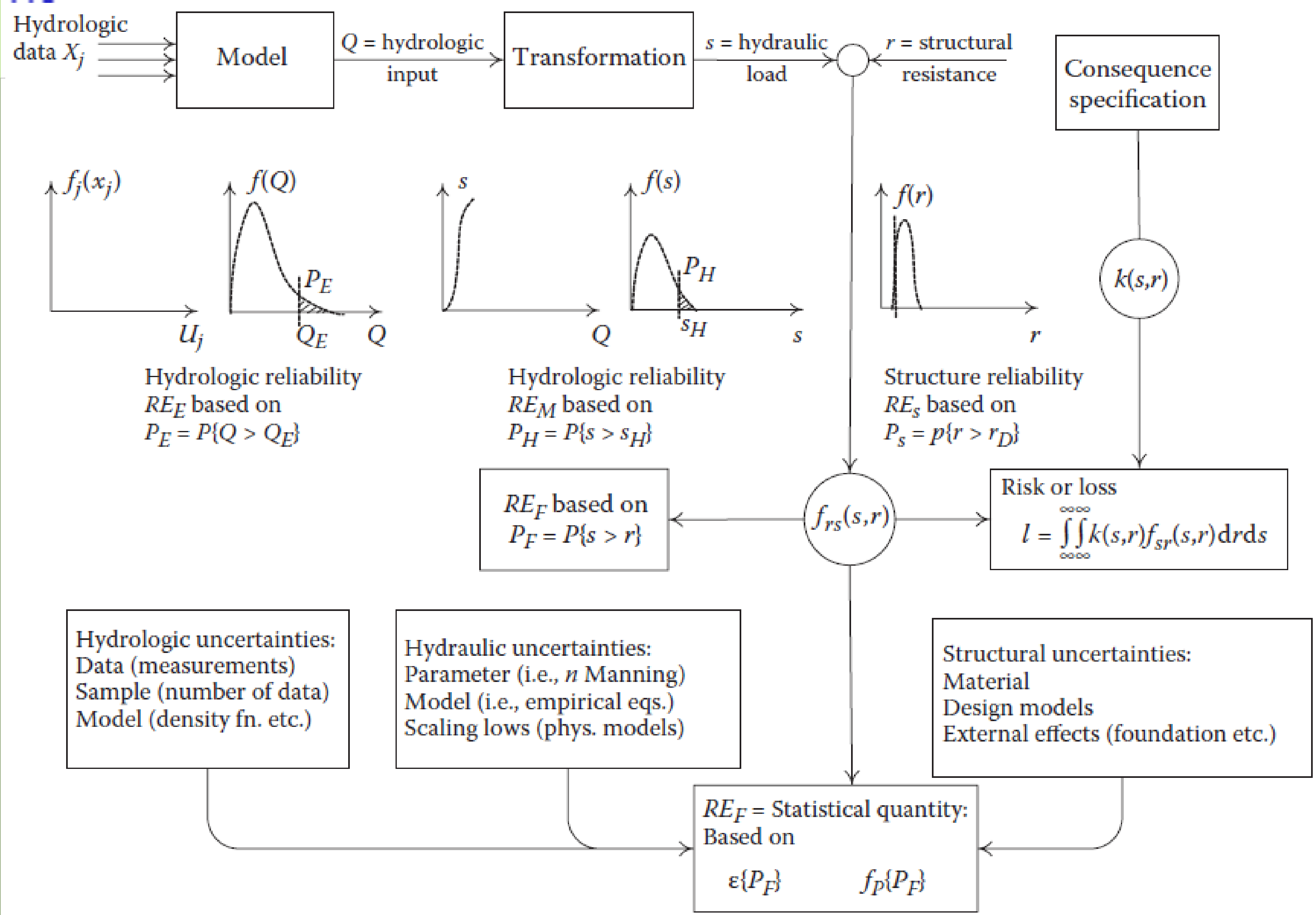
Hydrologic design Risk and Uncertainty Analysis

- **Hydrological design Risk**
  - Risk Analysis
  - Hydroeconomic Analysis
  - Composite Risk Analysis
  - Risk Analysis of safety factors and safety margins
- **Uncertainty Analysis**
  - First Order Analysis of uncertainty
  - Hydrologic design under natural and parameter uncertainty
    - Bayes risk
    - Opportunity Losses
    - Value of Sample Information

Hydrologic design Risk and Uncertainty Analysis

- Given the stochastic nature of stream flow, the planning, design, and operation of water-resource systems are necessarily subject to uncertainty.
- This is particularly so when dealing with design criteria that incorporate extremes.
- The future inflows to which the system is meant to respond are unknown and not predictable with a reasonable degree of reliability.

Generalized concept of risk & reliability analysis for Structures



AAiT

Addis Ababa University
Institute of Technology (AAiT)

Risk Analysis

- **Risk** is the probability of an event $X>X_T$ occurring per year, during a specific time period, n
- Consideration of Risk
  - Structure may fail if event exceeds T–year design magnitude

$$\bar{R} = 1 - \left[1 - \frac{1}{T}\right]^n$$

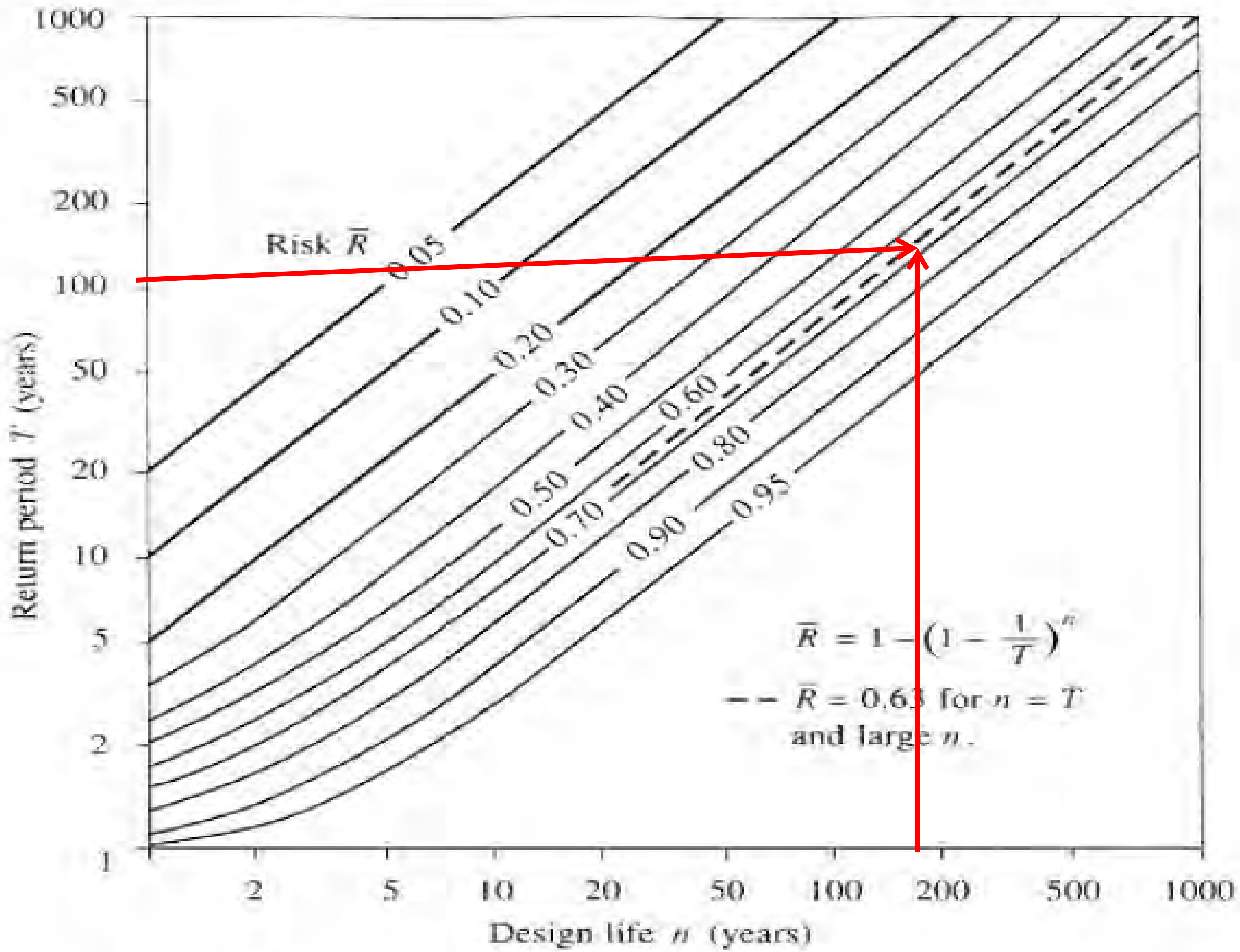
$\bar{R}$ = Probability that an event $X>X_T$ will occur at least once in *n* years

n = The expected life of the structure

T = Return period

- **Reliability** (R) is the complement of risk, or the probability that the loading will not exceed the capacity

Risk of at least one exceedance during the design life



Example

- Expected life of culvert = 10 yrs
- Acceptable risk of 10 % for the culvert capacity
- Find the design return period

$$R = 1 - \left(1 - \frac{1}{T}\right)^n$$

$$0.10 = 1 - \left(1 - \frac{1}{T}\right)^{10}$$

$$T = 95 \;\; yrs$$

- What is the chance that the culvert designed for an event of 95 yr return period will not have its capacity exceeded for 50 yrs?

The risk associated with failure of culvert when the flow exceeds 95 yr flood in the next 95 years is:

$$\overline{R} = 1 - \left(1 - \frac{1}{95}\right)^{50} = 0.41$$

The chance that the capacity will not be exceeded during the next 50 yrs, i.e. the reliability, is 1-0.41 = 0.59

Example

A cofferdam has been built to protect homes in a floodplain until a major channel project can be completed. The cofferdam was built for a 20-year flood event. The channel project will require 3 years to complete. What are the probabilities that

a. The cofferdam will not be overtopped during the 3 years (the reliability)?

b. The cofferdam will be overtopped in any one year?

c. The cofferdam will be overtopped exactly once in 3 years?

d. The cofferdam will be overtopped at least once in 3 years (the risk)?

e. The cofferdam will be overtopped only in the third year?

Solution

a. $\text{Reliability} = \left(1 - \frac{1}{20}\right)^3 = 0.95^3 = 0.86.$

b. $\text{Prob} = \frac{1}{T} = 0.05.$

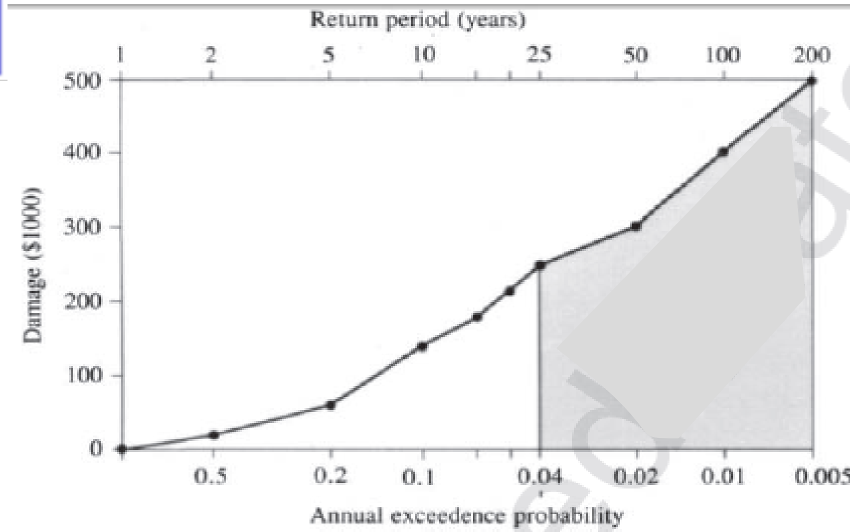
c. $P(x=1) = \binom{3}{1} p^1 (1-p)^2 = 3 \times 0.05 \times 0.95^2 = 0.135.$

d. $\text{Risk} = 1 - \text{Reliability} = 0.14.$

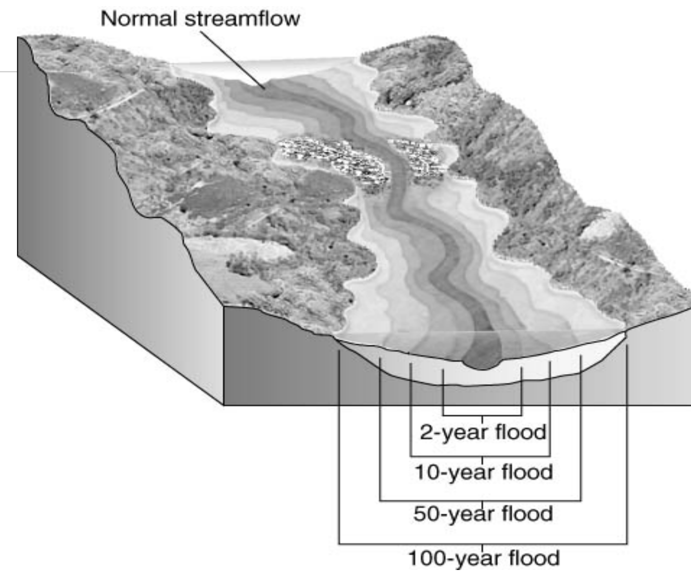
e. $\text{Prob} = (1-p)(1-p)p = 0.95^2 \times 0.05 = 0.045.$

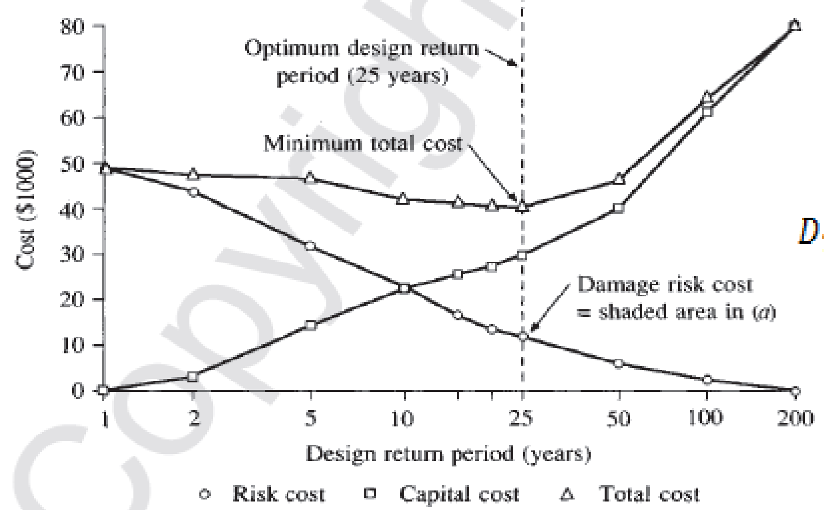
Hydroeconomic Analysis

- Probability distribution of hydrologic event and damage associated with its occurrence are known
- As the design period increases, capital cost increases, but the cost associated with expected damages decreases.
- In hydroeconomic analysis, find return period that has minimum total (capital + damage) cost.



(*a*) Damages for events of various return periods.





(*b*) Hydroeconomic analysis.

Damage risk cost

$$D_T = \sum_{i=1}^{\infty}\left[\frac{D(x_{i-1}) + D(x_i)}{2}\right][P(x \geq x_{i-1}) - P(x \geq x_i)]$$

Example

For events of various return periods at a given location, the damage costs and the annualized capital costs of structures designed to control the events, are shown in columns 4 and 7, respectively, in the Table below. Determine the expected annual damages if no structure is provided, and calculate the optimal design return period.

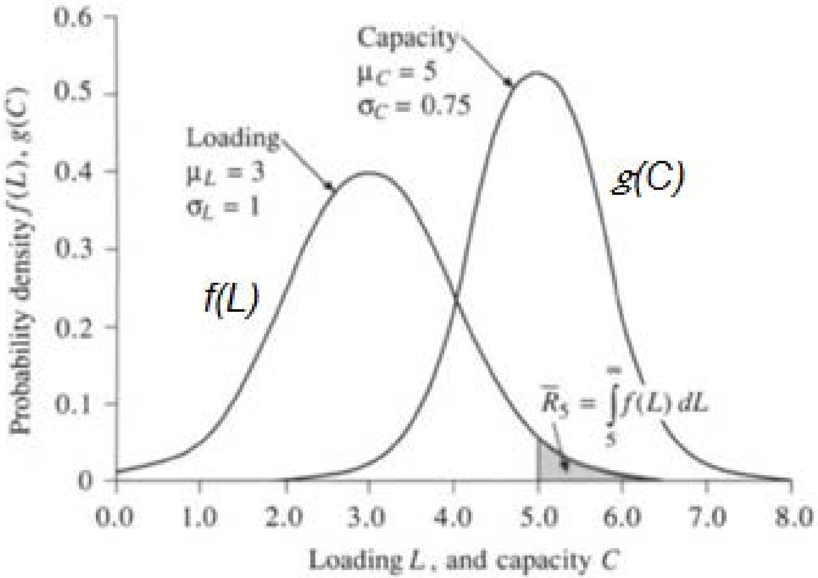
Column:	1 Increment i	2 Return period T (years)	3 Annual exceedence probability	4 Damage ($)	5 Incremental expected damage ($/year)	6 Damage risk cost ($/year)	7 Capital cost ($/year)	8 Total cost ($/year)
		1	1.000	0		49,098	0	49,098
	1	2	0.500	20,000	5,000	44,098	3,000	47,098
	2	5	0.200	60,000	12,000	32,098	14,000	46,098
	3	10	0.100	140,000	10,000	22,098	23,000	45,098
	4	15	0.067	177,000	5,283	16,815	25,000	41,815
	5	20	0.050	213,000	3,250	13,565	27,000	40,565
	6	25	0.040	250,000	2,315	11,250	29,000	*40,250*
	7	50	0.020	300,000	5,500	5,750	40,000	45,750
	8	100	0.010	400,000	3,500	2,250	60,000	62,250
	9	200	0.005	500,000	2,250	0	80,000	80,000

Annual expected damage = $49,098

Optimal design return period

Composite Risk

- Composite risk analysis is a method of accounting for the risks resulting from the various sources of uncertainty



Composite Risk p(L > C)

If capacity is known and fixed

$$P(L > C^*) = \int_{C^*}^{\infty} f(L)\,dL$$

If capacity is random variable

$$\bar{R} = \int_{-\infty}^{\infty} \left[\int_{C}^{\infty} f(L)\,dL \right] g(C)\,dC$$

Reliability p(L ≤ C) = 1 - Risk

$$R = \int_{-\infty}^{\infty} \left[\int_{0}^{C} f(L)\,dL \right] g(C)\,dC$$

Safety Margin (SM)

- SM is the difference between the actual design value (capacity, C) and the hydrologic design obtained by the risk analysis (Loading, L)

SM = C – L

- Reliability (R) = probability that C > L

  R = p(C-L>0) = p(SM > 0)

- Mean value of SM = $\mu_C - \mu_L$
- Variance of SM = $\sigma^2{}_C - \sigma^2 L$
- SM is assumed to be normally distributed and the reduced variate is given by: $$z = \frac{SM - \mu_{SM}}{\sigma_{SM}}$$
- The reliability then becomes: $$R = p\left(z \le \frac{\mu_{SM}}{\sigma_{SM}}\right) = F\left(\frac{\mu_{SM}}{\sigma_{SM}}\right)$$

Factor of Safety (FS)

- FS is the ratio C/L

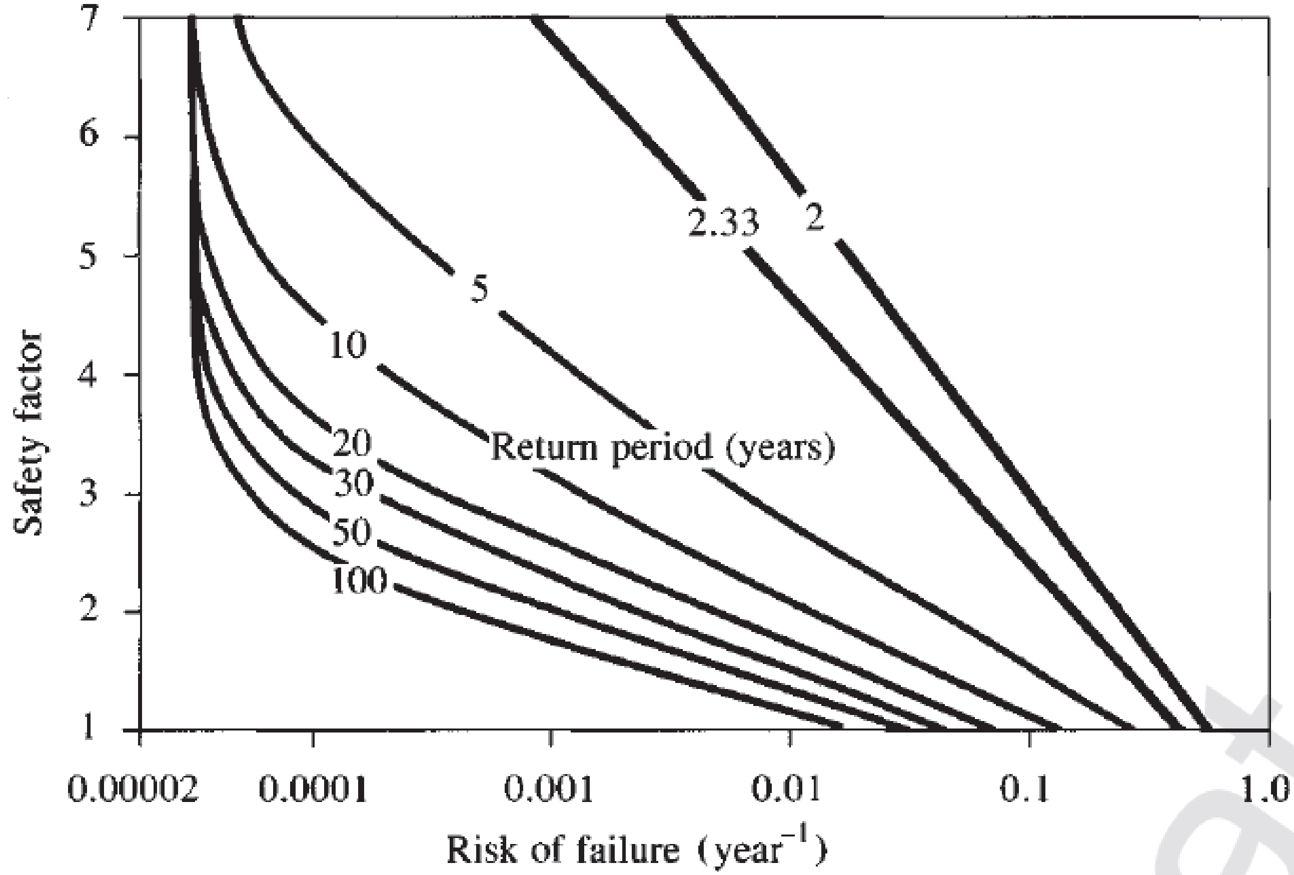
FS = C/L

Reliability (R) = probability that C > L

R = p(FS > 1) = p(ln(FS) > 0) = p(ln(C/L) > 0)

$$R = p\left(z \le \frac{\mu_{\ln FS}}{\sigma_{\ln FS}} \right) = F\left(\frac{\mu_{\ln FS}}{\sigma_{\ln FS}} \right)$$

Risk-Safety Factor-Return Period Relationship



- Inherent uncertainty of hydrologic phenomena can be addressed by the selection of the design return period
- Model and parameter uncertainty by the assignment of arbitrary safety factors or safety margins.

Example

The average surface runoff to a sewer is 3 m^3/s with a standard deviation of 1.2 m^3/s. The mean capacity of the sewer is estimated to be 4.5 m^3/s with a standard deviation of 0.8 m^3/s. Compute the reliability using safety margin approach.

Solution:

$\mu_L = 3;\quad \mu_C = 4.5;\quad \sigma_L = 1.2;\quad \sigma_C = 0.8$

$$\mu_{SM} = \mu_C - \mu_L = 1.5$$

$$\sigma_{SM}^2 = \sigma_C^2 + \sigma_L^2 = 2.08$$

Therefore,

And risk is,

$$R = \left(Z \le \frac{1.5}{\sqrt{2.08}} \right) = 0.851$$

$$= 1 - 0.851 = 0.149$$

Uncertainty

Uncertainties are caused by lack of perfect understanding of hydrologic phenomena and processes involved.

Sources of uncertainty

- Inherent randomness (natural and cannot be eliminated, e.g. weather)
- Model structural error that reflects the inability of a model to represent precisely the system's true behavior
- Model parameter value error
- Data error
- Oprational

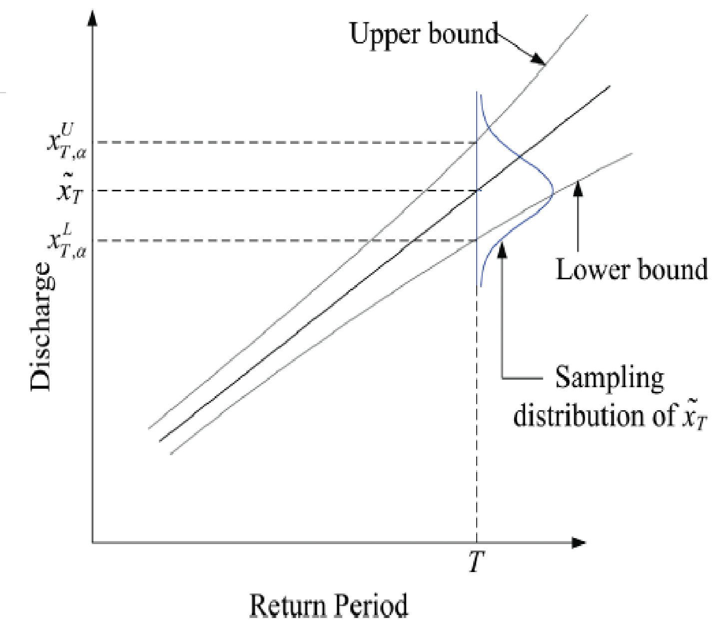
Uncertainty Measures

- Confidence interval:

Drawbacks include

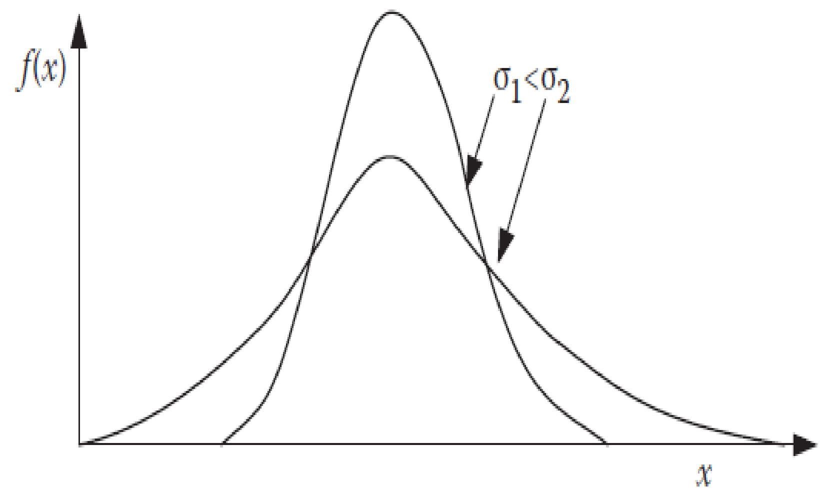
1. Parameter population may not be normally distributed

2. Difficulty of obtaining the overall confidence interval of the system by combining the confidence intervals of individual contributing random components



- Statistical moments,

  e.g. standard deviation



Larger standard deviation means larger uncertainty

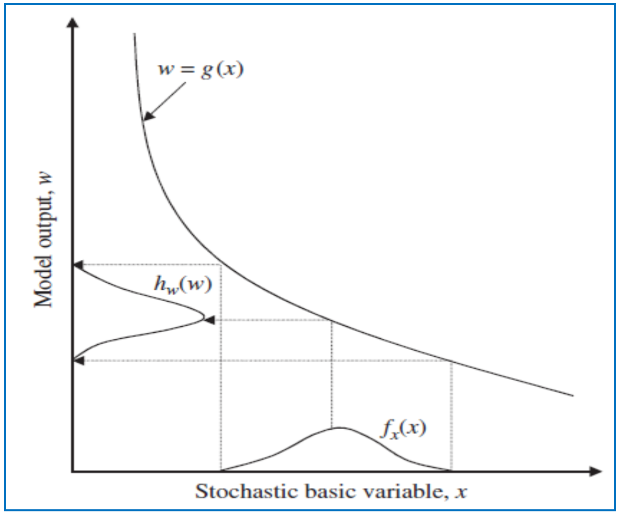
Uncertainty Analysis Methods

Analytic Methods	Approximation Methods	Sampling Methods
Derived distribution	First-order variance estimation	Monte Carlo simulation
Fourier, Laplace, and exponential transforms	Rosenblueth's probabilistic point estimation	Latin hypercube sampling
Mellin transforms	Harr's probabilistic point estimation	Correlated sampling
Estimations of probability and quantile using moments	Li's probabilistic point estimation	Antithetic variates

First Order Analysis

- First-order analysis is used to estimate the uncertainty in a deterministic model formulation involving parameters that are uncertain
- It enables assessment of the combined effect of uncertainty in a model formulation, as well as the use of uncertain parameters
- Consider a r.v. Y= g(X) where X = x_1, x_2,...., x_k are k random variables (r.v.)
- The effect of uncertainty of X on Y can be approximated by the 1st order Tailor series as:

$$y \approx g(\bar{X}) + \sum_{i=1}^{k} \left[\frac{\partial g}{\partial x_i} \right]_{\bar{X}} (x_i - \bar{X})$$



If the k r.v. are independent, then the variance σ^2_y and coefficient of variation Ω_y of y can be approximated as follows:

where $a_i = \left(\partial g / \partial x \right)_{\bar{X}}$

$$\sigma_y^2 = \mathrm{Var}[y] = \sum a_i^2 \sigma_{x_i}^2$$

$$\Omega_y = \left[\sum_{i=1}^{k} a_i^2 \left(\frac{\bar{x}_i}{\mu_y} \right)^2 \Omega_{x_i}^2 \right]^{1/2}$$

AAiT

Addis Ababa University
Institute of Technology (AAiT)

Monte Carlo Analysis

- Monte Carlo analysis uses a random number generator to select parameter values from a known or suspected distribution (probability density function).
- Monte Carlo simulation does not require linearization of the model equations and it is, in theory, nonparametric.
- Monte Carlo analysis is computationally expensive for large models, and it doesn't preserve the covariance structure among the parameters
- Parameter distributions may be determined by multiple samples collected in the field, determined in microcosm experiments in the laboratory, or from expert judgment.
- Typical distribution of parameters include normal and lognormal distributions

Monte Carlo Analysis

Steps

1. Determine the uncertainty distributions (pdf) for each parameter, input function, and variable that will be analyzed

2. Sample each of the parameter distributions using a random number generator that takes into account the probability of each value.

3. Use the set of parameter values selected as input for the first simulation. Run the model and save the output

4. Repeat steps 2 and 3 a large number of times, usually 100-200, or until the statistical output no longer changes by repeated realizations of the model.

5. Sort the stored output data and plot the output as the mean value plus or minus the probability range that is desired.

Bayesian Method

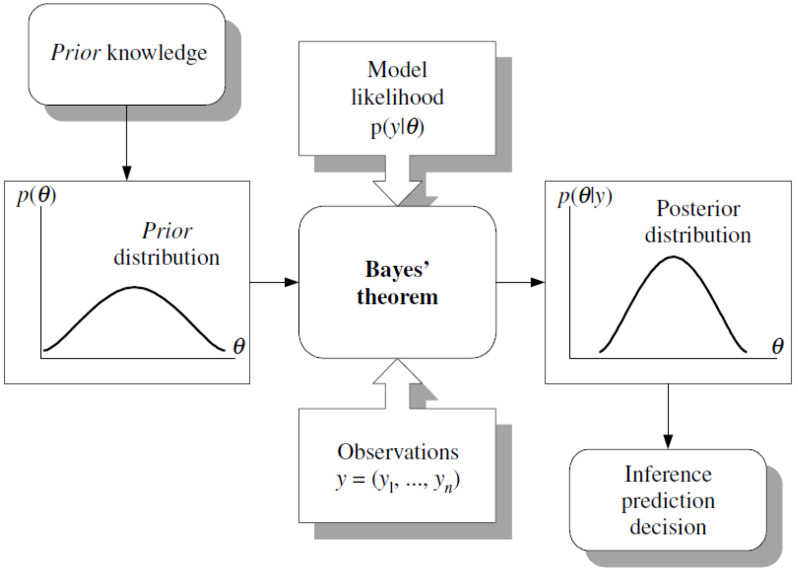
- The values that the parameters $q = (q_1, q_2, \ldots, q_p)$ defining the model state are random
- The uncertainty concerning the parameters is represented by a probability distribution called prior or a priori which is denoted $p(q)$.
- The prior distribution is established based on the information available a priori that does not result from a series of observations $y = (y_1, \ldots, y_n)$, but instead comes from other sources that can be either subjective (expert's or manager's knowledge…) or objective (previous statistical analyses).

AAiT

Addis Ababa University
Institute of Technology (AAiT)

Bayes' Theorem

It enables to update the a priori knowledge in light of the observations



Discrete

$= p\ (y \cap \theta)$

$$p(\theta \mid y) = \frac{p(y \mid \theta)\, p(\theta)}{\sum p(y \mid \theta)\, p(\theta)}$$

Continuous

likelihood

prior

$$p(\theta \mid y) = \frac{p(y \mid \theta)\, p(\theta)}{\int p(y \mid \theta)\, p(\theta)\, d\theta}$$

posterior

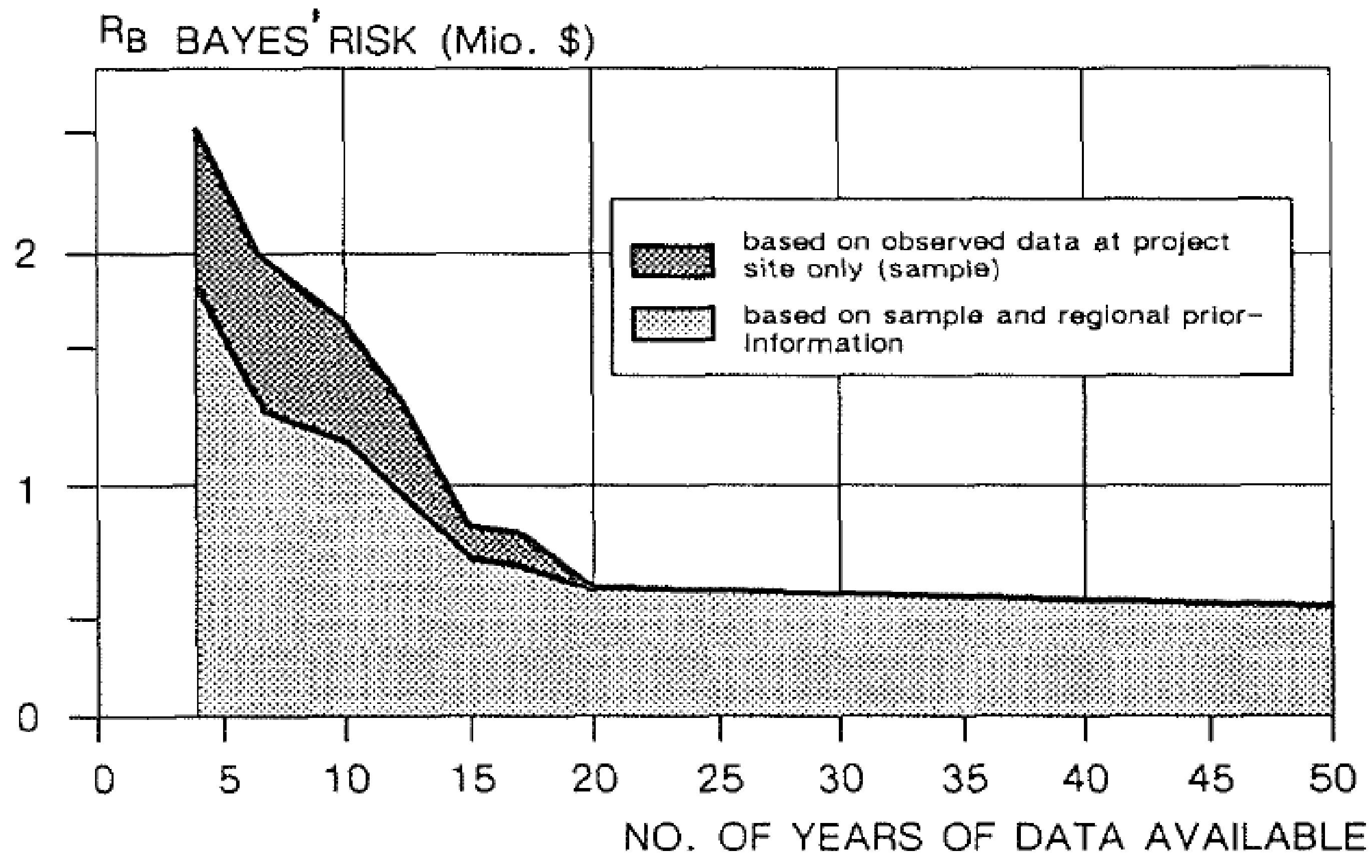
evidence

Bayes Risk

$$R(a) = \sum_y \sum_\theta L(a,\theta)\, p(\theta \mid y)$$

- Bayes Risk = Minimum R(a)
- Bayes decision rule: Minimize R(a)

- a: Action or decision to be taken
- L(a,θ) : Loss incurred if decision a is taken



Opportunity Loss and Expected Value of Sample Information

- Opportunity loss occurs when the outcome of a decision is worse than what could have been obtained if the correct decision had been made.

- Expected value of sample information (EVSI), estimates the benefits of reducing uncertainties by collecting additional data through a monitoring program or adaptive management experiment

Example- Decision in the context of inadequate data

Given

- Cost incurred for additional hydrologic data = 15 million monetary unit (MMU)
- Cost of construction of dam = 160 MMU
- Present value of dam's benefits = 200 MMU
- Possible decision options:
  - a1: build dam
  - a2: do not build dam
- Possible states or outcomes (Success/Failure):
  - Project successful ($\theta 1$) = 0.25
  - Project unsuccessful ($\theta 2$) = 0.75
- Possible indicators of values of additional hydrological data (Effectiveness)
  - X1: increase in effectiveness
  - X2: no change
  - X3: decrease in effectiveness

Required

- Opportunity loss for each decision
- Bayes risk
- Value of sample information

Example- Decision in the context of inadequate data

Joint probabilities of Effectiveness and Outcomes

$$P(X_j \cap \theta_j)$$

Outcome/Options	Increase effectiveness (X1)	No change (X2)	Decrease effectiveness (X3)
Successful	0.20	0.05	0.05
Failure	0.05	0.10	0.55
Marginal probabilities of effectiveness with additional data	0.25	0.15	0.60

AAiT

Addis Ababa University
Institute of Technology (AAiT)

Solution: Prior Opportunity Losses

Loss function $L(a, \theta)$ matrix

Decision	Success ($\theta 1$)	Failure ($\theta 2$)
Build dam (a1)	0	160
Do not build dam (a2)	200	0

- **Opportunity loss if dam is built (a1)**

$R(a1) = L(a1, \theta 1) \times p(\theta 1) + L(a1, \theta 2) \times p(\theta 2) = 120$ MMU

- **Opportunity loss if dam is not built (a2)**

$R(a2) = L(a2, \theta 1) \times p(\theta 1) + L(a2, \theta 2) \times p(\theta 2) = 50$ MMU

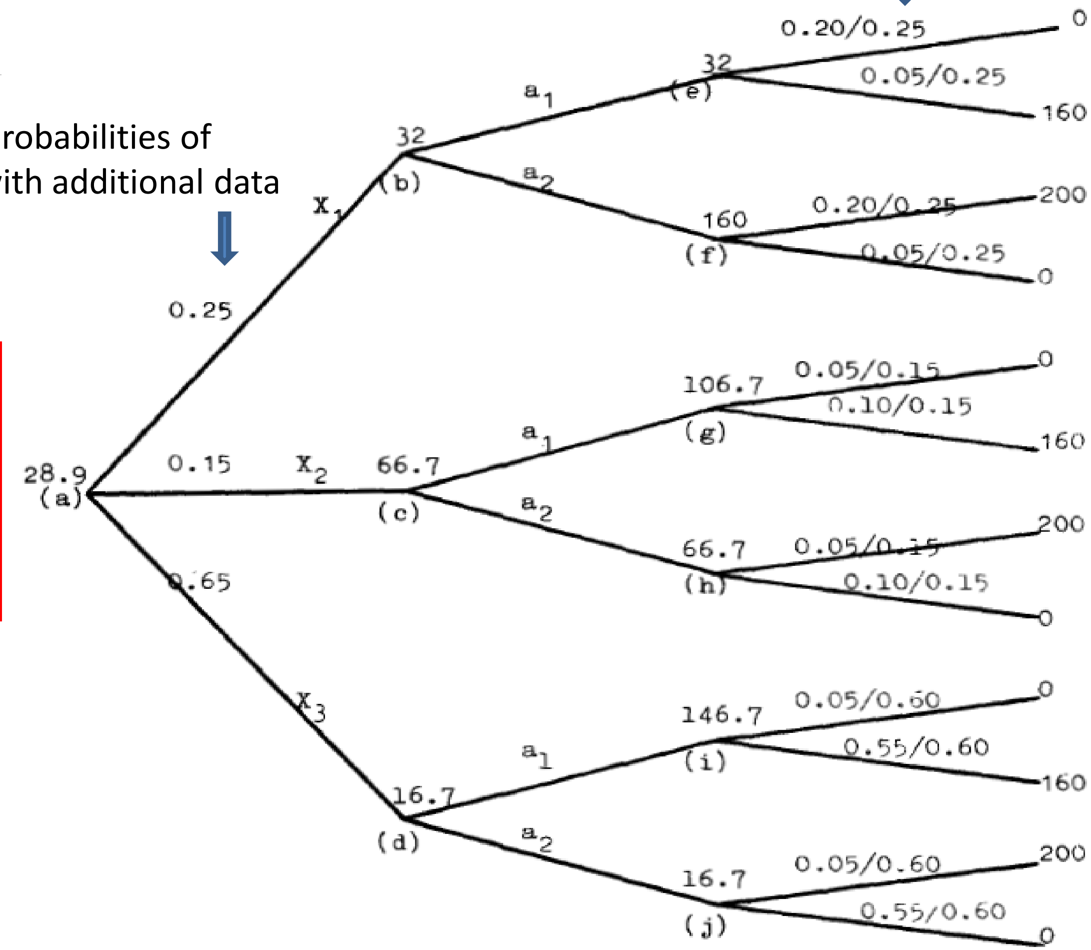
Bayes risk = min (R(a1), R(a2)) = 50 MMU

Best decision with no additional data: Do not build the Dam

Solution: Opportunity Los with additional data and Value of sample information

$$P(\theta_i / X_j) = \frac{P(X_j \cap \theta_i)}{P(X_j)}$$

Marginal probabilities of effectiveness with additional data



Opportunity loss with the additional hydrologic data is: 28.9 MMU

Value of Sample Information (VSI) is the difference between Opportunity loss with no information and Opportunity loss with information = 50 MMU – 28.9 MMU = 21.1 MMU > 15 MMU cost of data collection

→ **Decision: Invest on collection of additional data**