METHODS FOR THE RISK ASSESSMENT AND RISK-BASED MANAGEMENT OF AGING INFRASTRUCTURE

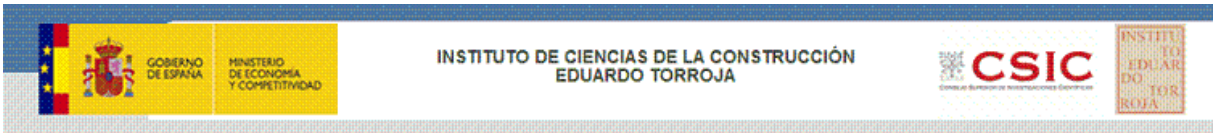
Milan Holický, Dimitris Diamantidis et al.

Lifelong Learning Programme

METHODS FOR THE RISK ASSESSMENT AND RISK-BASED MANAGEMENT OF AGING INFRASTRUCTURE

Editors:
Milan Holický, Czech Technical University in Prague, Klokner Institute
Dimitris Diamantidis, Ostbayerische Technische Hochschule Regensburg

Authors:
Milan Holický, Czech Technical University in Prague, Klokner Institute
Miroslav Sýkora, Czech Technical University in Prague, Klokner Institute
Dimitris Diamantidis, Ostbayerische Technische Hochschule Regensburg
Rafaelle Boccara, Ostbayerische Technische Hochschule Regensburg, and University of Pisa
Peter Tanner, E. Torroja Institute of Construction Sciences, CSIC, Madrid
Carlos Lara, E. Torroja Institute of Construction Sciences, CSIC, Madrid
Pietro Croce, University of Pisa, Department of Civil and Industrial Engineering
Maria Luisa Beconcini, University of Pisa, Department of Civil and Industrial Engineering
Daniele Pellegrini, University of Pisa, Department of Civil and Industrial Engineering
Selcuk Toprak, Pamukkale University, Denizli
Sevket M. Senel, Pamukkale University, Denizli
Mehmet Inel, Pamukkale University, Denizli



This project has been funded with support from the European Commission.
This publication [communication] reflects the views only of the author, and the Commission cannot be heldresponsible for any use which may be made of the information contained therein.

METHODS FOR THE RISK ASSESSMENT AND RISK-BASED MANAGEMENT OF AGING INFRASTRUCTURE

ISBN: 978-80-01-05611-0

Edited by:
Milan Holický, Czech Technical University in Prague, Klokner Institute
Dimitris Diamantidis, Ostbayerische Technische Hochschule Regensburg

Published by:
Czech Technical University in Prague, Klokner Institute
Šolínova 7, 166 08 Prague 6, Czech Republic

Pages: 188

Printed by:
Česká technika – nakladatelství ČVUT

1st edition

FOREWORD

The Leonardo da Vinci Project, "Innovation Transfer in Risk Assessment and Management of Aging Infrastructures", CZ/13/LLP-LdV/TOI/134014, addresses the urgent need to train students, young engineers and professionals in the assessment of existing infrastructures. The future of the construction industry lies in moving from new constructions towards the maintenance, repair and rehabilitation of existing structures, particularly of aging infrastructures. Risk management consisting of risk assessment and risk control of existing structure therefore plays thereby an important role.

The assessment of existing infrastructures is an imperative issue of great economic significance in most countries around the world, as more than 50 % of all construction activities concerns existing buildings, bridges and other civil engineering works. At present, the Eurocodes which will be used in all CEN Member countries are primarily focused on the design of new structures. Additional operational rules for existing structures are still missing. The international standard ISO 13822 provides only general principles for the assessment of existing structures which should be further developed for their effective operational use in practice.

The overall procedure of risk management of aging infrastructures is therefore an urgent issue of great economic significance in most countries around the world. This is primarily caused due to the fact that many infrastructures are affected by action effects, fatigue and environmental influences, and inevitably deteriorate. Moreover, available resources for the maintenance, repair or possible replacement of infrastructures are always limited. Typical questions and problems which must be solved in this context can be summarized as follows:

- What is the current state of a particular infrastructure?
- What is an acceptable level of the relevant risk?
- Which aspects are critical to sustainable performance?
- What is the optimum life-cycle cost?
- What is the best long-term funding strategy?

The proposed project highlights the above problems and focusses also on vocational training in risk management and assessment of aging infrastructure in the partner countries (CZ, DE, ES, IT and TR).

The current project addresses the urgent need for practical implementation of principles of the management and risk assessment of existing infrastructures in all the partner countries. The project is supported by the Czech Chamber of Chartered Engineers (ČKAIT). The project consortium, under the leadership of the Klokner Institute of the Czech Technical University in Prague (KI CTU), consists of five other partners, including research institutions and universities from three EU Member States (DE, ES, IT) as well as one associated country (TR), and also the Secondary Technical School of Civil Engineering in Ceske Budejovice (CZ). All researchers within the partnership are involved in research projects dealing with reliability assessment of existing structures. They participate in the national and international standardization activities within the international organizations CEN (European Committee for Standardisation) and ISO (International Standardization Organisation).

The project outcomes consist of vocational training materials based on international research and relevant standard committee's activities, and on the significant experience gained from case studies, and on selected results obtained from the previous projects of the Leonardo da Vinci programme. In particular, the experience gained by the partners of the present consortium from the completed project on vocational training related to assessment of

existing structures is utilized. Available innovations are transferred to undergraduate and graduate students, practising engineers and local government representatives. The developed training materials are of practical use in offering illustrative examples and several case studies.

All the project outcomes are based on documents of the international research organization, Joint Committee on Structural Safety JCSS and international research projects, selected outcomes of the previous project of the Leonardo da Vinci programme (developed by five partners of the present consortium in 2013-2015) and also on background documents to the new European and international standards.

Particular outcomes of the proposed project include practical handbooks, software tools, web-sites, e-learning, courses and seminars organized in the partners´ home countries. The project outcomes will also be offered to other European countries through code committees work, lectures and the project web-sites. It is foreseen that the outcomes being used by engineers, designers, technicians, and representatives of public authorities involved in planning and management will have an impact on the risk-based management of infrastructures.

One of the important project outcomes is the present Handbook 1 "Methods for the Risk Assessment and Risk-Based Management of Aging Infrastructure", which is focussed on general procedures of risk-management and assessment of aging infrastructures. The methodologies provided are independent of the type of structure and material, and are compatible with the background methodologies used in the Eurocodes. Operational techniques for the assessment of aging infrastructures supplemented by illustrative case studies are presented in Handbook 2 and Handbook 3, which also result from the mentioned project.

Handbook 1 consists of ten chapters and one annex. Chapter 1 gives a brief overview of the basic concepts used in the management and risk assessment of existing structures, supplemented by the basic definitions and terminology. Chapter 2 summarizes current applicable standards, codes and recommendations related to actions and environmental influences on existing structures. Chapter 3 reviews the most important degradation models due to environmental influences. Fatigue effects are treated separately in Chapter 4. The following Chapters, 5, 6, 7 and 8, are devoted step by step to the entire procedure of risk management, including probabilistic reliability analysis, consequence analysis, risk evaluation and risk assessment based on risk acceptance criteria, and consequently provide important information supporting decisions concerning the safety of the structure. The last two Chapters, 9 and 10, illustrate instructive case studies concerning risk management and risk assessment of road tunnels. Annex A presents basic statistical methods and techniques used in the evaluation and updating of available data required for probability and consequence analysis.

It is believed that the material of this handbook is presented in an understandable way, illustrated by several realistic examples. A number of references mentioned in each chapter provide additional background materials, further guidance and information useful for particular practical applications.

Prague and Regensburg, 2014

CONTENTS

CHAPTER 1: BASIC CONCEPTS

Milan Holický[1] and Dimitris Diamantidis[2]

[1]Klokner Institute, Czech Technical University in Prague, Czech Republic
[2]OTH Regensburg, Germany

Summary

Decision concerning different kinds of infrastructure is always associated with various types of risk, where risk is defined as the expected consequences (probability times consequence) of undesirable events. Risk assessment, as a key part of the risk management, provides decision makers with procedures to determine whether or not and in what manner it is appropriate to treat risks. The described procedure allows to compare alternative solutions on a rational basis or indicates the optimum way of risk reduction. The assessed risk is evaluated using criteria for acceptability or tolerability to support final decision. The basic concepts for risk assessment of aging infrastructures are summarized in this chapter. The terminology is provided in the annex given at the end of this chapter.

1 INTRODUCTION

1.1 Background Documents

Principles of risk assessment and the common tools applied for analysing the risk of civil engineering systems described in this text are based on the common concepts presented in International Standard ISO and other documents [1] to [10]. Particularly the new ISO document [10] "General principles on risk assessment of systems involving structures" provides valuable fundamentals including terminology. In addition, publications [11] to [17] and working materials of the JCSS [18] offer supplementary information that can be utilized in practical applications.

1.2 General Principles

Within the broader context of risk management, risk assessment provides decision makers with procedures for determining risk and for comparing possible alternatives of risk treatment. Risk management includes generally two major procedures: risk assessment and risk control. The risk assessment itself combines the determination of possible consequences of foreseen events and the assessment of the associated probabilities. The assessed risks are finally evaluated using appropriate criteria for acceptability or tolerability to support final decision.

2 RISK MANAGEMENT

2.1 Standardization

It appears that the standardization of general principles and rules for the risk assessment of civil engineering systems including infrastructures is of the utmost importance. In 2004 the International Standard Organization ISO, technical committee TC98, subcommittee SC2 established a new working group responsible for developing a new document on the general principles on risk assessment of systems involving structures [10]. The new International Standard should provide a common methodology and clearly defined terminology that would improve not only the common practice of risk assessment, but also risk communication as an extremely important component of the whole risk management. The focus is on strategic and operational decision-making related to design, assessment, maintenance and decommissioning of structures.

The aim of the submitted handbook is to describe important aspects of risk assessment that are valuable for practical applications to various civil engineering systems including infrastructures.

2.2 General Framework

A number of national standards and some parts of newly developing European standards for structural design [11, 15] often refer to methods of risk assessment, particularly in the case of structures exposed to accidental actions when the consequences of adverse events are significant. Accidental design situations seem to occur more and more frequently and extents of unfavourable events are of increasing importance. However, up to now no internationally accepted document for the risk analysis and risk evaluation of civil engineering systems has not been available.

The general framework of risk management based on the Standards CAN/CSA [2], technical documents of CIB [6] and ISO [7, 9] is schematically indicated in Figure 1. The main parts of risk management, risk assessment and risk control including the basic terminology used in this text are indicated in Figure 1 (adopted from [2]).

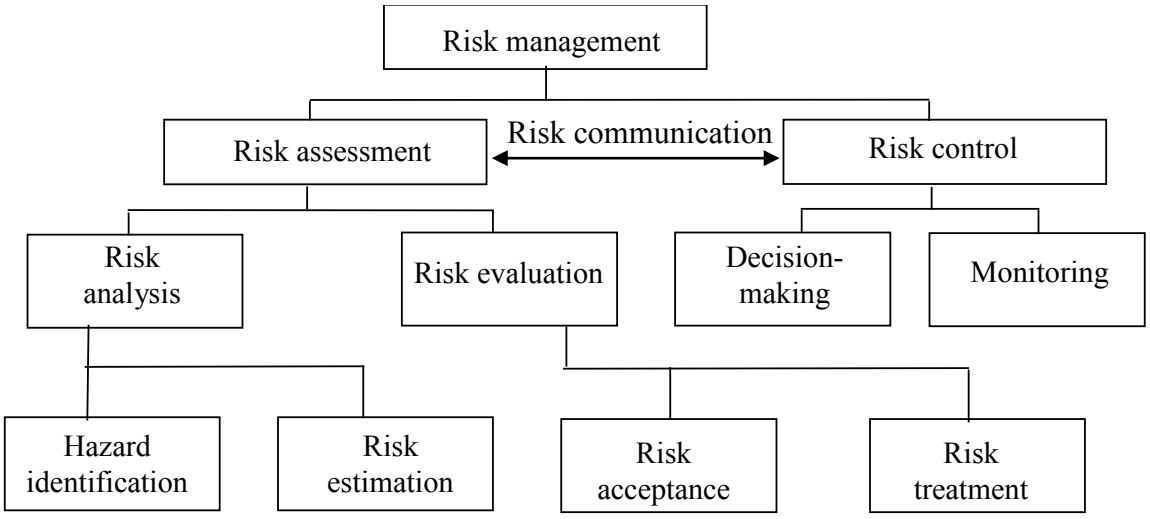


Figure 1 A framework for risk management (adopted from [2])

3 RISK ASSESSMENT PROCESS

3.1 General Procedure

It follows from Figure 1 that the risk assessment of a system is an important part of the whole risk management. The risk assessment consists of risk analysis and risk evaluation. The risk analysis of a system consists of the use of all available information to estimate the risk to individuals or populations, property or the environment, from identified hazards. The risk assessment further includes risk evaluation (acceptance or treatment) as indicated in Figure 1 (adopted from [2]). The whole procedure of risk assessment is typically an iterative process as indicated in Figure 2 (adopted from [7]). The first step in the risk analysis involves the context (scope) definition related to the system and the subsequent identification of hazards.

3.2 Definition of the System

The system is understood [2] as a bounded group of interrelated, interdependent or interacting elements forming an entity that achieves in its environment a defined objective through the interaction of its parts. In the case of technological hazards related to civil engineering works, a system is normally formed from a physical subsystem, a human subsystem, their management, and the environment. Note that the risk analysis of civil engineering systems (similar to the analysis of most systems) usually involves several interdependent components (for example human life, injuries, environmental and economic loss).

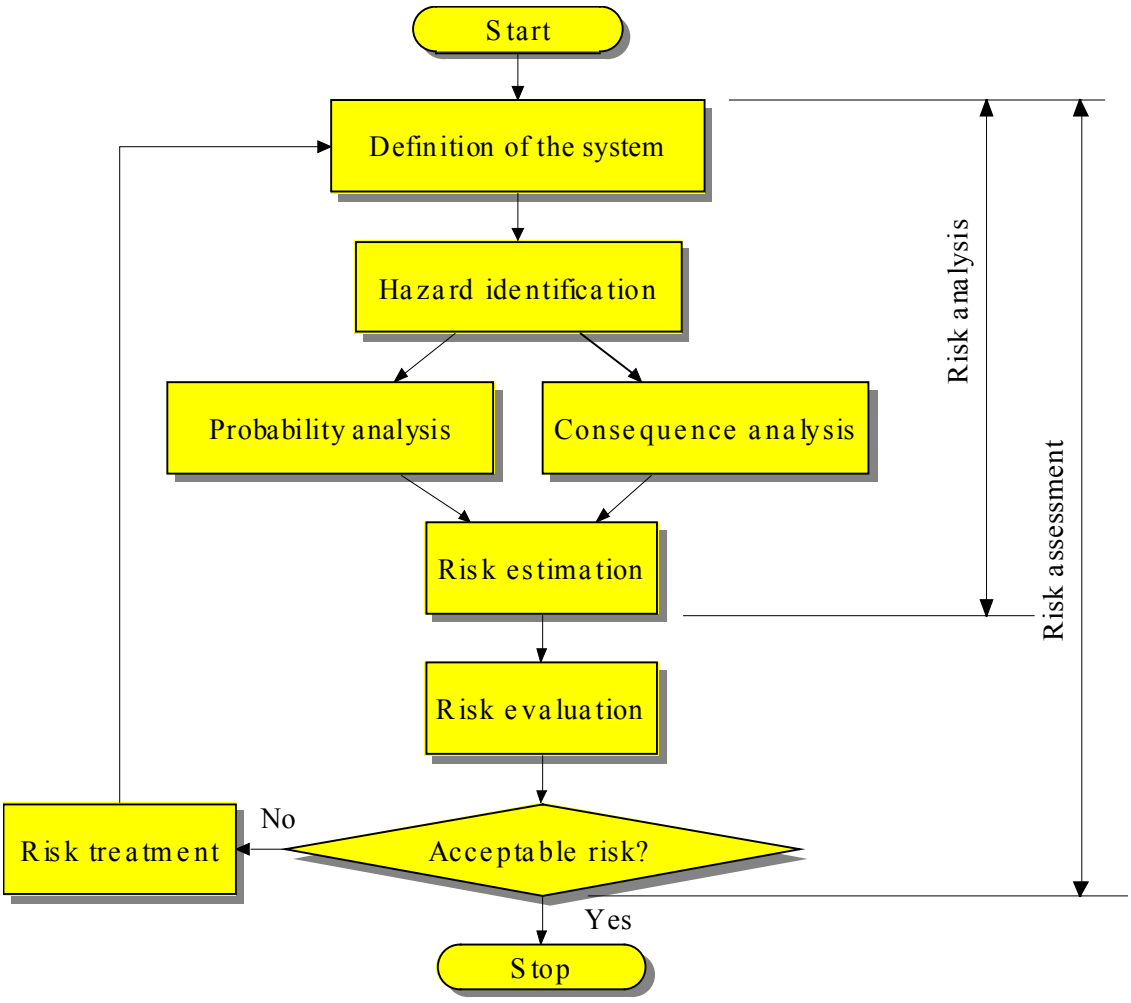


Figure 2 Flowchart of iterative procedure for risk assessment (adopted from [7])

Any technical system may be exposed to a multitude of possible hazard situations. In the case of civil engineering facilities, hazard situations may include both environmental effects (wind, temperature, snow, avalanches, rock falls, ground effects, water and ground water, chemical or physical attacks, etc.) and human activities (usage, chemical or physical attacks, fire, explosion, etc.). As a rule, hazard situations due to human errors are more significant than hazards due to environmental effects.

3.3 Hazard Identification and Scenarios

A hazard is a set of circumstances, possibly occurring within a given system, with the potential for causing events with undesirable consequences. For instance, the hazard of a civil engineering system may be a set of circumstances with the potential to cause an abnormal action (for example fire, explosion) or environmental influence (flooding, tornado) and/or insufficient strength or resistance or an excessive deviation from intended dimensions. In the case of a chemical substance, the hazard may be a set of circumstances likely to cause its exposure [2].

Hazard identification and modelling is the process of recognising the hazard and defining its characteristics in time and space. In the case of civil engineering systems the hazards H_i may be linked to various design situations of the structure (as defined in [7]) including persistent, transient and accidental design situation. As a rule, H_i are mutually exclusive situations (for example persistent and accidental design situations of a building). Then if the situation H_i occurs with the probability $\mathrm{P}\{H_i\}$, it holds $\sum \mathrm{P}\{H_i\} = 1$. If the situations H_i are not mutually exclusive, then the analysis becomes more complicated.

Note that in some documents (for example in the recent European document EN 1990 [11]) the hazard is defined as an event, while in risk analysis [2] it is usually considered a condition with the potential for causing events, thus as a synonym for danger.

A hazard scenario is a sequence of possible events for a given hazard leading to undesired consequences. Identifying what might go wrong with the system or its subsystem is a crucial task for risk analysis. It requires detailed examination and understanding of the system [2]. Nevertheless, a given system is often part of a larger system. Consequently, modelling and subsequent analysis of the system is a conditional analysis.

The modelling of relevant scenarios may be dependent on specific characteristics of the system. For this reason a variety of techniques have been developed for the identification of hazards (for example PHA HAZOP) and for the modelling of relevant scenarios (fault tree, event tree/decision trees, causal networks). A detailed description of these techniques is beyond the scope of this text, but may be found in literature (for example [12, 18]).

3.4 Estimation of Probabilities

Probability is generally the likelihood or degree of certainty of a particular event occurring during a specified period of time. In particular, the reliability of a structure is often expressed as the probability related to a specific requirement and a given period of time, for example 50 years [11].

Assuming that a system may be found in mutually exclusive situations H_i, and the failure F of the system (for example of the structure or its element) given a particular situation H_i occurs with the conditional probability $\mathrm{P}\{F|H_i\}$, then the total probability of failure p_f is given by the law of total probability (see for example [11]) as:

$$p_f = \sum_i \mathrm{P}\{H_i\}\mathrm{P}\{F \mid H_i\} \tag{1}$$

Equation (1) can be used for the modification of the partial probabilities $\mathrm{P}\{H_i\}\mathrm{P}\{F|H_i\}$ (appropriate to the situations H_i) with the aim to comply with the design condition $p_f < p_t$, where p_t is a specified target probability of failure. The target value p_t may be determined using the probabilistic optimisation of an objective function describing, for example, the total cost.

The conditional probabilities $\mathrm{P}\{F|H_i\}$ must be determined by a detailed probabilistic analysis of the respective situations H_i under relevant scenarios. The traditional reliability methods [2] assume that the failure F of the system can be well defined in the domain of the vector of basic variables $\boldsymbol{X}$. For example, it is assumed that a system failure may be defined by the inequality $\mathrm{g}(\boldsymbol{x}) < 0$, where $\mathrm{g}(\boldsymbol{x})$ is the so-called limit state function, where $\boldsymbol{x}$ is a realisation of the vector $\boldsymbol{X}$. Note that $\mathrm{g}(\boldsymbol{x}) = 0$ describes the boundary of the limit state, and the inequality $\mathrm{g}(\boldsymbol{x}) > 0$ the safe state of a structure.

If the joint probability density $\mathrm{f}_{\boldsymbol{X}}(\boldsymbol{x}|H_i)$ of basic variables $\boldsymbol{X}$ given the situation H_i is known, the conditional probability of failure $\mathrm{P}\{F|H_i\}$ can then be determined using the integral

$$\mathrm{P}\{F|H_i\} = \int_{\mathrm{g}(\boldsymbol{x})<0} \mathrm{f}_{\boldsymbol{X}}(\boldsymbol{x} \mid H_i)\mathrm{d}\boldsymbol{x} \tag{2}$$

It should be mentioned that the probability $\mathrm{P}\{F|H_i\}$, calculated using Equation (2), suffers generally from two essential deficiencies:

– Uncertainty in the definition of the limit state function $\mathrm{g}(\boldsymbol{x})$
– Uncertainty in the theoretical model for the density function $\mathrm{f}_{\boldsymbol{X}}(\boldsymbol{x}|H_i)$ of basic variables $\boldsymbol{X}$

These deficiencies are most likely the causes of the observed discrepancy between the determined probability p_f and actual frequency of failures; this problem is particularly evident in the case of fire. Yet, the probability requirement $p_f < p_t$ is generally accepted as a basic criterion for the design of structures.

In a risk analysis we need to know not only the probability of the structural failure F, but probabilities of all events having unfavourable consequences. In general, the situations H_i may cause a number of events E_{ij} (for example excessive deformations, full development of the fire). The required conditional probabilities $\mathrm{P}\{E_{ij}|H_i\}$ must be estimated by a separate analysis using various methods, for example the fault tree method or causal networks.

More details related to computation of probabilities and especially probabilities of failure can be found in Chapter 4. In addition, the statistical evaluation of data is presented in detail in Annex A of this handbook.

3.5 Estimation of Consequences

Consequences are possible outcomes of a desired or undesired event that may be expressed verbally or numerically to define the extent of human fatalities and injuries or environmental damage and economic loss. A systematic procedure to describe and/or calculate consequences is called the consequence analysis. Obviously, consequences are generally not one-dimensional. However, in specific cases they may be simplified and described by several components only, for example by human fatalities, environmental damage and costs. At present only various costs have usually been included. It is assumed that adverse consequences of the events E_{ij} can normally be expressed by several components $C_{ij,k}$,

where the subscript k denotes the individual components (for example the number of lost lives, the number of human injuries and the damage expressed in a certain currency). The classification of consequences is further discussed in Chapters 5 and 6.

3.6 Estimation of Risk

Risk is a measure of the danger that undesired events represent for humans, the environment or economic values. It is commonly expressed in the probability and consequences of the undesired events. It is often estimated by the mathematical expectation of the consequences of an undesired event. Then it is the product "probability × consequences". However, a more general interpretation of risk involves probability and consequences in a non-product form. This presentation is sometimes useful, particularly when a spectrum of consequences, with each magnitude having its own probability of occurrence, is considered [9].

The estimation of risk is a process used to produce an estimate of the measure of risk. As already stated above, the risk estimation is based on the hazard identification and generally contains the following steps: scope definition, frequency analysis, consequence analysis, and their integration [10]. If there is one-to-one mapping between the consequences (utility) $C_{ij,k}$ and the events E_{ij}, then the total risk R_k related to the considered situations H_i is the sum

$$R_k = \sum_{ij} C_{ij,k} \mathrm{P}\{E_{ij} \mid H_i\} \mathrm{P}\{H_i\} \tag{3}$$

If the dependence of consequences on events is more complicated than just one-to-one mapping, then equation (3) will have to be modified. A practical example of equation (3) can be found in [14, 17, 18], where an attempt to estimate the risk due to persistent and fire design situation is presented.

In some cases it is possible to deal with one-component risk R only. Then the subscript k in equation (3) may be omitted. Moreover, the probability of undesired events may depend on the vector of basic variables $\boldsymbol{X}$. Then the total risk R may be formally written as

$$R = \int C(\boldsymbol{x}) \mathrm{f}_{\boldsymbol{X}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{4}$$

where $R(\boldsymbol{x})$ denotes the degree of risk as a function of basic variables $\boldsymbol{X}$, and $\mathrm{f}_{\boldsymbol{X}}(\boldsymbol{x})$ denotes the joint probability density function $\boldsymbol{X}$.

Risk assessment technologies are described in Chapter 5 and the terminology of risk assessment is defined in the annex to this introductory chapter.

4 DECISION PROCESS

The decision-making is generally based on the process of risk acceptance and option analysis (see Figure 1) that is sometimes referred to as the risk evaluation. The risk acceptance is based on various criteria of risk that are the reference points against which the results of the risk analysis are to be assessed. The criteria are generally based on regulations, standards, experience, and/or theoretical knowledge used as a basis for the decision about the acceptable risk. Acceptance criteria and the criteria of risk may sometimes be distinguished [6]. Various aspects may be considered, including cultural, social, psychological, economic and other aspects [6]. Generally, the acceptance criteria may be expressed verbally or numerically.

Assuming for example that the acceptance limits $C_{k,\mathrm{d}}$ of the components C_k are specified, then it is possible to design the structure on the basis of acceptable risks using the criterion $C_k < C_{k,\mathrm{d}}$, which may supplement the probability requirement $p_f < p_\mathrm{t}$.

It should be noted that various levels of risk might be recognised, for example acceptable risk, tolerable risk, and objective risk [6] (see the definitions of these terms). It is a remarkable fact that the public seems to be generally better prepared to accept certain risks than to stand for specified probabilities of failure [6, 12]. The derivation of acceptance criteria is further analytically described in Chapter 6.

5 CONCLUDING REMARKS

The risk is commonly estimated by the mathematical expectation of the consequences of an undesired event that often leads to the product "probability × consequences". As a rule, the risk of civil engineering systems including infrastructures is a multidimensional quantity having several components.

Risk analysis is based on hazard identification and generally contains the following steps: the scope definition, hazard identification, definition and modelling of hazard scenarios, estimation of probabilities, estimation of consequences, estimation of risk and decision-making.

The most important contribution of risk analysis and assessment consists of the systematic consideration of various consequences. Several techniques are available at present and are described in Chapter 5 of this handbook and provide a transparent, logical and effective tool for analysing engineering systems. It should, however, be underlined that any analysis of an engineering system is always dependent on the assumed input data, often of a very uncertain nature. The input data should be estimated with due regard to the specific technological and economic conditions of a given system. In particular, the economic, social and environmental consequences of adverse events should be further investigated.

It appears that the methods of risk analysis and assessment may significantly contribute to a further improvement of current engineering design including assessment of infrastructures. The remarkable fact that the public is better prepared to accept certain risks than to stand for specified probabilities of failure will make the application of risk assessment easier. It is therefore anticipated that in the near future probabilistic methods in engineering will be supplemented by criteria for acceptable risks. Obviously the proposed new International Standard ISO will be extremely useful.

REFERENCES

[1] NS 5814 Requirements for risk analysis (1991).
[2] CAN/CSA - Q634-91 Risk analysis requirements and guidelines (1991).
[3] ISO 2394 General principles on reliability of structures (1998).
[4] ISO/DIS 8930 General principles on reliability of structures - List of equivalent terms (1999).
[5] TNO report 96-CON-R1599 Proposal for a framework in behalf of developing terminology with regard to the process of the probabilistic design and/or assessment of building and civil engineering structures with reference to ISO 8930 (1996).
[6] CIB TG 32, Report 259 Risk assessment and risk communication in civil engineering, CIB secretariat (2001).

[7] ISO/IEC Guide 73: 2002, Risk management – Vocabulary - Guidelines for use in standards (2002).
[8] ISO 9000: Quality management systems – Fundamentals and vocabulary (2000).
[9] ISO/IEC Guide 51: Safety aspects – Guidelines for their inclusion in standards (1999).
[10] ISO ISO 13824: Bases for design of structures - General principles on risk assessment of systems involving structures (2009).
[11] EN 1990 Eurocode – Basis of structural design. CEN (2002).
[12] Stewart M. G. and Melchers R. E. (1997) Probabilistic risk assessment of engineering system. Chapman & Hall, London.
[13] Jensen F. V (1996) Introduction to Bayesian networks. Aalborg University, Denmark.
[14] Holický M. (2009) Reliability analysis for structural design. Sun Press.
[15] Gulvanessian H., Calgaro J.-A. and Holický M. (2002) Designer's guide to EN 1990, Eurocode: Basis of structural design. Thomas Telford, London.
[16] Holický M. (2013) Introduction to probability and statistics for engineers, Springer.
[17] Holický M. (2001) Prospects for advanced engineering design based on risk assessment. *Acta Polytechnica*, Vol. 41, No. 4-5, pp. 8-12.
[18] JCSS: http://www.jcss.byg.dtu.dk/

ANNEX: TERMINOLOGY OF RISK ASSESSMENT

The following definitions are provided to assure consistent understanding of selected terms within the scope of this Leonardo da Vinci project. The aim of this annex is to improve general knowledge and understanding within the fields of safety, risk, reliability and quality assurance for all types of civil engineering works and activities. The following documents have been used for the preparation of this draft; the definitions provided in recent CIB TG 32 document [8] ISO/IEC Guide 73 [9] are acknowledged in particular:

[1] NS 5814 Requirements for risk analysis (1991).
[2] CAN/CSA - Q634-91 Risk analysis requirements and guidelines (1991).
[3] ISO 2394 General principles on reliability of structures (1998).
[4] ISO/DIS 8930 General principles on reliability of structures - List of equivalent terms (1999).
[5] TNO report 96-CON-R1599 Proposal for a framework in behalf of developing terminology with regard to the process of the probabilistic design and/or assessment of building and civil engineering structures with reference to ISO 8930. 1996.
[6] Stewart M. G. and Melchers R. E. (1997) Probabilistic risk assessment of engineering system. Chapman & Hall, London.
[7] EN 1990: Basis of design. CEN TC 250, draft (November 1999).
[8] CIB TG 32, Report 259 Risk assessment and risk communication in civil engineering, CIB secretariat (2001).
[9] ISO/IEC Guide 73: 2002, Risk management – Vocabulary - Guidelines for use in standards (2002).
[10] ISO 9000: 2000, Quality management systems – Fundamentals and vocabulary (2000).
[11] ISO/IEC Guide 51: 1999, Safety aspects – Guidelines for their inclusion in standards (1999).

Considering the main area of their application the terms are subdivided into the following four groups: general terms, terms related to risk communication, terms related to risk assessment and terms related to risk control. General terms and terms related to risk communication concern all activities of risk management indicated in Figure 1. Terms related to risk assessment and risk control are primarily applicable in the corresponding areas of risk assessment and risk control.

Mutual links between different terms are illustrated in Figure 1 indicating a framework for risk management and in Figure 2 showing a fundamental flowchart of the risk assessment procedure.

1 GENERAL TERMS

1.1 Hazard: An event or a combination of events with the potential for undesirable consequences.

Note 1: For instance an abnormal action or environmental influence and/or insufficient strength or resistance or an excessive deviation from intended dimensions, in the case of a chemical, the potential that the substance has for causing adverse effects at various levels of exposure. [2].

Note 2: In some documents (for example in the recent draft of EN 1990 [11]) the hazard is defined as an event, while in risk analysis [2] it is considered a condition with a potential for causing an event. Thus, in risk analysis the hazard is a synonym to danger.

1.2 **Hazard scenario:** A sequence of possible events related to a given hazard leading to undesired consequences.
Note: To identify what might go wrong with the system or its subsystem is crucial to a risk analysis. It requires the system to be examined and understood in considerable detail [6].

1.3 **Event**: Occurrence of a particular set of circumstances.
Note: An undesired event is an event, which can cause negative consequences like human fatalities and injuries or environmental damage and economic losses.

1.4 **Probability:** The likelihood or degree of belief of a particular event occurring within a specified reference (time, number of repetitions, etc.).
Note: The probability may depend significantly on the time period during which the particular event may occur.

1.5 **Objective probability:** The probability determined using theoretical arguments or adequate statistical data.

1.6 **Subjective probability:** The probability determined using intuition and relevant experience.

1.7 **Consequence:** The utility assigned to the event in accordance with the preferences of the decision maker.
Note 1: There can be more than one consequence from one event.
Note 2: Consequences can range from positive to negative.
Note 3: Consequences can be expressed qualitatively or quantitatively.

1.8 **Risk:** The expected consequences associated with an activity. Risks may be related to adverse events for humans, qualities of the environment or economic values. In general the risk is the combination of probability of an event and its consequence [7].
Note 1: The risk is often estimated by the mathematical expectation of the consequences of an undesired event. In such case it is the product "probability × consequences". However, a more general interpretation of the risk involves probability and consequences in a non-product form. This presentation is sometimes useful, particularly when a spectrum of consequences, each having its own corresponding probability of occurrence, is considered [9].
Note 2: Various levels of risk may be recognised, for example acceptable risk, tolerable risk and objective risk [9] (see the definition of these terms).

1.9 **Objective risk:** An estimate of the system risk, obtained using theoretical arguments or adequate statistical data (for example the annual expected fatalities from car accidents) or from quantified risk analysis methods (QRA, PRA).

1.10 **Reliability:** The ability of a structure or structural element to fulfil the specified requirements during a given period of time (for example design life).

Note 1: The reliability is often expressed as a probability related to a specific requirement and a period of time [11].
Note 2: In respect of ultimate limit states, the reliability is often referred to as safety; in respect of serviceability limit states, the reliability is often referred to as serviceability [11].

1.11 **Safety:** The state of being protected against hurt or injury, freedom from danger or hazard.
Note: In structural reliability safety is often understood as the reliability with regard to the ultimate limit state (see the definition of Reliability).

1.12 **System**: A bounded group of interrelated, interdependent or interacting elements forming an entity that achieves a defined objective in its environment through the interaction of its parts.
Note 1: This definition implies that the system is identifiable, is made up of interacting elements or subsystems, all elements are identifiable, and the boundary of the system can be identified [2].
Note 2: In terms of technological hazards, a system is normally formed from a physical subsystem, a human subsystem, their management and environment [2].

2 TERMS RELATED TO RISK COMMUNICATION

2.1 **Risk communication:** The exchange or sharing of information about risk between the decision-maker and other stakeholders.
Note: The information can relate to the existence, nature, form, probability**,** severity, acceptability, treatment or other aspects of risk.

2.2 **Stakeholder:** Any individual, group or organisation that can affect, be affected by, or perceive itself to be affected by a risk [7].
Note 1: The decision-maker is also a stakeholder.
Note 2: The term “stakeholder” includes but has a broader meaning than the interested party (which is defined in ISO 9000:2000 [8]).

2.3 **Interested party:** A person or group having an interest in the performance or success of an organisation [9].
Examples: Customers, owners, people in an organisation, suppliers, bankers, unions, partners or society.
Note: A group can comprise an organisation, a part thereof, or more than one organisation.
(ISO 9000: 2000, definition 3.3.7 [8]).

2.4 **Risk perception:** The way in which a stakeholder views a risk, based on a set of values or concerns [7].
Note 1: The risk perception depends on the stakeholders’ needs, issues, knowledge and preferences.
Note 2: The risk perception can be significantly subjective.

2.5 **Criteria of risk:** The reference points against which the results of the risk analysis are to be assessed. The criteria are generally based on regulations, standards, experience, and/or theoretical knowledge used as a basis of the decision on acceptable risk.
Note: Various aspects may be considered, including cultural, social, psychological, economic and other aspects [6]. The acceptance criteria may be expressed verbally or numerically [6].

2.6 **Acceptable risk:** A level of risk, which is generally not seriously perceived by society, and which may be considered as a reference point in criteria of risk.
Note: It is expected that various aspects including cultural, social, psychological, economic and other aspects will influence the risk perception in society (see also the definition of risk criteria).

3 TERMS RELATED TO RISK ASSESSMENT

3.1 **Hazard identification:** A process of recognising the hazard and defining its characteristics.

3.2 **Causal analysis:** A systematic procedure for describing and/or calculating the probability of causes for desired or undesired events.

3.3 **Consequence analysis:** A systematic procedure to describe and/or calculate consequences.

3.4 **Risk analysis:** The use of available information concerning relevant hazard situations for estimating the risk for individuals or populations, property or environment.
Note: The risk analysis generally involves the context (scope) definition, hazard identification, and risk estimation [2].

3.5 **Risk assessment:** A process of risk analysis, risk acceptance and option analysis.
Note: In some documents [11] the risk assessment is defined as risk analysis and risk evaluation, where the risk evaluation covers risk acceptance and option analysis (see the definition of risk evaluation).

3.6 **Risk estimation:** A process used to produce the estimate of the risk measure.
Note: The risk estimation is based on hazard identification and generally contains the following steps: scope definition, probability analysis, consequence analysis, and their integration [2].

3.7 **Risk evaluation:** A process of risk acceptance and option analysis.

3.8 **Sensitivity analysis:** A systematic procedure to describe and/or calculate the effect of variations in the input data and underlying assumptions in general on the final result.

3.9 **Option analysis:** A process used to identify a range of possible alternatives for managing the risk.

4 TERMS RELATED TO RISK MANAGEMENT AND CONTROL

4.1 **Risk management:** The complete process of risk assessment and risk control.
Note: The entire risk management is schematically indicated in Figure 1 (adopted from [10]).

4.2 **Risk treatment:** A process of selection and implementation of measures to modify risk [9].
Note 1: The term "risk treatment" is sometimes used for the measures themselves.
Note 2: The risk treatment measures can include avoiding, optimising, transferring or retaining risk.

4.3 **Safety management:** A systematic process undertaken by an organisation in order to attain and maintain a level of safety that complies with the defined objectives.

4.4 **Tolerable risk:** A level of risk which an individual or society is willing to accept to secure certain benefits, assuming that the risk will be properly controlled.
Note: The tolerable risk may not be negligible, but it should be kept under review and permanent control.

4.5 **Risk control:** Actions implementing risk management decisions.
Note: The risk control may involve monitoring, revaluation, and compliance with decisions.

4.6 **Risk optimisation:** A process, related to a risk, to minimise the negative and to maximise the positive consequences and their respective probabilities [7].
Note 1: In the context of safety, the risk optimisation is focused on reducing the risk.
Note 2: The risk optimisation depends upon risk criteria, including costs and legal requirements.
Note 3: A risk associated with risk control can be considered.

4.7 **Risk reduction:** Actions taken to lessen the probability, negative consequences, or both, associated with a risk [9].

4.8 **Mitigation:** Limitation of any negative consequence of a particular event [7].

4.9 **Risk avoidance:** The decision not to become involved in, or action to withdraw from, a risk situation.
Note: The decision may be taken based on the result of risk evaluation.

4.10 **Risk transfer:** Sharing with another party the burden of loss or the benefit of gain, for a risk [9]**.**
Note 1: Legal or statutory requirements can limit, prohibit or mandate the transfer of a certain risk.
Note 2: The risk transfer can be carried out through insurance or other agreements.
Note 3: The risk transfer can create new risks or modify existing risk.
Note 4: Relocation of the source is not the risk transfer.

4.11 **Risk financing:** Provision of funds to meet the cost of implementing risk treatment and related costs [9].

Note: In some industries, the risk financing refers to funding the financial consequences related to the risk only.

4.12 **Risk retention:** Acceptance of the burden of loss, or the benefit of gain, from a particular risk [9].
Note 1: The risk retention includes the acceptance of risks that have not been identified.
Note 2: The risk retention does not include treatments involving insurance, or transfer by other means.
Note 3: There can be variability in the degree of acceptance and dependence on risk criteria.

4.13 **Risk acceptance:** The decision to accept a risk.
Note 1: The verb "to accept" is chosen to convey the idea that acceptance has its basic dictionary meaning.
Note 2: The risk acceptance depends on risk criteria.

4.14 **Residual risk: A** risk remaining after risk treatment [8].
Note: See also ISO/IEC Guide 51 [9] for safety aspects.

CHAPTER 2: ACTIONS AND INFLUENCES

Mehmet Inel, Sevket Murat Senel and Selcuk Toprak

Pamukkale University, Denizli, Turkey

Summary

Assessment of aging infrastructure and its risk based management for the future must consider available information regarding the actions and their influences on the infrastructure and its components. This chapter provides guidance about the actions and their influences on the structure and its components. Classification of actions and their involvement in the assessment are included.

1 INTRODUCTION

Infrastructures are the key facilities for the sustainability of public services, especially in urban areas. Transportation systems (highways, railways, tunnels, etc.), energy plants and transmission systems, water management structures and communication systems are the typical examples of infrastructures. Hospitals, schools, libraries, theatres, police stations, military centres and court houses are also the examples of infrastructures constructed to meet social and governmental needs of public.

The need for infrastructures is a result of inevitable requirements of urbanization. Water transmission systems in ancient cities such as Alhambra and Rome are the typical examples of first generation infrastructures which were constructed to deliver water to the urban areas. Aging and consequently deteriorating infrastructures are suffered from various actions such as earthquakes, floods, fires, etc. In some cases these actions may cause excessive damages that interrupt the services provided. Fukushima nuclear disaster triggered by 2011 Tohoku earthquake, Japan, failure of Kopru Dam, Turkey in 2012 caused by diversion tunnel broke after a period of heavy rain, Eschede, Germany train disaster in 1998 caused by a single fatigue crack in one wheel, Veligonda railway, India bridge disaster in 2005 caused by the flood action are typical examples of infrastructure failures occurred in recent years.

Estimation of intensity and probability of various actions have been one of the important problems for engineers. Long term measurements and engineering studies have provided valuable information about the actions that should be considered in design. Codes and standards can be treated as resulting documents of these studies and are used to reflect current knowledge level for practicing engineers. On the other hand, it should be reminded that codes and standards are mainly prepared and used for the design of new structures. While assessing the performance of existing structures and/or infrastructures, various factors that can affect the building performance should be considered and adopted. In other words codes or regulations should be reviewed and revised to cover not only design but also assessment purposes.

Infrastructures are the key facilities for the relationship between people and their environment, playing a critical role both in determining vulnerability to particular hazards as

well as directly increasing or decreasing the impact of hazards. Therefore, revision and the adoption of codes in order to cover assessment issues is very important subject for the hazard mitigation in aging infrastructures. Assessment of aging infrastructures and its risk based management for the future must consider available information regarding the actions and their influences on the infrastructures and their components. Table 1 summarizes critical actions and possible influences for existing infrastructure.

Table 1: Critical actions and possible influences for existing infrastructure

Type of Infrastructure	**Critical Action**	**Possible Influence**
Public and Cultural Services Infrastructure (Hospitals, schools, museums, libraries etc.)	Permanent action -indirect action (uneven support settlement) Variable action - Imposed floor loads (unexpected loads due to unnecessary refurnishing) -Seismic loads	Disruption of public and cultural services Potential loss of historical and cultural heritages.
Governmental Services Infrastructure (Police stations, military service centers, courthouses and prisons etc.)	Permanent action -indirect action (uneven support settlement) Variable action - Imposed floor loads (unexpected loads due to unnecessary refurnishing) -Seismic loads Accidental actions -Fire -Explosions	Disruption of governmental services Potential threat for public safety Temporary increase of criminals
Transportation Infrastructures (Highway networks, railways, airports, tunnels, bridges etc.)	Permanent action -self weight due to unexpected surfacing Variable action -Imposed loads due to heavy traffic Accidental actions -Impact from vehicles	Disruption of transportation of people and goods Pollution and environmental disasters
Energy Infrastructure (Electric transmission network, natural gas pipelines, nuclear power plants, dams etc.)	Variable action -Wind load -Snow load Accidental actions -Fire	Temporary power outage Disruption of public services Disruption of telecommunication Decrease in living standards Potential threat for human life
Water Management Systems (Drinking water supply systems, waste water systems, water drainage systems etc)	Variable action -Imposed loads due to heavy traffic	Lack of water Man-made contaminants of water sources or water carrying systems Potential threat for human life

General principles for classification of actions on structures including environmental impacts are provided in existing documents such as Eurocode EN 1990 [1]. Detailed description of individual types of actions is given in various Parts of Eurocode EN 1991 [2]. Actions are sets of forces and they can be classified according to different criteria. EN 1990 classifies actions by their variation in time (permanent, variable or accidental), by their origin (direct or indirect), by their spatial variation (fixed or free) and by their nature and /or the structural response (static or dynamic) as shown in Figure 1.

In most cases, the actions are classified by their variation in time as permanent, variable and accidental actions. These categorizations draw together families of circumstances or conditions that the structure might experience during its life.

Permanent action (G) is defined as action that is likely to act throughout a given reference period with negligible variation in magnitude or monotonic variation until the action attains a certain limiting value while action whose variation in magnitude with time is neither negligible nor monotonic is defined as variable action (Q). Action acting short duration with significant magnitude is termed as accidental action (A). This action is unlikely to occur on a given structure during its design life.

Considering its variation in time and space, self-weight of an infrastructural element is classified as permanent action while imposed load as variable action. However, if the permanency of a self-weight is not clear, then the load shall be treated as variable imposed load. EN 1991 [2] generally treats the imposed load as static load considering dynamic magnification factor. If an imposed load causes significant acceleration on the structure, dynamic analysis shall be applied according to EN 1990 [1].

The most common types of actions for existing infrastructure are given in Table 2. Similar to design situation given in EN 1990 [1], the most critical load combination of permanent, variable and accidental including seismic loading needs to be considered in assessment of existing infrastructures. Also, the most critical load cases should be determined taking into account the most unfavourable influence area of every single action for a given structural element and loading considered. Since permanent loads are on the structure, the most unfavourable influences are primarily due to arrangements of imposed loads. Although this general principle concerns primarily load arrangements of imposed loads, it may concern also self-weight, in particular when structural and non-structural elements or stored and materials may be removed or added.

The classification listed in Table 2 includes general loadings and the actions on the existing infrastructures are not limited to them. Depending on actual loading and site conditions, certain actions, such as seismic actions may be considered as either accidental or variable actions. In some countries or regions of Europe, earthquakes are not rare events and they can be mathematically treated as variable actions.

Table 2 Classification of actions considered in this study

Permanent action	Variable action	Accidental action
(a) Self-weight of infrastructures, fixed equipment, surfacing	(a) Imposed floor loads	(a) Fire
(b) Indirect action, e.g. settlement of supports	(b) Wind loads	(b) Explosions
	(c) Snow loads	(c) Impact from vehicles
	(d) Seismic loads	
	(e) Indirect action, e.g. temperature effects	

EN 1991-1-6 gives principles and general rules for the determination of actions to be taken into account during the execution of buildings and civil engineering works. It may also be used as guidance for the determination of actions to be taken into account during structural alterations, reconstruction and partial or full demolition of existing infrastructures.

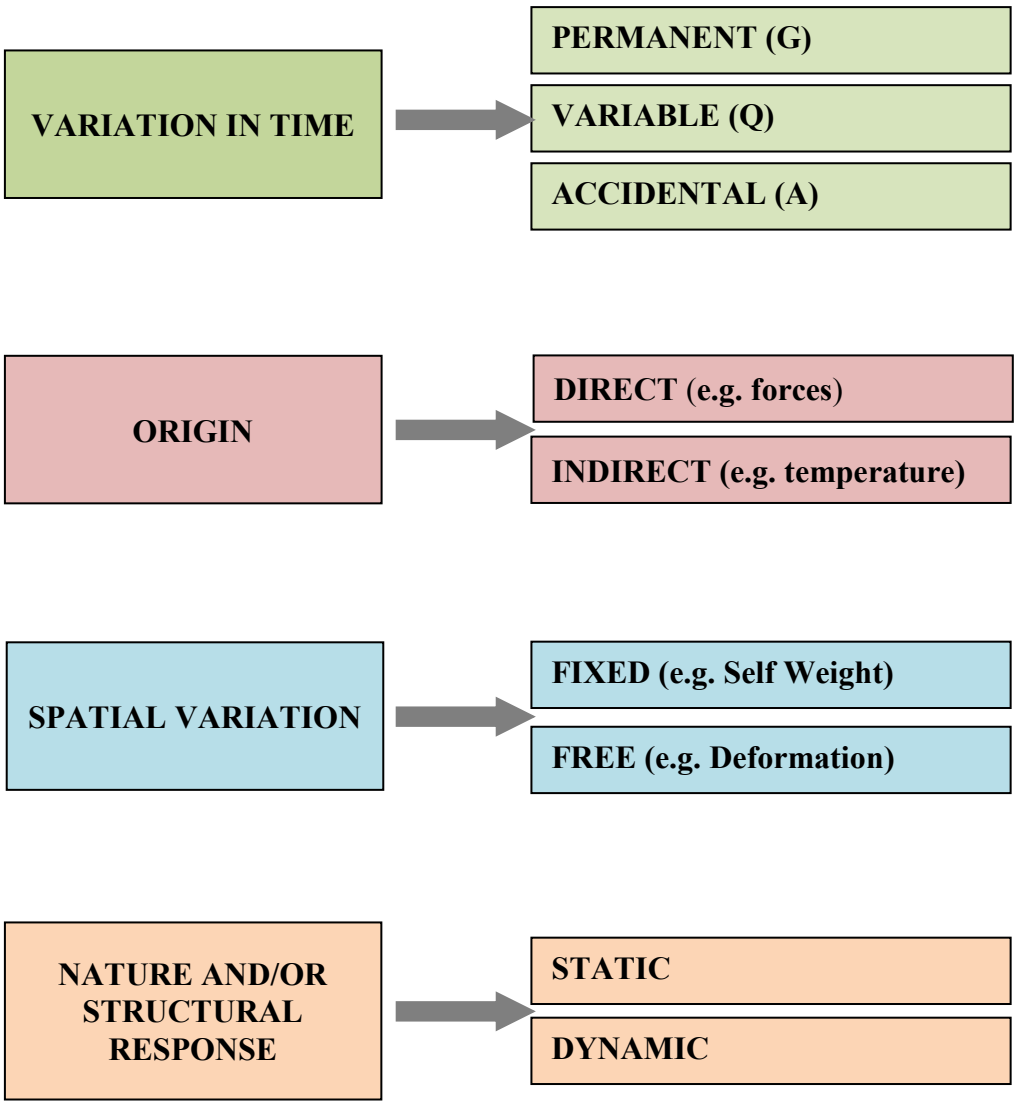


Figure 1 Classification of actions

In the case of seismic actions, it is clear that in some countries or regions of Europe, earthquakes are not rare events and they can be mathematically treated as variable actions. For example, in the particular case of railway bridges, two levels of magnitude may be defined : a moderate level, corresponding to a rather short return period (e.g. 50 years), for which the tracks supported by the bridge deck shall not be damaged, and a rather high level (e.g. corresponding to a 475 years return period), for which the tracks may be damaged, the bridge structure remaining usable (with more or less repair needed). Another example can be found in EN 1991-1-3: Snow loads. This Eurocode allows to treat the snow loads as accidental actions for some cases. This applies for specific climatic regions where the local drifting of snow on roofs is considered to form exceptional snow loads because of the rarity with which they occur, and are treated as accidental loads in accordance with EN 1990: Basis of structural design [1].

2 COMBINATION OF ACTIONS

The effects of actions are determined by combining the values of actions considered by the engineer to be capable of occurring simultaneously for the assessment of an existing structure (e.g. self-weight and imposed loads). The load combinations should consider permanent, variable, accidental and possible seismic loadings. Eurocode 8 [3] provides detailed information for design of structures for earthquake resistance. The load combinations provided in Eurocode EN 1990 [1] can be adapted for the assessment depending on possibility of variable and accidental actions. The input data should be adjusted for actual data of existing structure by updating the values according to the measured ones. The remaining lifetime should be taken into account for the assessment.

3 PERMANENT ACTIONS

A permanent action is defined as an action that is likely to act throughout a given reference period with negligible variation in magnitude e.g. self-weight of structures, weight of fixed equipment and road surfacing, as well as indirect actions caused by shrinkage and uneven settlements. The variability of permanent actions can usually be assumed to be small if the coefficient of variation during the design life is not greater than 0.05 to 0.1 (recommended value in EN 1990 [1]) depending on the type of structure.

3.1 Self-Weight of Construction Elements

The magnitude of self-weight of construction elements is known with greater certainty, as it is closely linked to density and quantity of the construction materials. These have a low variance and the engineer assessing the infrastructure is normally responsible for specifying these components.

Construction elements cover both structural elements (load bearing frames and supporting structures) and non-structural elements including fixed machinery. If it is intended to add or remove structural or non-structural elements after construction, critical load cases need to be identified and taken into account. Non-structural elements shall include surfacing and coverings, partition and linings, hand rails, safety barriers, parapets and curbs, wall cladding, suspended ceilings, thermal insulation, roofing and fixed services. Typical fixed services are equipment for lifts, moving stairways, heating, ventilation and air conditioning equipment, electrical equipment, pipes without their contents etc.

Self-weight of construction elements shall be determined considering nominal dimensions (given in design documentation) and characteristic (nominal) values of densities. If the variability is negligible, a single characteristic value, G_k, mean value, is used. However, two values, an upper and a lower value, have to be used when the variability is not small. For example, the upper and lower characteristic values of densities should be taken into account for ballast on railway bridges or fill above buried structures (see EN 1991-1-1).

The low variability of permanent actions with coefficient of variation ranging from 0.05 to 0.1 is usually intended for self-weight of common buildings. Different values are expected for other structures. For example, bridges, and more particularly long span bridges, the variability of self-weight effects may be in a smaller range, for example between 0.02 and 0.05 [4].

Self-weight G of construction members is usually determined as a product of the volume Ω and the density γas given in Equation (1). Both the volume Ω and the density γ, are random variables that may be described as by normal distribution, with a mean value very

close to their nominal values. Figure 2 illustrates the mean μ_G, lower value $G_{k,lower}$ and upper $G_{k,upper}$, values of self-weight for normal distribution case (see JCSS report [5]). The mean of the volume Ω is approximately equal to the nominal value (as a rule slightly greater), the mean of the density γ is usually well defined by the material producer. Mean values for a large number of different materials are given in EN 1991-1-1, Annex A, as shown in Table 3. In existing infrastructures, the actual self-weight can be specified using samples from the existing structure.

$$G = \Omega \ \gamma \tag{1}$$

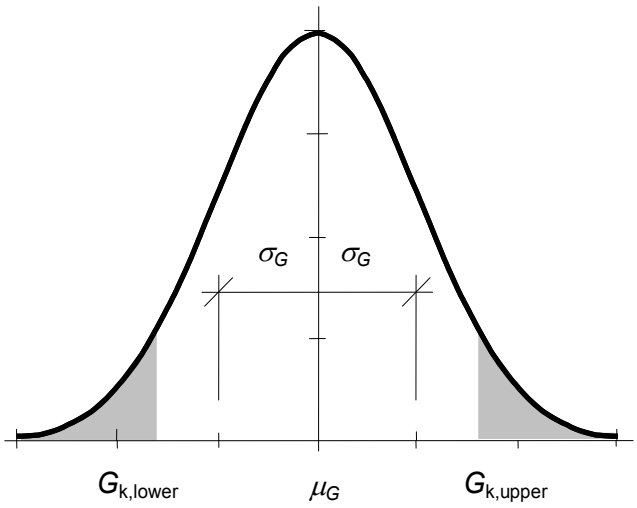


Figure 2 Definition of lower ($G_{k,inf}$) and upper ($G_{k,sup}$) characteristic values of permanent actions based on a normal distribution (adapted from [5])

Table 3 Mean and coefficient of variation (COV) values for weight density [5]

Material	Mean Value (kN/m^3)	COV
Concrete	24	0.04
Steel	77	<0.01
Masonry	-	0.05
Timber	4.4-6.8	0.10

The renovation of an infrastructure or change of building use requires some additional surfacing or elements to the structure resulting in a considerable increase in its self-weight. This is a common case in school buildings or similar public buildings in Turkey. The existing surface elements are not removed and new ones are added during renovation of these buildings. This case sometimes causes additional 5 cm surface thickness in existing public buildings, approximately 1 kN/m^2 additional self-weight of the structure. Sometimes change of purpose of use may cause additional weight or missing details of code. Figure 3 shows collapse of Sampoong department store in Seoul, South Korea on June 29, 1995, mainly due to change of building use originally designed as an office building and redesigned to be a department store when it was halfway through its construction.

Figure 3 Collapse of Sampoong Department Store in Seoul, South Korea (1995)

3.2 Indirect Actions: Uneven Settlements

Settlement in a structure refers to the distortion or disruption of parts of a building due to either; unequal compression of its foundations, shrinkage such as that which occurs in timber framed buildings as the frame adjusts its moisture content, or by undue loads being applied to the building after its initial construction. Settlement should not be confused with subsidence which results from the load-bearing ground upon which a building sits reducing in level, for instance in areas of mine workings where shafts collapse underground.

Some settlement is quite normal after construction has been completed, but unequal (differential) settlement may cause significant problems for buildings. The causes of foundation settlement are rarely due to the design (or under-design) of the structure itself. More commonly, damage is caused as changes occur within the foundation soils that surround and support the structure. For bridges, uneven settlements of the foundations of abutments or piers lead to important deformations and internal forces in the girders. Figure 4 illustrates an example for settlement of a centre pier in the middle of the river settled (about 50 to 80 mm).

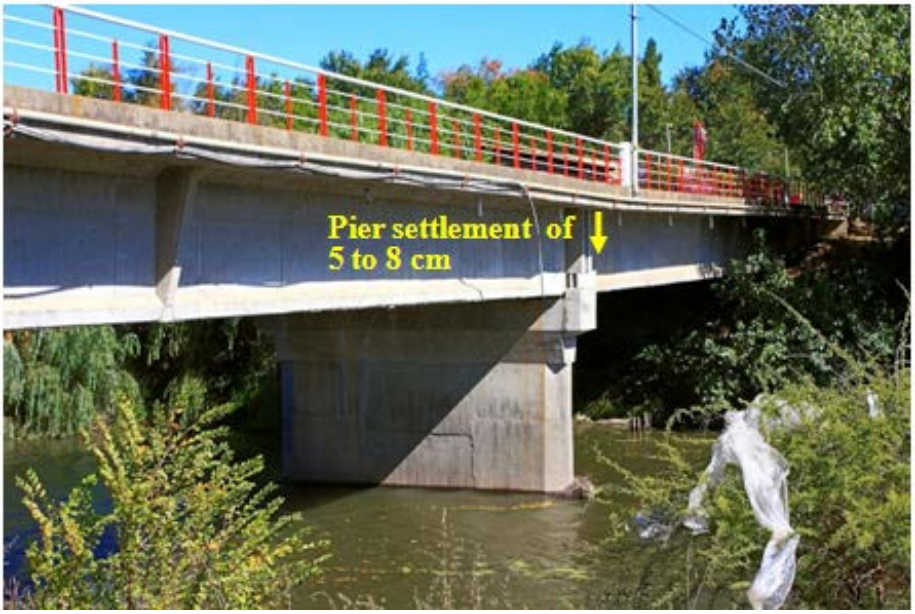


Figure 4 Lateral spreading and ground failure at Raqui 2 Bridge (Chile 2010, [6])

4 VARIABLE ACTIONS

In most cases, actions on infrastructures can be dependent on various factors. The function of use of infrastructure, the region where the structure is constructed and the shape and size of infrastructure can be counted among these factors. Imposed floor loads, snow loads, wind loads, earthquake actions and some indirect actions such as temperature changes are typical examples of variable actions on buildings.

4.1 Imposed Floor Loads

Imposed floor loads are the live loads created by persons, equipment, machines and objects which are stored or attached to infrastructures. These types of loads arise with the structure and act during its use and occupancy. Live loads can be movable and variable depending on the use of infrastructure and sometimes they can also cause dynamic effects. Generally it is suitable to differentiate these loads as sustained and transient loads.

Sustained live loads are generally long term loads caused by the equipment which are used with the structure such as machines, furniture, etc. On the other hand, transient live loads corresponds to short term loads caused by such as heavy duty service vehicles, temporary equipment and persons.

A lot of studies have shown that sustained live loads can be better represented by Gamma distribution. Gamma distribution gives better correlation with the observations with respect to Normal and Log-Normal distribution functions. It is also shown that maximum sustained live load corresponding to a given reference period can be represented by Type I extreme value distribution [7]. Depending on the numerical convenience, Type I extreme value distribution is generally preferred instead of Gamma distribution. Transient live loads, on the other hand, are generally represented by exponential probability distribution function as been suggested in JCSS Report [5]. Total live load is the sum of transient and sustained live loads and it is assumed that the distribution of total live load is also better represented by

Type I extreme value distribution function. Typical values of sustained and transient loads for different usage categories and their details are presented in JCSS Report [5].

4.2 Snow Loads

Various factors are considered for the calculation of snow loads at the roof level of structure. Climate conditions (including win speed and temperature levels during winter) and the altitude of the region where the structure is situated are the examples of these factors. Roof shape and the roofing material are also important parameters that affect the influence of snow loads.

In Eurocode 1-Part 3 (EN 1991-1-3) the value of loads due to snow to be used for the structural design of buildings and civil engineering works are presented. In this code it is stated that Eurocode 1-Part 3 can be applied for the sites altitudes below 1500 m unless otherwise specified. National annexes are referenced for altitudes above 1500 m. In other documents alternative definition of snow loads can be found such as probabilistic model Code prepared by JCSS [5].

In these documents definitions and formulations about the calculation of snow loads at the ground level, conversion of snow loads from ground to roof level and regional coefficients for coastal and mountain regions are provided. In Figure 5 typical relation between snow depth and altitude level based on recorded observations is shown. This figure clearly indicates the interdependency between snow depth and altitude.

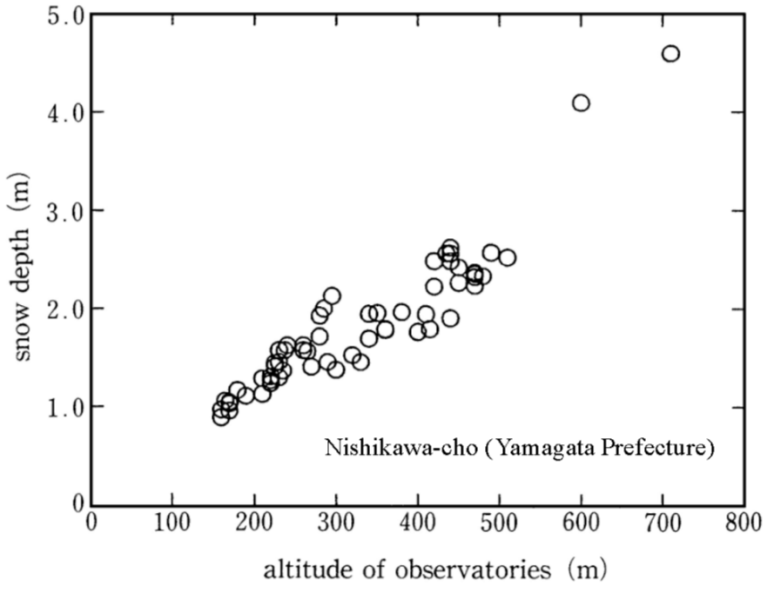


Figure 5 Relation between snow depth and the altitude of observatories [8]

Probabilistic definition of snow load at the ground level is generally represented by two probability distribution functions which define the total duration of loading and the maximum load intensity. These two cases are represented by gamma distribution functions and the parameters of the functions can be determined by using local observations. Conversion of snow loads from ground to roof is affected from various parameters. The exposure of the structure to the wind effect changes the action of snow load. The exposure level of building at the roof level is controlled by the slope and the shape of the roof. In Figure 6 accumulation of snow load at the roof level depending on the roof shape and wind

effect is schematically presented. Thermal effects at the roof level (heated, unheated or isolated roof systems etc.) may be also included in the calculation of snow loads.

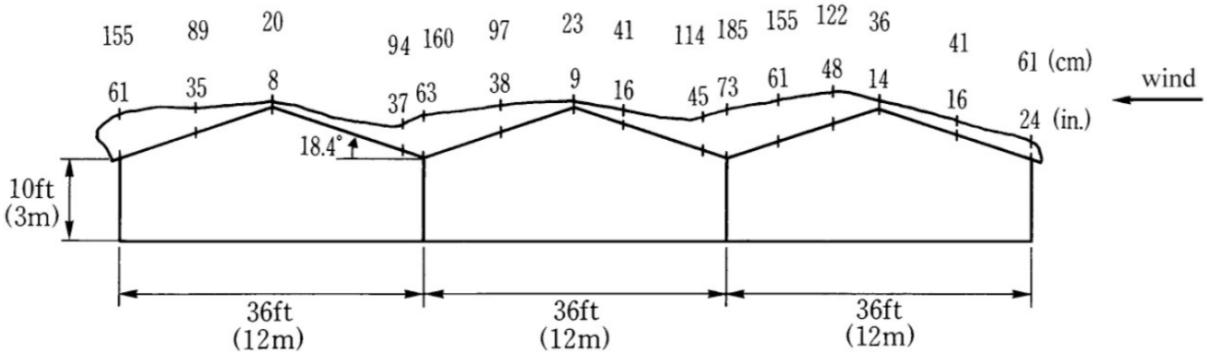


Figure 6 Effects of wind and the roof shape on the accumulation snow loads [9]

However, in some cases codes cannot reflect or formulate all the details of snow actions. Especially aging infrastructures, constructed before modern code regulations, should be assessed to check whether they meet to the requirements of updated regulations or not. It should be reminded that majority of today's infrastructures have been designed using climatic design values derived from historical climate data. Increasing snow actions due to climate changes will require modifications to how infrastructures are engineered, maintained and operated. It should also be considered that structural forms and shapes that can cause the accumulation of snow loads can increase the local damage potential (Figure 7)

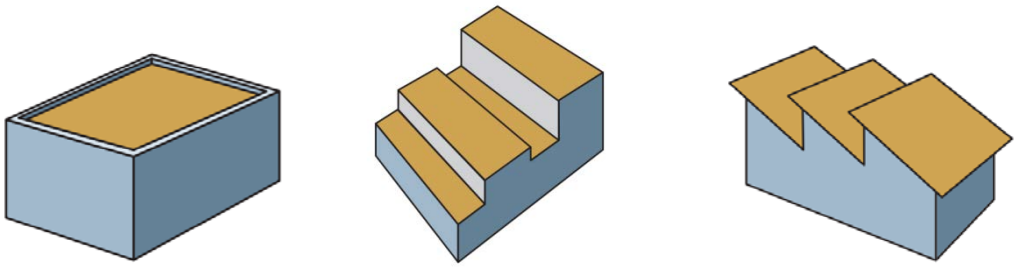


Figure 7 Common roof configurations that affect the severity of snow actions [10]

4.3 Wind Loads

Wind loads on structures are affected by wind climate, shape and the dimensions of structure. In fact wind forces are fluctuating forces. Code recommendations are used to reflect probable maximum fluctuating force effect on infrastructures in terms of static loads. Probabilistic methods are applied to calculate equivalent static load that reflects wind load effect. In Eurocode 1-Part 4 (EN 1991-1-4) natural wind action for structural design of building and civil engineering works is given. The whole structure or parts of the structure or elements attached to the structure such as components, cladding units and their fixings, safety and noise barriers are included in this code.

EN 1991-1-4 is applicable to buildings and civil engineering works with heights up to 200 m. Bridges having no span greater than 200 m, provided that they satisfy the criteria for dynamic response. Alternative approach for the estimation of wind actions can be found in the Probabilistic Model Code prepared by JCSS [5]. In these documents mean wind velocity and it's probabilistic distribution is defined and expressed. During the calculation of mean wind

velocity various factors such as terrain category and roughness, height and the shape of structure e.g. are considered. Then, wind pressure on surfaces is calculated by considering the shape and exposure factors.

However, recent failures occurred in non-building type infrastructures such as wind turbines, solar energy plants and electric transmission systems indicated that code based estimation of wind actions may not provide safer designs for these kind of structures.

In structural engineering calculations dynamic wind loads are assigned to the structures in terms of equivalent static forces. However, dynamic response of flexible structures can strongly be affected from large drift demands caused by wind forces. Time dependent variation of suction and pressure forces also increases the proximity and hence decreases the accuracy of calculations. In Figure 8, steel stadium collapsed in 2009 in Dallas (U.S.A) under wind action is presented. Investigations made by NIST [11] have shown that wind load design of this structure was performed by making enclosed structure assumption. However, excessive deformations during the wind and the tears in fabric membrane surfaces interrupt the enclosure conditions and increase the internal pressures in the facility.

Figure 8 Wind load failure of steel stadium in Dallas

This type of failures shows that great care must be taken in the design of such kind of infrastructures even using the methods presented in the code. Assumptions made during the design (such as enclosed or partially enclosed for internal wind pressures) are vital and can make the difference between a structure standing or collapsing.

4.4 Seismic Actions: Earthquake

Earthquakes are multi-hazard events that have the potential for causing major socioeconomic impacts and losses with little or no warning. In just a matter of seconds an earthquake can cause billions of euros in damage, and leave thousands of people dead, injured, or homeless. Disruption of lifelines, transportation systems, and communication systems can be critical.

The principal geologic hazards associated with moderate- to large-magnitude earthquakes include ground shaking, surface fault rupture and tectonic subsidence, soil liquefaction and related ground failure, landslides, and liquefaction.

Earthquakes can occur anywhere but major earthquakes tend to happen along fault lines in the Earth's crust. Although there are thousands of earthquakes each year only a few damage towns or cities. There is still no reliable way to predict earthquakes despite 40 years

of research, but statistics and geological knowledge can indicate where large earthquakes are likely and what their effects will be.

Records show that some seismic zones experience moderate to major earthquakes approximately every 50 to 70 years, while other areas have recurrence intervals for the same size earthquake of about 200 to 400 years.

Earthquakes are low probability, high consequence events. Although earthquakes may occur only once in the lifetime of a particular asset, they can have devastating effects. Moderate earthquakes occur more frequently than major earthquakes. Nevertheless, a moderate earthquake can cause serious damage unreinforced masonry buildings and buildings non-conforming to existing seismic codes, building contents and non-structural systems, and serious disruption in building operations. Major earthquakes can cause catastrophic damage including collapse and massive loss of life.

Earthquake risk in southern Europe is high; Italy, Greece, Turkey, and Romania are key seismically-exposed countries. The most notable earthquake in European history was the great Lisbon earthquake of 1755 where up to a third of the city's population is estimated to have perished. Even northern Europe has significant loss potential; seismicity is more moderate but more uncertain and insured values are high.

Major infrastructures have experienced serious damages in past earthquakes in regions that were not classified as seismic areas and structures built without seismic provisions. The 2002 Molise earthquake having magnitude of 5.9 is a typical example for such a case; many of the municipalities affected by the earthquake were not classified as seismic areas, and structures were built without seismic provisions. Therefore, the damage exceeded the expected damage in an earthquake of moderate magnitude. The collapse of the primary school Iovene in San Giuliano, caused the death of 27 children and one teacher in the same earthquake, alerting the country to the vulnerability of critical structures. Damage examples from important infrastructures are provided in Figures 9-13. The first three photos are from important public buildings while the others are damages on the highway or secondary road during an earthquake.

No. 3.—The City Police Station.

Figure 9 Collapse of the City Police Station during 1935 Quetta earthquake

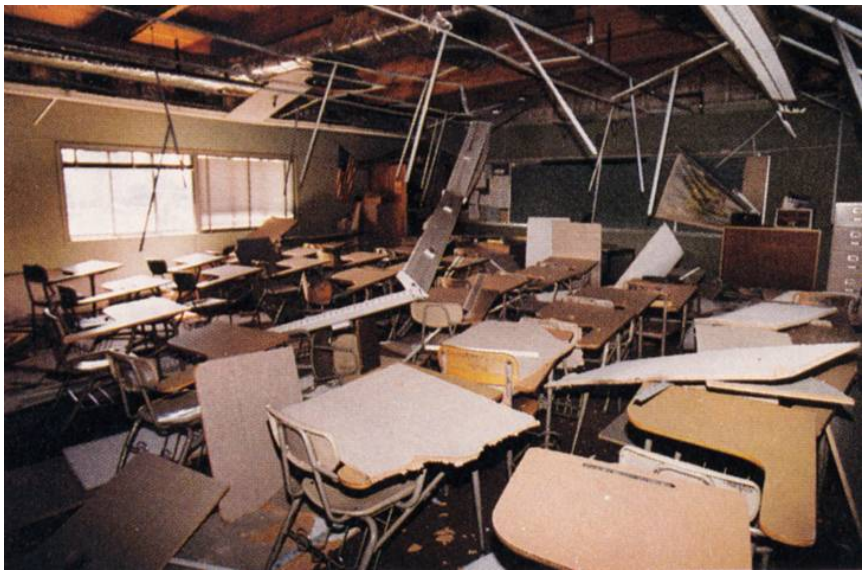


Figure 10 Classroom damage during the 1994 Northridge earthquake

Figure 11 Collapse of Gedikbulak primary school during 2011 Van earthquake

Figure 12 Embankment failure due to liquefaction (Peru 2007)

Figure 13 Retaining wall failure on a secondary road (Peru 2007)

4.4.1 Modelling of Seismic Load

Earthquake loading can be represented by deterministic or probabilistic approaches. Since the deterministic approach requires more detailed site information, the probabilistic approach is more common. JCSS document [5] provides a modelling approach for earthquake loading.

Probabilistic seismic hazard assessment (PSHA) is related to the prediction of the strong motion likely to occur at a given site. The variables for the prediction include:

– Source-to-site distance identification and magnitude model
– Fitting attenuation relationship on peak ground motion parameter
– Determining the peak ground motion parameter at the given site with probability of non-exceedance during a specified period

Recurrence relations for European earthquake vary by seismic source type and have undergone statistical testing for robustness. The model for Turkey employs time-dependent recurrence probabilities to reflect the increased risk of an event since the previous occurrence. Earthquake models in Europe generally use spectral acceleration as the hazard parameter defining shaking intensity.

The usual expression relating to earthquake magnitudes with their occurrence rates, called as magnitude-frequency relation, is provided by Gutenberg and Richter (1944) [12] as given in Equation (2)

$$\text{Log } N\,(M) = a\ b\ M \tag{2}$$

where N(M) is the mean number of earthquakes per unit volume and per unit time having magnitude greater than M. The parameters a and b are zone dependent constants; a depends on the observation period and level of seismicity varying significantly from region to region while b remains within a narrow range around 1.0.

An exponential distribution is suitable for the magnitude M in a region [5] as given in Equation (3).

$$F(M) = \frac{1 - \exp[-\beta(M - M_1)]}{1 - \exp[-\beta(M_2 - M_1)]} \tag{3}$$

$$\beta = b \ln 10 \approx 2.30 \,(\text{if b} = 1) \tag{4}$$

In Equation (3), M_1 is the smallest magnitude considered in the model while M_2 is the largest magnitude event expected in the region. M_2 can also be considered as a random variable or approximately obtained by the following equation.

$$M_2 = (1 + r^{-1})Y_1 - r^{-1}Y_r \tag{5}$$

where r is the number of considered earthquakes, Y_1 is the largest value of magnitude until now and Y_r is the largest value of r earthquakes considered.

Attenuation relations vary by region and by earthquake source type. Attenuation variables include magnitudes, fault mechanisms, focal depths and source-to-site distance. These parameters are reflected in attenuation laws which are based primarily on instrumental and other observed data. There are many attenuation relations in literature. JCSS report [5] provides a general form of attenuation relation for the horizontal peak ground acceleration, PGA in Equation (6) in fraction of g.

$$PGA = b_1 e^{b_2 M} (R + k)^{-b_3} \varepsilon_A \tag{6}$$

In Equation (6), R is the distance between site and source in km, k (in km), b_1(in g), b_2 and b_3 variables provided in Table 4 for different attenuation relations. ε_A is a lognormal error term which takes into account the variability of attenuation law [13-24]. Since no information about residual error is given for the attenuation relations, the values of $\mu(\varepsilon_A)$=1 and the coefficient of variation $V(\varepsilon_A)$ not less than 0.2 can be assumed.

Table 4 Attenuation relations for peak ground acceleration (in g) [5]

Region	**Literature**	**b_1 [g]**	**b_2 [-]**	**b_3 [-]**	**k [km]**	**V (ε_A)**
California USA	Donovan (1973)	1.1	0.51	1.32	25	0.71
California	Mc Guire (1974)	0.48	0.64	1.3	25	0.22
California and Central America	Esteva & Villaverde (1973)	5.7	0.8	2	40	0.64
California [*1]	Campbell (1981)	0.016	0.868	1.09	2.007	0.24
Canada	Milne & Davenport (1969)	0.04	0.99	1.39	0	-
Japan	Katayama & Saeki (1978)	0.02	0.7	0.8	0	-
Taiwan	Mau & Kao (1978)	0.38	0.876	1.836	20	-
South Africa	Donovan & Bornstein (1977)	0.073	0.756	1.01	25	-
Central Europe	Ahorner & Rosenhauer (1975)	1.28	0.8	2	13	-
Greece	Makropoulos (1978)	2.2	0.7	1.8	20	0.5
Switzerland [*1]	Ziegler (1985)	0.016	0.868	0.95	2.649	-
Italy [*2]	Sabetta & Pugliese (1987)	0.014	0.363	0.05	25	0.2

[1] R is epicentral distance
[2] $PGA = b_1 e^{b_2 M} (R^2 + k)^{-b_3} \varepsilon_A$ is used.

Recent attenuation relations include more parameters. Although there are many attenuation relations, two of the most used relations are given below as Bommer (1996) [25] and Campbell (1997) and & Campbell & Bozorgnia (1994) [26, 27]. Ground motion model by Bommer is given in Equation (7).

$$\ln PGA = a + bM + d \ln R + qh \tag{7}$$

where h is focal depth, PGA is in g, $a = -1.47$, $b = 0.608$, $d = -1.181$, $q = 0.0089$ and $h = 0.54$.

Horizontal component of ground motion model by Campbell is obtained:

$$\begin{aligned} \ln PGA = {} & a_1 + a_2 M + a_3 \ln \sqrt{R^2 + \left[a_4 \exp(a_5 M)^2\right]} + \left[a_6 + a_7 \ln R + a_8 M\right] F \\ & + \left[a_9 + a_{10} \ln R\right] S_{SR} + \left[a_{11} + a_{12} \ln R\right] S_{HR} \end{aligned} \tag{8}$$

where PGA is in g, $a_1 = -3.512$, $a_2 = 0.904$, $a_3 = -1.328$, $a_4 = 0.149$, $a_5 = 0.647$, $a_6 = 1.125$, $a_7 = -0.112$, $a_8 = -0.0957$, $a_9 = 0.440$, $a_{10} = -0.171$, $a_{11} = 0.405$, $a_{12} = -0.222$. SSR = 0, S_{HR} = 1 for hard rock, S_{SR} = 1, S_{HR} = 0 for soft rock and S_{SR} = 0, S_{HR} = 0 for alluvium or firm soil. F corresponds to coefficient for fault mechanism. F=0 strike-slip, F=1 reverse, reverse-oblique, thrust, and thrust-oblique and F = 0.5 (recommended) for normal or unknown mechanisms.

4.4.2 Seismic Loading on Buried Pipelines

Earthquake damage to buried pipelines can be attributed to transient ground deformation (TGD) or to permanent ground deformation (PGD) or both. TGD occurs as a result of seismic waves and often stated as wave propagation or ground shaking effect. PGD occurs as a result of surface faulting, liquefaction, landslides, and differential settlement from consolidation of cohesionless soil. The effect of earthquake loading on pipelines can be expressed in terms of axial and flexural deformations. At locations where the pipeline is relatively weak because of corrosion, etc., breaks and/or cracks may be observed on the pipelines. If deformations are high, the damages can be in the form of separations of joints, wrinkling, buckling and tearing of pipelines.

Pipeline damage due to TGD effects is correlated with various seismic parameters whereas damage due to PGD effects is correlated with amount of ground movement or deformation. Pipeline damage commonly expressed as repair rate, which is the number of pipeline repairs in an area divided by the length of the pipelines in the same area. Toprak (1998) [28] evaluated pipeline damage correlations developed before the 1994 Northridge earthquake, USA, and used various seismic parameters to explore new relationships between seismic intensity and pipeline damage. The seismic parameters that he tried are Modified Mercalli Intensity (MMI), peak ground acceleration (PGA), peak ground velocity (PGV), peak ground displacement, spectral acceleration, spectral velocity, spectrum intensity, and Arias intensity. Among the various seismic parameters, the most statistically significant correlations were found for PGV (Toprak, 1998 [28]; O'Rourke et al., 1998 [29]). PGV has a more direct physical interpretation in terms of its effects on buried pipelines. PGV is correlated with axial strains experienced in the soil due to seismic wave propagation as expressed in the following general equation (Committee on Gas and Liquid Fuel Lifelines, 1984 [30]):

$$\varepsilon_g = \frac{V_{\max}}{C} \tag{9}$$

where ε_g is the maximum seismic ground strain, $V_{\max}$ is the maximum ground velocity, and C is the seismic wave-propagation velocity. Depending on the slippage developed between a pipeline and the surrounding soil, a certain percentage of the soil strain is transferred to the pipeline. Because of this relationship, a good correlation between PGV and pipeline damage would be expected. If PGD and TGD are similar in magnitude, they will contribute to pipeline damage at comparable levels. Large TGD is often observed when there are near-field large peak ground velocity pulses or where site response characteristics result in amplified transient motion.

Toprak et al. (2014) provided a comprehensive literature review and presented new developments on seismic response of underground lifeline systems in their theme lecture at the Second European Conference on Earthquake Engineering and Seismology (2ECEES) [31]. Some standards or guidelines which can be used in the assessment of pipelines are European Standard EN 1993-4-3 (Eurocode 3: Design of steel structures - Part 4.3: Pipelines) [32], Eurocode 8, Part 4 (Design of structures for earthquake resistance -Part 4: Silos, tanks and pipelines) [33], ISO 16134: 2006 (Earthquake- and subsidence-resistant design of ductile iron pipelines) [34], and ALA (2001) (Seismic fragility formulations for water systems) [35]. Applications on pipelines will be provided in Handbook 2 of this project.

4.5 Traffic Loads

Loads due to the road traffic, consisting of cars, lorries and special vehicles (*e.g.* for industrial transport), give rise to vertical and horizontal, static and dynamic forces. Traffic load models adopted in practicing codes for bridges have significantly evolved. Traffic flow information and statistics for European countries play an important role in modelling. Daily traffic for cars and lorries, and details of lorries are necessary for proper modelling.

In modern codes, static models, representing physically existing heavy vehicles, have been replaced by ideal models in order to reproduce the target values of the effect induced in the bridges by the real traffic. EN 1991-2 [36] includes remarkable traffic load models.

Traffic loads are critical issue for bridges, especially for the old ones. Traffic loads have significantly increased as the needs of human being and technology increased. Thus, current traffic loads may not be considered for a 100 years old bridge during its design phase. While assessing the performance of existing bridges or infrastructures with traffic loads, the current codes of practice should be considered and adopted. Figure 14 shows a photo of The Lake County Highway (SR 29) Bridge in Napa County, California is a three span stone arch that was built in 1902 and registered as historic in 2005. Erosion is damaging the bridge foundations and arches, heavy truck traffic is damaging the deck, and the bridge is too narrow to safely carry two lanes of traffic. Figure 15 illustrates another photo of damaged bridge due to heavy trucks. The heavy traffic also damages the road surfaces as shown in Figure 16. The last two pictures are taken from a road on a gas well site in One Denton, Texas. In some cases, imposed load and traffic load can be controlled by imposing restrictions on the existing structures [37].

Figure 14 Lake County Highway (SR 29) Bridge (taken from http://www.bphod.com/2014/01/napa-county-california-bridges-garnet.html)

Figure 15 Damaged bridge due to heavy trucks (http://www.marcellus-shale.us/road_damage.htm)

Figure 16 Pulverized blacktop road due to heavy traffic (http://www.marcellus-shale.us/road_damage.htm)

4.6 Indirect Variable Actions: Thermal Effects

EN 1990 [1] provides guidance for relevant design situations for thermal actions. Similarly, the assessment of existing infrastructure shall check structural elements to ensure that thermal movement does not cause overstressing of the structure, either by the provision of sufficient gaps for joint movements or by including the effects in the assessment.

The degree of the thermal effects depends on local climatic conditions, the orientation of the structure and its overall mass. In the case of building structures, the thermal effects also depend on heating and ventilation regimes and thermal insulation.

The temperature distribution within an individual structural element is decomposed into the following four essential constituent components as shown in Figure 17:

(a) A uniform temperature component, ΔT_u
(b) A linearly varying temperature difference component about the z-z axis
(c) A linearly varying temperature difference component about the y-y axis
(d) A self-equilibrating non-linear temperature difference component, ΔTE

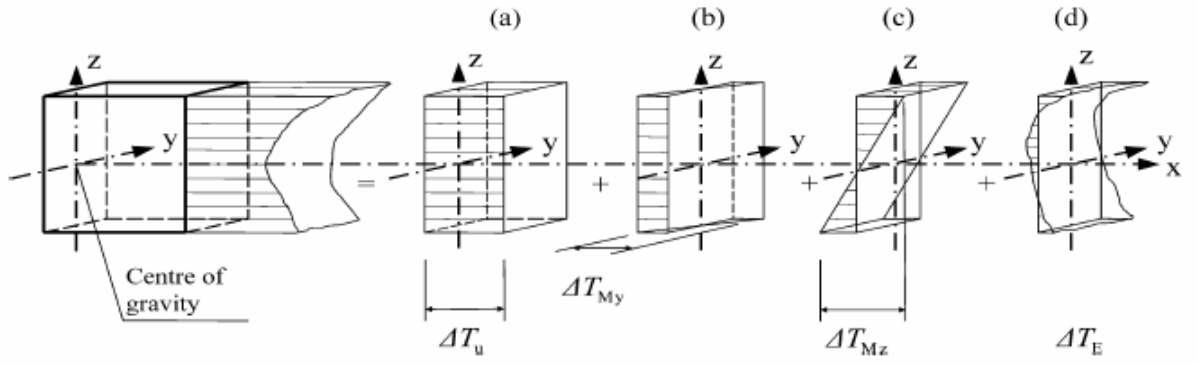


Figure 17 Representation of constituent components of a temperature profile

Regulations provide characteristic values whose probability of being exceeded is 0.02 corresponding to a return period of 50 years. The fundamental quantities on which thermal actions are based are the extreme air temperatures value, that is, the maximum and minimum, in the shade at the building site. Such values are furnished by the national meteorological service of each member state. The shade air temperature is measured by a device known as a "Stevenson Screen", which is simply a thermometer set in a white painted wooden box with louvered sides. Eurocode 1991-1-5 [2] does not include maps for extreme temperature determinations: such task is left up to the national meteorological services.

The uniform temperature component of a structural element ΔT_u is defined as:

$$\Delta T_u = T - T_0 \tag{10}$$

In Equation (10) T is an average temperature of a structural element due to climatic temperatures in winter or summer season and due to operational temperatures while T_0 is the temperature of a structural element at the relevant stage of its restraint (on completion). When elements of one layer are considered and when the environmental conditions on both sides are similar, T may be approximately determined as the average of inner and outer environment temperature T_{in} and T_{out}. Actual values should be updated according to measurements on existing infrastructures.

5 ACCIDENTAL ACTIONS

General principles for classification of actions including accidental actions are given in EN 1990 [1] while a detailed description of individual actions is given in various parts of Eurocode 1, EN 1991 [2]. Accidental actions are covered in Part 1.7 of EN 1991, including guidelines how to handle accidental loads.

5.1 Impact from Vehicles

In mechanics, an impact is a high force or shock applied over a short time period when two or more bodies collide. Such a force or acceleration usually has a greater effect than a lower force applied over a proportionally longer time period of time. The effect depends critically on the relative velocity of the bodies to one another.

At normal speeds, during a perfectly inelastic collision (ship impact on existing bridges), an object struck by a projectile will deform, and this deformation will absorb most, or even all, of the force of the collision. Viewed from the conservation of energy perspective, the kinetic energy of the projectile is changed into heat and sound energy, as a result of the deformations and vibrations induced in the struck object. However, these deformations and vibrations cannot occur instantaneously. A high-velocity collision (an impact) does not provide sufficient time for these deformations and vibrations to occur. Thus, the struck material behaves as if it were more brittle than it is, and the majority of the applied force goes into fracturing the material.

In the case of hard impact, the values for horizontal actions due to impact on vertical structural elements may be obtained from Table 5. The forces $F_{d,x}$ and $F_{d,y}$ correspond to the forces in driving and perpendicular directions, respectively [38]. In existing structures, it is advised to find actual impact loading by considering site specified conditions.

Table 5 Horizontal equivalent static forces due to impact on supporting vertical structural elements [38]

Type of road	Type of vehicle	Force $F_{d,x}$ (kN)	Force $F_{d,y}$ (kN
Motorway	Truck	1000	500
Country road	Truck	750	375
Urban area	Truck	500	250
Courtyards	Passengers cars only	50	25
Courtyards	Truck	150	75

Consider a structural element in the vicinity of a road or track. Impact on the structural element occurs if some vehicle, travelling over the track, leaves its intended course at some critical place with sufficient speed (see Figure 18 taken from [5]).

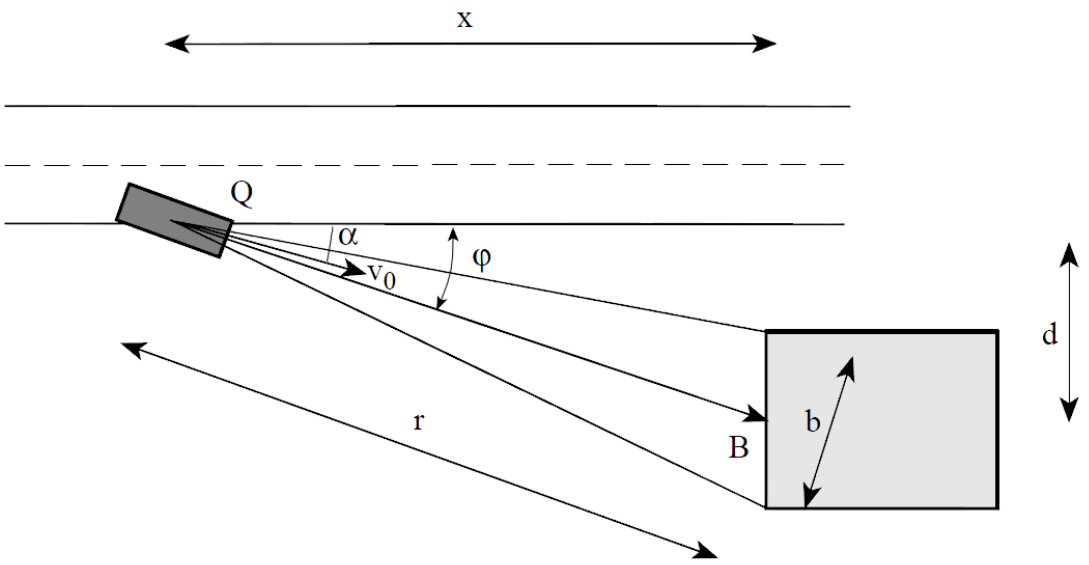


Figure 18 Impact of a vehicle leaving the intended course at point Q with velocity v_0 and angle α on a structural element at distance r that is hit with velocity v_r [5]

The collision force probability distribution based on Figure 18, neglecting the variability in y direction is given by:

$$P(F_c > X) = nT\lambda\Delta x P\left[\sqrt{mk(v_0^2 - 2ar)} > \mathrm{X}\right] \tag{11}$$

n = number of vehicles per time unit
T = period of time under consideration
λ= probability of a vehicle leaving the road per unit length of track
Δx = part of the road from where collisions may be expected
v_o = velocity of the vehicle when leaving the track
a = deceleration
r =d/sinα = the distance from "leaving point" to "impact point"
d = distance from the structural element to the road
α = angle between collision course and track direction
$\lambda\, \Delta x$ is the probability that a passing vehicle leaves the road at the interval Δx, which is approximated by

$$\Delta x = \frac{b}{\sin \mu(\alpha)} \tag{12}$$

Numerical values and probabilistic models can be found in Table 6 [5]. The value of b depends on the structural dimensions. However, for small objects such as columns a minimum value of b follows from the width of the vehicle, so b > 2.5 m. Similar to other loads, actual load should be updated according to measurement on the bridges.

Table 6 Numerical values for vehicle impact [5]

variable	designation	type	mean	stand dev
λ	accident rate	deterministic	10^{-10} m^{-1}	-
α	angle of collision course	rayleigh	10 °	10°
v	vehicle velocity - motorway - urban area - court yard - parking garage	 lognormal lognormal lognormal lognormal	 80 km/hr 40 15 10	 10 km/hr 7 6 5
a	deceleration	lognormal	4 m^2/s	1.3 m/s^2
m	vehicle mass - truck - car	 normal normal	 20.000 kg* 1500 kg	 12.000 kg* 400 kg
k	vehicle stiffness	lognormal	300 kN/m	60 kN/m
*Combined with F = k$\sqrt{}$mv these estimates are quite conservative. One might consider possible reductions due to transformation of energy into rotational movements, etc. e.g. by the concept of "effective mass"				

5.2 Fire Actions

Fire safety of infrastructures is an important issue for the sustainability of goods and services. Damage risk on buildings due to fire is covered in the codes. However, additional civil engineering studies should be taken into consideration for infrastructures such as bridges, tunnels, etc. In Eurocode 1- Part 2 (EN 1991-1-2) there are specifications which regulate the structural fire design rules for concrete, steel, masonry and timber structures. These regulations deal with the load bearing capacity of elements and structure. Whether or how long a structure can withstand to the fire is the main question. Therefore, ultimate limit state is considered in design.

Fire design of structures according to Eurocode can be summarized as shown in Figure 19 [39]. Fire is an accidental action and the therefore verification of structure against fire should be performed under ultimate limit states rather than serviceability limit state. For this reason, cases that can be related with the total or partial collapse of infrastructure (such as excessive deformations in members and system, loss of stability, formation of mechanism) due to fire should be discussed.

Characteristic of fire loads for various occupancy classes are given in EN1992-1-2. In this code fire growth rate and fire load values are tabulated for building type infrastructures such as hospitals, schools, shopping centres, hotels, etc. However, especially for non-building type infrastructures definition and formulation of fire actions becomes much more complex. Life lines in urban areas such as natural gas, energy and water transmission systems are the examples of such kind of facilities. Fire in such kind of infrastructures can interrupt the critical public services. Space and water heating, wired and wireless communication and electric supply services are the examples of these critical public services.

In some cases fire can be the side effect of other dangerous actions such as earthquakes, explosions and accidents. Fires reported after 1994 Northridge (U.S.A), 1999 Kocaeli (Turkey) and 2011 Tohoku (Japan) earthquakes are the examples of such kind of fires. It should be noted that during these fires a lot of infrastructures were damaged and out of order for operation.

It is clear that expecting the estimation of deterministic formulation of fire action by using the codes is impossible. On the other hand critical issues that must be considered for the assessment of fire safety of aging infrastructures can be defined.

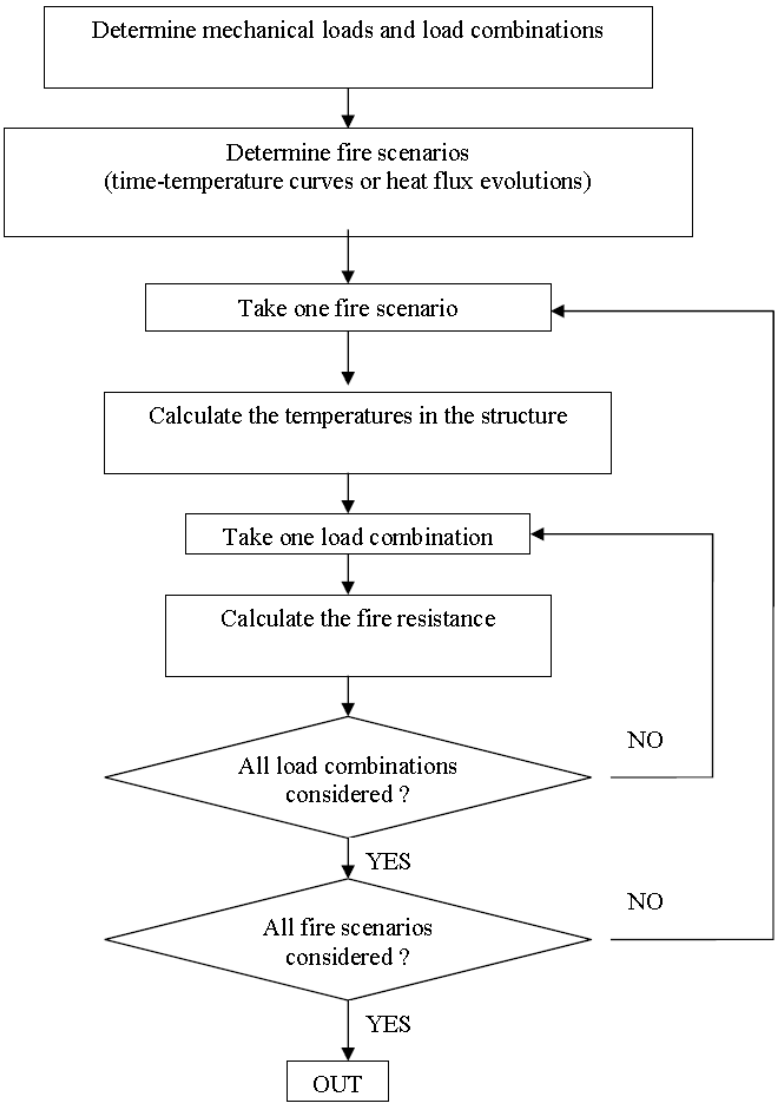


Figure 19 Analysis steps for structures against fire

REFERENCES

[1] EN 1990 Basis of structural design. European Committee for Standardisation, CEN/TC 250 (April 2002).

[2] EN 1991 Actions on structures, Part 1.1 Densities, self-weight and imposed loads on buildings. European Committee for Standardisation, CEN/TC 250 (April 2002).

[3] Eurocode 8: Design of structures for earthquake resistance, Part 1: General rules, seismic actions and rules for buildings, EN1998-1-2004. European Committee for Standardization, CEN, Brussels (2004).

[4] Gulvanessian H., Calgaro J.-A. and Holický M. (2002) Designer's guide to EN 1990, Eurocode: Basis of structural design. Thomas Telford, London.

[5] JCSS: Probabilistic model code. JCSS working materials, http://www.jcss.ethz.ch/ (2001).

[6] Post-earthquake reconnaissance report on transportation infrastructure: Impact of the February 27, 2010, Offshore Maule Earthquake in Chile, U.S: Department of Transportation, Federal Highway Administration, FHWA-HRT-11-030 (2011).

[7] Melchers R. E. (1987) Structural reliability analysis and prediction. Ellis Horwood Limited. Chichester, John Wiley and Sons.

[8] Takahashi T., Fukaya M., Mihashi H. and Izumi M. (1992) Influence of altitude and sea area ratio for geographic distribution of snow depth, Summaries of technical papers of annual meeting, Vol. B, AIJ

[9] Taylor D. A. (1980) Roof snow loads in Canada, *Canadian Journal of Civil Engineering*, Vol.7, No.1, pp.1 - 8.

[10] Snow load safety guide. FEMA P-957, FEMA Risk management series (January 2013)

[11] Gross J. L. et al. (January 2010). Final report on the collapse of the Dallas Cowboys indoor practice facility, May 2, 2009. National Institute of Standards and Technology (NIST).

[12] Gutenberg B. & Richter C. (1944) Frequencies of earthquakes in California, *Bull. Seism. Soc. Amer.* (34), pp. 185-188.

[13] Donovan N. C. (1973) A statistical evaluation of strong motion data including the February 9, 1971 San Fernando earthquake. Proceedings of the 5th World Conference on Earthquake Engineering (Vol. 1), pp 1252–1261.

[14] McGuire R. K. (1974) Seismic structural response risk analysis, incorporating peak response regressions on earthquake magnitude and distance. Massachusets Inst. of Technol., Dep. of Civ. Eng., Res. Rpt. R74-51.

[15] Esteva L. & Villaverde R. (1973) Seismic risk, design spectra and structural reliability, Proc. 5th WCEE, Rome, pp. 2586-2597.

[16] Campbell K. W. (1981) Near-source-attenuation of peak horizontal acceleration, *BSSA*, Vol.71, No.6.

[17] Milne W. G. & Davenport A. G. (1969) Distribution of earthquake risk in Canada, *Bull. Seism. Soc. Amer*. (59), No. 2, pp. 754-799.

[18] Katayama T. & Saeki M. (1978) Statistical analysis of earthquake acceleration response spectra, Proc. of ISCE, No. 275 (in Japanese).

[19] Mau S. T. & Kao C. S. (1978) A risk model for seismic zoning in Taiwan, Proc. of the 2nd Intern. Conf. on Microzonation, Vol. I, San Francisco, pp. 367-378.

[20] Donovan N. C. & Bornstein A. E. (1977) The problems of uncertainties in the use of seismic risk procedures in use of probabilities in earthquake engineering, ASCE, Special Publication, pp. 1-36.

[21] Ahorner L. & Rosenhauer W. (1975) Probability distribution of earthquake accelerations with applications to sites in the northern Rhine area, Central Europe, *Journ. Geophys*. (41), pp 581-594.

[22] Makropoulos K. (1978) The statistics of large earthquake magnitude and an evaluation of Greek seismicity, PhD thesis, University of Edinburgh, Scotland.

[23] Ziegler A. (1985) Bemessungsbeben für Stauanlagen im schweizerischen Alpenraum, Abhandlung zur Erlangung des Titels eines Doktors der Technischen Wissenschaften

der Eidgenössischen Tecnischen Hochschule Zürich, Diss. ETH Nr. 7767, Schweiz.
[24] Sabetta F. & Pugliese A. (1987) Attenuation of peak horizontal acceleration and velocity from Italian strong-motion records. *Bull. Seismol. Soc. Amer.* (77), pp. 1491-1523.
[25] Bommer J. J., Hernandez D. A., Navarrete J. A. & Salazar W. M. (1996) Seismic hazard assessments for El Salvador. *Geofisica Internacional*, 35 (3), pp. 227-244.
[26] Campbell K. W. (1997) Empirical near-source attenuation relationships for horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra. *Seismological Research Letters*, 68 (1), pp. 154-179.
[27] Campbell K. W., Bozorgnia Y. (July 1994) Near-source attenuation of peak horizontal acceleration from worldwide accelerograms recorded from 1957 to 1993. Proceedings of the 5th U.S. National Conference on Earthquake Engineering, Vol. III., pp 283-292*:*
[28] Toprak S. (1998), Earthquake effects on buried lifeline systems, Ph.D. thesis, Cornell University, Ithaca, NY.
[29] O'Rourke T. D, Toprak S., Sano Y (1998) Factors affecting water supply damage caused by the Northridge Earthquake, Proc., 6th U.S. National Conf. on Earthquake Engineering, EERI, Oakland, CA.
[30] American Society of Civil Engineers, ASCE (1984) Guidelines for the seismic design of oil and gas pipeline systems, Committee on Gas and Liquid Fuel Lifelines.
[31] Toprak S., Nacaroglu E., Koc A. C. (2014) Seismic response of underground lifeline systems, The 2nd European Conference on Earthquake Engineering and Seismology (2ECEES).
[32] EN 1993-4-3, 2007, Eurocode 3: Design of steel structures - Part 4-3: Pipelines.
[33] EN 1998-4:2006, Eurocode 8: Design of structures for earthquake resistance – Part 4: Silos, tanks and pipelines.
[34] ISO 16134:2006, Earthquake - and subsidence-resistant design of ductile iron pipelines.
[35] ALA 2001, 2005, Seismic guidelines for water pipelines, prepared by ASCE, FEMA and NIBS.
[36] EN 1991-2: Eurocode 1: Actions on structures - Part 2: Traffic loads on bridges (September 2003).
[37] Croce P., Holicky M., Markova J., Arteaga A., Diego A., Tanner P., Lara C., Diamantidis D., Vrouwenvelder T. (2010), Guidebook 2, Design of bridges, Czech Technical University in Prague, Klokner Institute.
[38] Development of skills facilitating implementation of Eurocodes, Leonardo da Vinci Pilot Project CZ/02/B/F/PP-134007, Handbook 4: Design of bridges.
[39] Introduction to fire design (2008), Raul Zaharia, http://www.ct.upt.ro/users/RaulZaharia/Introduction_fire_design.pdf.

CHAPTER 3: DEGRADATION MODELLING

Carlos Lara, Peter Tanner

Eduardo Torroja Institute for Construction Sciences, IETcc-CSIC, Madrid, Spain

Summary

During their service life, infrastructures are subject to numerous environmental influences which over time may cause deterioration of their materials. Degradation usually influences both, resistance and appearance of an infrastructure. In order to efficiently assess and manage aging infrastructures, the most relevant, physical and chemical, degradation mechanisms of infrastructure systems are discussed in this chapter. Degradation models for construction materials such as reinforced concrete, structural steel and masonry are provided. Inspections may be used as a tool to reduce the uncertainty in the predicted deterioration process. Some hints are therefore presented about a risk-based approach for inspection and maintenance planning of deteriorating infrastructures.

1 INTRODUCTION

1.1 Context

The structural assessment of an existing infrastructure and its risk-based management for the future use is based on the available information regarding the actions on the structure, structural behaviour and the resistance of its components. Usually established structural analysis methods are applied and the degree of detail of the analysis depends on the stage of the assessment process which has been described in the previous chapter.

According to the risk assessment process reported in Chapter 1, the investigation involves the acquisition of all relevant information concerning:

- The original design and structural concept of the infrastructure, as well as the code rules used, if any
- The sequence of structural modifications during its previous service period, addition or demolition of structural parts and/or deep maintenance interventions
- Actual material properties
- Actual damage and degradation
- Required performance level

Time is a decisive factor within the framework of assessment and management of infrastructures, influencing residual service life, maintenance, inspection, repair or the life-cycle cost. The reliability level for important limit states has to be analysed frequently utilising mathematical models for prediction of degradation.

During their service life, infrastructures are subject to numerous environmental influences (e.g., moisture, temperature, chemical substances, biological processes, etc.) which, over time, may cause deterioration of their materials. Degradation usually influences

both, resistance and appearance of an infrastructure. Serviceability and safety of the structure may be seriously compromised by degradation mechanisms. In order to efficiently assess and manage aging infrastructures, models predicting the initiation and propagation of deterioration processes are needed.

Modelling the deterioration process involves significant uncertainties associated with environmental conditions, material properties, limitations of predictive models, and inadequacy of material testing, detection, and inspection methods [1]. Therefore, deterioration processes acting on infrastructures are best described in probabilistic terms. The basic concepts and procedures of the probabilistic theory of reliability are provided in Chapter 5.

The present chapter gives a short review of degradation mechanisms and the corresponding models which may be applied within the framework of structural assessment methods and procedures for aging infrastructures. Initially, a classification of degradation processes according to their nature (physical or chemical) is included (Section 2). Following this classification, probabilistic models for the most relevant degradation mechanisms for reinforced concrete (RC), structural steel, and masonry are presented in Sections 3, 4 and 5, respectively.

Inspections may be used as a tool to reduce the uncertainty in the predicted deterioration process and/or as a mean of identifying deterioration before it becomes critical. A risk-based approach for inspection and maintenance planning of deteriorating infrastructures is briefly presented in Section 6. Case studies about modelling of deterioration processes are illustrated in Handbook 2.

1.2 Background Documents

Degradation processes described in this chapter have been classified following recommendations of the International Standard ISO 13823 [2], which specifies general principles and recommends procedures for the verification of the durability of structures subject to known or foreseeable environmental actions, including mechanical actions, causing material degradation. The presented models for deterioration processes affecting RC infrastructures are in accordance with models described in structural codes such as Eurocode 2 [3] and Fib Model Code 2010 [4]. Similarly, the presented models for deterioration processes affecting steel infrastructures are in accordance with rules and models adopted in Eurocode 3 [5]. Other International Standard such as [6] and [7] provide general principles and guidance for the assessment and management of existing infrastructures. In addition the JCSS Probabilistic Model Code [8] offers supplementary information about probabilistic models for resistance and environmental variables that can be used in practical applications.

2 CLASSIFICATION OF DETERIORATION PROCESSES

Deterioration processes in infrastructures depend on the interaction between the system and its environment. The importance of different environmental parameters in this interaction depends on the particular deterioration process considered. Environmental parameters which affect all (or most) of the deterioration processes include [1]:

– Temperature
– Moisture/humidity
– Wind, solar radiation

Amongst the most significant effects of temperature variation belong the deformations induced in materials. These deformations can result in damage to components and loss of performance. Although most chemical processes are accelerated by an increase in temperature, some are accelerated by a decrease. This may occur in the corrosion of metals, where low surface temperatures promote condensation on the surface, providing an environment favourable to corrosion. However, if the surface temperature is high, drying of the surface impedes corrosion [2].

Water in the atmosphere in its various states (gas, liquid or solid) interacts with material surfaces in several forms: adsorbed moisture (low relative humidity); condensed water (high relative humidity, e.g. dew); and precipitation (e.g. fog, rain, snow). Water plays a major role in the corrosion of metals, the decay of wood and the deterioration of other materials, either as an active agent or as a means for transfer of other agents [2].

Other environmental agents influencing specific deterioration processes are, for example:

- Chloride content in air or sea water (affects corrosion of reinforcing steel in concrete and atmospheric corrosion of steel)
- Concentration of carbon dioxide, CO_2 (affects carbonation and, consequently, possible corrosion of reinforcing steel in concrete)
- Concentration of sulphur dioxide, SO_2 (affects atmospheric corrosion of steel)
- Presence of termites (may cause deterioration of timber structures)
- Etc.

Table 1 (adapted from [2]) summarizes the main environmental actions and action effects that may affect reinforced concrete, steel and masonry structures or components, along with a description of the agents and conditions causing degradation.

Depending on the nature of the agents causing environmental actions that deteriorate and deform materials leading to damage and failure of components, deterioration processes may be also classified in physical and chemical.

Physical processes are those induced by the influences of physical actions such as temperature and moisture, the fluctuation of which can result in damage or separation of components that allow transfer of moisture or humid air, increasing the risk of failure due to material deterioration [2].

Chemical processes are those in which chemical actions are continuously active in many component materials. These actions may be beneficial (e.g. increase of concrete strength during and after curing) or detrimental (embrittlement of plastics exposed to UV radiation). Chemical actions depend on material constituents of a component and agents such as moisture, temperature and oxygen. Frequently, the environmental action due to one agent is dependent on the presence of another, for example moisture combined with oxygen, as can be seen in Table 1 [2].

The most relevant physical and chemical processes that affect RC, steel and masonry structures are presented in next sections.

Table 1 Environmental actions and actions effects in structural materials (adapted from [2])

Type of structure	Material	Environmental action	Action effects	Agents conditions
Reinforced concrete	Concrete	Freeze-thaw cycles	Disintegration, appearance	High moisture content during freeze-thaw cycles, aggravated by chlorides and lack of drainage
		Sulphate attack	Expansion followed by disintegration	Sulphates in groundwater, coal stockpiles or seawater
		Alkali-aggregate reaction	Expansion followed by disintegration	Silica or dolomite aggregates, requires moisture
		Shrinkage	Cracking, damage of adjacent components	High w/c ratio, high moisture content during construction (concrete blocks)
	Reinforcing steel	Corrosion in concrete environment	Loss of bond, failure of reinforcement, cracking and delamination of concrete	Sustained moisture, oxygen. chlorides or pH reduced by carbonation
Steel	Steel	Corrosion in atmospheric environment	Connector failures, appearance, damage due to rust expansion	Sustained moisture, oxygen, aggravated by acid and hygroscopic impurities
		Corrosion in marine environment	Corrosion of piles in splash zone	Sustained moisture, oxygen, aggravated by chlorides
		Corrosion in soil environment	Pile failures, pipe failures	Sustained moisture, oxygen or anaerobic bacteria, aggravated by soluble salts, stray electric currents
Masonry	Stone	Acid attack (leaching)	Disintegration, disfigurement	Carbonates in stone, acid rain, lack of drainage, orientation
		Movements due to moisture change	Bowing of panels	Type of stone, thickness
		Freeze-thaw cycles	Spalling, disintegration	Lack of drainage, high moisture content during freeze-thaw cycles, aggravated by non-breathing surface coatings
	Clay	Freeze-thaw cycles	See under "stone masonry"	
		Salt crystallization	Efflorescence, occasionally spalling	High moisture content and presence of salts in brick, mortar or adjacent materials
		Movements due to moisture or temperature variations	Cracking	Restraints
	Concrete blocks and mortar	See table entries for concrete		
	Reinforcing steel	Corrosion in masonry environment	Failure of connectors, cracking of masonry	Sustained moisture, oxygen, aggravated by salts

3 REINFORCED CONCRETE STRUCTURES

3.1 General

Among the degradation processes that affect the reinforced concrete structures, the corrosion of reinforcing steel is generally considered as the main cause of its degradation [6]. For this reason, much of this section will focus on the modelling of this degradation mechanism. Models for the evolution of the corrosion process, for the main effects of reinforcement corrosion on RC structures, as well for the residual load bearing capacity of corrosion damaged members are presented in Section 3.3.1. The estimation of model uncertainties for the resistance of RC deteriorated elements is also presented in Section 3.3.1. Furthermore, the modelling of other chemical deterioration processes such as sulphate attack and the alkali-aggregate reactions will be briefly described in Sections 3.3.2 and 3.3.3 respectively, while the modelling of physical processes such as the thaw freezing and abrasion will be summarized in Sections 3.2.1 and 3.2.2, respectively.

3.2 Physical Processes

3.2.1 Thaw Freezing

The degradation of concrete under the action of freezing and thawing is a very complex phenomenon and many theories have been presented to explain individual aspects of damaging effects. Among the most important actions, which may lead to disintegration of concrete, the following mechanisms must be considered [9]:

- Generation of hydraulic pressure due to freezing in capillaries
- Diffusion of gel water into capillaries followed by freezing
- Differential strains due to localised shrinkage and swelling as well as thermal strains
- Osmotic pressures resulting from partial freezing in capillaries of solutions with a local salt concentration

For the development of frost damage it is essential, however, that a high degree of saturation of the capillaries with water is prevailing. The ease with which water can move into and within concrete depends on the pore structure characteristics of the concrete. The destructive stress is produced by the restricted flow of displaced water away from the region of freezing, the pressure being due to the viscous resistance to such flow through the permeable structure of concrete. Although the pore size distribution in concrete mainly dictates the extent of damage to concrete during freeze thaw cycles, the amount of freezable water present in the capillary pores is also important. This amount depends on the degree of saturation, the minimum temperature reached by the water and the composition of the pore solution [9].

When combined with de-icing salt, freeze-thaw attack is also affected by materials factors such as aggregate type and reactivity. Besides moisture content, factors such as the minimum freezing temperature, the rate of freezing and the cation types in the de-icing agent are important. Freeze-thaw attack combined with de-icing salts result in physical, but also chemical changes in the pore solution, binder paste matrix and aggregates, inducing finally damage at structural level.

A service life model to describe the internal damage caused by freeze-thaw attack was developed by Fagerlund [10]. The model is based on the observation that a critical water

saturation degree S_{CR} exists, above which the material is damaged by frost. Below S_{CR} no severe damage occurs. S_{CR} is a material parameter and may be obtained from a suitable test.

Further information can be found in [11]. However, according to [4], at present, no validated time-dependent models exist for the calculation of the resistance of a given concrete in a structural component to the action of frost or frost combined with de-icing agents.

3.2.2 Abrasion

Abrasion may lead to a progressive loss of material at the concrete surface. The underlying cause is generally dry attrition due to the rubbing or grinding of another solid object against the concrete surface. Mechanical abrasion is usually characterised by long shallow grooves in the concrete surface and spalling along construction joints.

Concrete resistance to abrasion and erosion depends on the quality of the concrete (low porosity, high strength) and in particular on the aggregate particles used in the mix.

3.3 Chemical Processes

3.3.1 Corrosion of Reinforcing Steel

3.3.1.1 Main Effects of Reinforcement Corrosion

Under normal conditions concrete protects embedded reinforcing steel against corrosion. These protective properties of concrete are attributed to a passive oxide film which forms on the surface of steel in highly alkaline environment provided by the concrete pore solution. However, carbonation or penetration of chloride ions negates the protective properties of concrete and may lead, over time, to corrosion of reinforcing steel [1]. Corrosion leads primarily to a reduction in the cross-section and a decline in material ductility. Geometric discontinuities induced by uneven corrosion may generate stress concentrations, particularly as regards aggressive chloride ion-mediated pitting corrosion. Furthermore, cracking in the concrete cover or even spalling due to corrosion product expansion may have an adverse effect on the composite interaction between concrete and steel due to deterioration of the bond. For all these reasons, reinforcement bar corrosion can affect the ultimate as well as the serviceability limit state performance of concrete structures.

3.3.1.2 Evolution

In order to describe a corrosion process, a two-phase model is normally used [12] (Figure 1, [1]), including an initiation phase lasting until the concentration of the aggressive agent exceeds the threshold value for depassivation at the location of the reinforcement (t_i), and a propagation phase from the steel depassivation until a certain deterioration level is developed in the structure. The deterioration of the reinforcing steel is modelled by a time dependent function (Figure 1). The time of the appearance of first corrosion-induced cracking on the concrete surface is denoted as t_{cr1}, while the time of excessive cracking (serviceability failure) is denoted as t_{cr}. The time when the strength reduction is such that the RC structure does not satisfy anymore an ultimate limit state is denoted as t_u. For a given concrete structure, as well the penetration rate of an aggressive agent as the propagation rate of corrosion depend on environmental parameters such as temperature, humidity, wind, among others, which are variable in space and time. Therefore both, initiation and propagation are stochastic processes. Modelling of these two phases is described in the following sections.

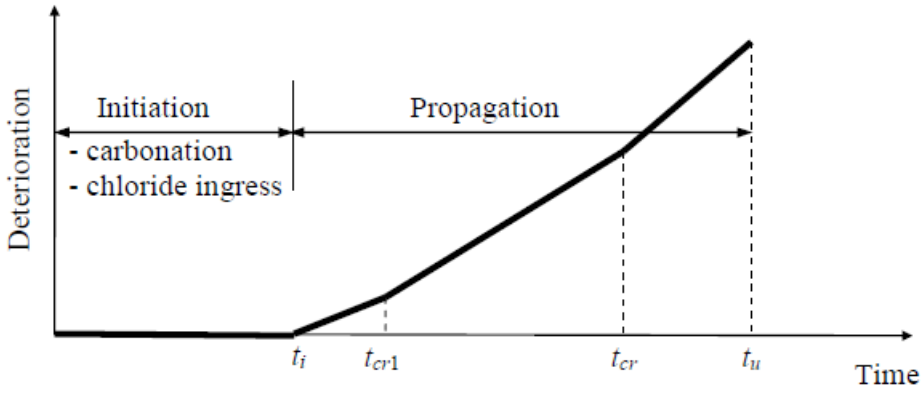


Figure 1 Schematic representation of the development of corrosion in RC structures [1]

3.3.1.3 Models for the Corrosion Process

Initiation Phase

Mathematical models exist for describing the initiation and propagation phases of the corrosion process, which depend on a large number of parameters, mainly related to the material properties and environmental conditions. The initiation phase is the period between the time of exposure of a RC structure to aggressive environment and the time of corrosion initiation. The corrosion starts when [1]:

- The carbonation front reaches the reinforcing steel (carbonation)
- The chloride concentration near reinforcing steel reaches a threshold value (chloride ingress)

The modelling of the initiation phase of corrosion due to carbonation and chloride ingress is considered in the next paragraphs. A particular type of localized attack is the stress corrosion cracking on pre-stressing steel which may develop in pre-stressed tendons. This type of corrosion will not be dealt with in the present handbook.

Carbonation Induced Corrosion

The exposure of concrete structures to atmospheric CO_2 results in the carbonation of the hydration products accompanied by a reduction in pH of the pore solution which can induce corrosion of the steel reinforcement. The penetration of the carbonation front depends on the concentration of CO_2 in the atmosphere and the amount of hydration products able to react with CO_2. Carbon dioxide penetrates into concrete mainly due to gaseous diffusion through air-filled pores.

From 1988 several models have been developed for the calculation of the carbonation front. The models are commonly based on diffusion theory, taking into account the binding of carbon dioxide by the hydration products along with other factors (e. g. relative humidity variations, drying and wetting, lack of homogeneity of concrete, etc.). All models assume that the concrete microstructure does not alter due to interaction with the exposure environment and that temperature and carbon dioxide concentration are constant [11]. The majority of these models have been formulated by applying suitable modifications to the Fick's 1st law of diffusion. The most important of them have been identified and evaluated within the

framework of the European Project DuraCrete [13, 14, 15], finally leading to the proposal summarized below.

If gas diffusion is assumed, the carbonation depth is proportional to the square root of time. According to [4], the propagation of the carbonation front from the concrete surface may be described by:

$$x_c(t) = \sqrt{2 \cdot k_e \cdot k_c \cdot R_{NAC,0}^{-1} \cdot C_S} \cdot \sqrt{t} \cdot W(t) \tag{1}$$

where

$x_c(t)$	carbonation depth at the time t in mm;
t	time in years;
k_e	environmental function which accounts for the moisture conditions;
k_c	execution transfer parameter;
C_S	CO_2-concentration in the air in kg/m³;
$W(t)$	weather function;
$R_{NAC,0}^{-1}$	inverse effective carbonation resistance of concrete in natural conditions in (mm²/years)/(kg/m³).

In order to reduce test time it is recommended to determine $R_{NAC,0}^{-1}$ using an accelerated carbonation test (ACC-test [16]). A relationship between $R_{NAC,0}^{-1}$ and the inverse carbonation resistance of concrete determined in the ACC-test, $R_{ACC,0}^{-1}$, is found by a linear regression analysis:

$$R_{\boldsymbol{NAC},0}^{-1} = k_t R_{\boldsymbol{ACC},0}^{-1} + \varepsilon_t \tag{2}$$

where k_t is the regression parameter for the test effect of the ACC-test and ε_t is the error term for inaccuracies which occur conditionally when using the ACC-test method in (mm²/years)/(kg/m³).

Equation (1) has been developed in the European research project DuraCrete [13] and slightly revised in the research project "Durable and Reliable Tunnel Structures. Deterioration Modelling" DARTS, [17]. Further information on parameters and functions considered in Equations (1) and (2) may be found in [18].

The probability of corrosion initiation, P_{corr}, is obviously time-dependent and can be expressed as

$$P_{corr}(t) = \Pr\left[c - x_c(t) < 0\right] \tag{3}$$

where $x_c(t)$ is the depth of carbonation, calculated according to Equation (1), and c the thickness of the concrete cover [1].

$R_{ACC,0}^{-1}$ can be described by a normal distribution and the thickness of the concrete cover, c, by a lognormal distribution [1]. More detailed description of the probabilistic models for these and the remaining random variables and parameters involved in the calculation of the probability of corrosion initiation, P_{corr}, may be found in [1, 8, 16, 18].

Chloride Induced Corrosion

The chloride ions may be present in the concrete if they are added in the mix (admixtures, water or aggregates). However, this must be avoided and is fortunately not common. The most frequent is that chlorides penetrate from outside, either due to the structure being situated in marine environments or because de-icing salts are used [19]. Penetration of chlorides changes the chemical composition of the pore solution of concrete adjacent to the steel reinforcement causing corrosion. This is a complex process involving two transport mechanisms, ion diffusion and convection, which depend on a large number of factors including the properties of concrete (i.e., its composition, porosity and microstructure), the degree of concrete pore saturation, and the exposure conditions at the concrete surface sometimes referred to as "microclimate" (i.e., the surface chloride content, temperature, and humidity). A number of these factors are inter-, time-, spatial-, and temperature-dependent [1].

In the last decades, a number of models have been proposed to describe the process of chloride ingress into the concrete. However, most of these models involve a large number of parameters which depend on numerous factors and usually can only be determined experimentally with a large degree of uncertainty. For these reasons, semi-empirical models, which assume that the chloride ingress can be treated as a pure diffusion process, are mostly employed [1]. Assuming that chloride ingress is a pure diffusion process, it can be described by the 2nd Ficks´s Law of diffusion:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} \tag{4}$$

where C is the total concentration of chloride ions at distance x from the surface after the time t of exposure to chlorides and D the chloride diffusion coefficient.

In the framework of the abovementioned European joint research projects DuraCrete and DARTS [13, 17] a model for the prediction of time- and depth-dependent chloride content has been also developed and validated (Equation 5). Assuming that the surface chloride concentration remains constant with time, the solution of Equation (4) is expressed by [4]:

$$C(x,t) = \left(C_0 + \left(C_{s,\Delta x} - C_0 \right) \cdot \left[1 - \mathrm{erf}\left(\frac{x - \Delta x}{2 \cdot \sqrt{D_{app,C} \cdot t}} \right) \right] \right) \tag{5}$$

where

$C(x,t)$	chloride content of concrete in % by mass of cement;
x	depth in m;
t	concrete age in s;
C_0	initial chloride content of concrete in % by mass of cement;
$C_{s,\Delta x}$	chloride content at a depth of Δx in % by mass of cement;
erf	error function;
Δx	depth of the convection zone in m;
$D_{app,C}$	apparent chloride diffusion coefficient in concrete in m²/s

with

$$D_{app,C}(t) = k_e \cdot D_{RCM,0} \cdot k_t \cdot A(t) \tag{6}$$

where

$D_{RCM,0}$ is the chloride migration coefficient in m²/s;
k_e is the environmental variable;
k_t is the test method variable;
$A(t)$ is the ageing function, calculated by

$$A(t) = \left(\frac{t_0}{t}\right)^a \tag{7}$$

being

t_0 reference concrete age in s;
a age exponent.

As indicated in [16], $D_{app,C}$ is usually determined by use of the so-called "Chloride profiling method". The determined $D_{app,C}$ is a constant average value representing the period from start of exposure to the moment of inspection when the profile is taken (time of interest). While assessing existing structures, $D_{app,C}$ might be derived directly from chloride profiles taken from the chloride exposed structure at different points in time, although such a procedure is highly impracticable in most real cases. Another way for determining $D_{app,C}$ is from test samples stored under conditions which are expected in practise. However, the determination of $D_{app,C}$ on test samples (for the design of new structures) is very time consuming method. Consequently a second, empirically derived approach is offered in [16]. The chloride migration coefficient, $D_{RCM,0,}$ is one of the governing parameters for the description of the material properties in the chloride induced corrosion model. $D_{RCM,0}$ varies significantly depending on the water/cement ratio and type of binder. The exponent a also varies considerably with the cement type and exposure.

The chloride content at a depth of Δx, $C_{s,\Delta x}$, depends on material properties and on geometrical and environmental conditions. When assessing existing structures exposed to a chloride rich environment $C_{s,\Delta x}$ might be derived directly from chloride profiles from the structure. Further information on parameters and functions involved in Equations (5), (6) and (7) may be found in [16].

The probability of initiation of chloride induced corrosion, P_{corr}, is time-dependent and can be expressed as

$$P_{corr}(t) = \Pr\left[C_{crit} - C(x=c,t) < 0\right] \tag{8}$$

where C_{crit} is the threshold chloride concentration and c the thickness of the concrete cover [1].

In this context, the critical chloride content C_{crit} is defined as the total chloride content which leads to the depassivation of the reinforcement surface and initiation of iron dissolution, irrespective of whether it leads to visible corrosion damage on the concrete surface. Detailed information about the quantification of the parameters and probabilistic models for the random variables involved in the calculation of the probability of corrosion initiation, P_{corr}, may be found in [1, 8, 16, 18].

Propagation phase

The propagation phase is defined as the time period that extends from the moment at which steel depassivation is produced until the structure reaches an unacceptable deterioration. This period can be determined from the so-called attack penetration function, $P_x(t)$, representing the loss of rebar diameter at time t. This function can also be implemented in damage functions for structural performance.

In a similar way as for the initiation phase, the propagation of corrosion is time-dependent and the following limit state can be considered:

$$P_f = \Pr\left[P_x(t) > P_{crit}\right] \tag{9}$$

where P_{crit} is the critical loss of rebar, which mainly depends on structural geometry [18], as well as internal forces and moments.

In a rigorous form, the attack penetration function, $P_x(t)$, is expressed by:

$$P_x(t) = \int_{t_i}^{t} V_{corr}(\tau) \cdot d\tau \tag{10}$$

where $V_{corr}(\tau)$ is the corrosion rate at the instant τ, and t_i is the initiation period. Expression (10) gives the evolution of corrosion with time as a function of the corrosion rate V_{corr} in mm/year.

The corrosion rate determines the rate at which the bar is losing its section, generating oxides and causing other effects. Alternatively the corrosion rate can be defined by the corrosion current density I_{corr}, with units of $\mu A/cm^2$. The conversion between the two parameters is achieved through the density of the steel and the Faraday's Law.

The corrosion rate, V_{corr}, is usually not constant along the propagation phase and may vary due to different events (e. g. continuous arrival of chlorides, environmental changes) suffered by the structure [18]. There are numerous parameters related with the concrete properties and environment which may influence the value of the corrosion rate. They are inter-related in such a way that, for instance, the temperature influences the humidity within the concrete, the resistivity, and the electro-chemical reactions. Consequently, the prediction of its influence on V_{corr} is a difficult task. Four main parameters, which include the effect of others, have been selected in [13] to evaluate the corrosion rate behaviour, namely the resistivity, galvanic effects, chloride content and humidity/temperature.

The evolution of corrosion rate with the variation of these parameters has been detailed in [13, 19] and a brief description may be found in Chapter 2: Non-destructive Testing, Inspection Techniques and Monitoring of [20].

A representative value of the corrosion rate, V_{corr}^{rep}, has to be initially determined. The main difficulty regarding the modelling of the damage function $P_x(t)$ lies in the establishment of the value of V_{corr}^{rep}, being able to represent the particular conditions of the corrosion process. According to [18] three practical approaches may be adopted for modelling the propagation phase:

- Assuming values of the corrosion rate as a function of the exposure class and of the type of attack
- Estimate V_{corr}^{rep} from direct measuring of I_{corr} in specimens (design phase) or on-site (existing structures)
- Using an empirical relationship based on the variables governing the process (e. g. resistivity, which also governs the initiation period)

In existing structures, the I_{corr} value can be measured on-site by means of a corrosion portable meter. The procedure for measuring the corrosion rate has been described in Chapter 2 of [20].

As it has been mentioned earlier, propagation of corrosion is a stochastic process. For more accurate measurements the influence of climatic variations in that particular structure and environment has to be taken into account by conducting measurements in several locations of the structure for at least a period of one year. However, this is not always possible due to the need to complete the assessment, the type and location of the structure or the accessibility, among other factors. Therefore, a clear distinction must be drawn between propagation rate at a given point of the structure and at a given time, and the mean propagation rate. The inference is that corrosion rate measurements must be interpreted carefully.

The values obtained for the corrosion rate are used for the modelling of the following consequences of reinforcement corrosion in concrete elements:

- Reduction in bar cross section
- Change of reinforcing bar material properties
- Concrete cover deterioration
- Bond deterioration

Reduction in Bar Cross-Section

Two types of corrosion are possible: homogenous and pitting. Homogenous corrosion affects a substantial area of reinforcement with more or less uniform metal loss over the perimeter of reinforcing bars and occurs in carbonated concrete. Pitting (or localized) corrosion, in contrast to homogenous corrosion, concentrates over small areas of reinforcement and is usually caused by chloride penetration.

The attack penetration (loss of bar radius), in mm, after a certain period, t_p, in years, following the initiation of the corrosion process, can be obtained by:

$$P_x = 0.0116 \cdot I_{corr}^{rep} \cdot t_p \tag{11}$$

where

P_x average value of the attack penetration (loss of bar radius), in mm;
I_{corr}^{rep} representative value of the corrosion rate during the time t, in $\mu A/cm^2$;
t_p time elapsed since the aggressive reached the bar (propagation period), in years.

Once the depth of the attack penetration has been obtained, the residual bar diameter can be estimated by means of the following Expression [13, 19]:

$$\phi_t = \phi_0 - \alpha P_x \tag{12}$$

where

ϕ_t residual diameter at time t_p, in mm;
ϕ_0 initial (nominal) diameter, in mm;
α coefficient that depends on the type of attack; when homogeneous corrosion occurs, α is equal to 2; however, when localized corrosion occurs, α may reach values of up to 10 (Figure 2);
P_x average value of the attack penetration (loss of bar radius), in mm.

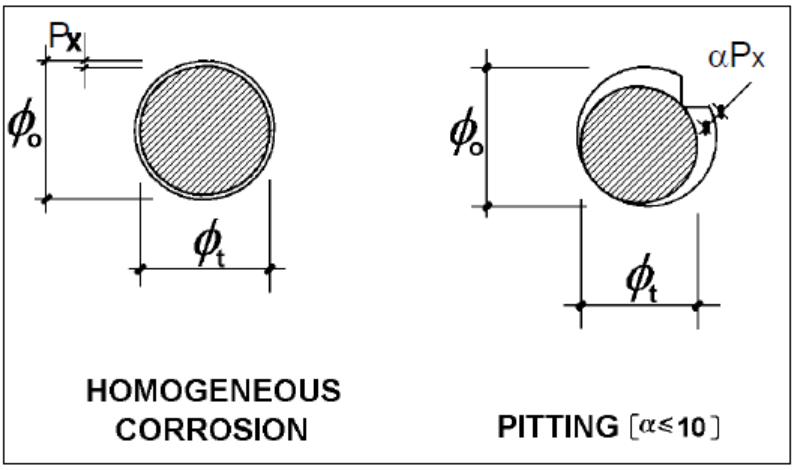


Figure 2 Residual reinforcing bar section for homogenous and pitting corrosion [19]

Change on Reinforcing Bar Material Properties

There exists experimental evidence that the mechanical properties of reinforcing steel may be affected by corrosion [13]. A significant reduction of the elongation at maximum load was found in some research studies [21, 22]. Reductions in the elongation of 30 and 50% for cross section losses of 15 and 28 % were reached in these studies. This reduction may affect the capacity of the bending moment redistribution in corroded structures.

As noted in [13] experimental results should be treated with caution in cases of pitting, because the reinforcing bars normally fail at pit location due to reduction of cross sections when, perhaps, the strain of the bar in zones outside the pit is still low.

According to [1] available experimental data shows that homogenous corrosion does not affect such properties of reinforcing steel as strength and ductility, whereas there is strong evidence that pitting corrosion may affect both properties. It has been suggested that yield strength, f_y, of a reinforcing bar reduces linearly with steel section loss caused by pitting corrosion and that the mechanical behaviour of reinforcing bars changes from ductile to non-ductile (brittle) as pitting corrosion loss increases [23]. While there is a gradual transition from ductile to brittle behaviour with an increase in corrosion loss. For simplicity it can be assumed that the complete loss of ductility in corroding reinforcing bars occurs after the percentage corrosion loss (measured in terms of reduced cross sectional area) exceeds a threshold value. The literature shows that it is reasonable to quantify this threshold value of corrosion loss at 20%, although more research is needed to more accurately quantify this important variable [1].

As can be noted, in spite of the experimental evidences, the results furnished are, uncertain and in some case contradictory. Therefore, the modelling of the deterioration in reinforcing steel properties should be treated with caution.

3.3.1.4 Load Bearing Capacity

A number of simplified models for estimating the performance of concrete beams and columns with corroded reinforcement bars are described in the literature [19]. The load bearing capacity of such members can be determined by applying the basic principles set out in structural design codes such as Eurocode 2 [3] and strength models can be established based on the lower bound theorem of the theory of plasticity [24]. For such models, the effect of deterioration on the load bearing and deformation capacities must be quantified and dealt

with as updated geometrical or construction material properties. Since in existing structures many characteristics can be measured, updating geometric and material parameters is relatively straightforward even in deteriorated members.

Bond between possibly corroded reinforcing bars and concrete and hence the mechanism involved in longitudinal stress transfer are quite another matter, however. Despite its complexity, an understanding of bond performance in corrosion-damaged structures is essential for the assessment of structural reliability, since load transfer failure between the reinforcement and the surrounding concrete is conducive to potentially brittle structural behaviour. An experimental and numerical study was conducted to address the existing lack of knowledge about bond performance in corrosion-damaged structures. The experimental procedure and most prominent findings are discussed and compared to the results of numerical simulation in a previous study [25].

Although further research is needed to fully predict bond performance in corroded steel bars, further to the above all the information relevant to establishing the load bearing capacity of corrosion-damaged reinforced concrete structures can be assumed to be available or obtainable: material as well as geometrical properties, including the remaining bar and concrete cross-sections, from site surveys, and bond strength from studies such as cited above [25]. The residual bending capacity of corrosion-damaged beams, for example, can be quantified by using the aforementioned updated information in combination with suitably adapted and modified strength models described in structural codes, such as Eurocode 2 [3] or the Model Code [4]. The upper limit for residual strength is estimated by entering the reduced area of the reinforcement bar cross-sections in such bending strength models. Where relevant, stress concentrations must also be taken into consideration, particularly in the presence of pitting corrosion, whereas steel embrittlement may be disregarded in most reinforced concrete members. In addition to these alterations, where spalling has occurred, the remaining concrete cross-section must be addressed in the model. Even where no spalling is observed, the possible cracking of the concrete cover in the compression zone as a result of corrosion product expansion may not be ruled out. The effects of such cracks on concrete strength are assumed to be equivalent to the effects of cracks due to laterally imposed strain on the concrete in the stress field of a sound structure. In such circumstances, the effective concrete strength, f_{ce}, may be determined with sufficient accuracy as a function of the lateral strain, ε_1, and the characteristic compressive strength of test cylinders, f_{ck} [26]:

$$f_{ce} = \frac{1}{0.8 + 170 \cdot \varepsilon_1} \cdot f_{ck} \leq f_{ck} \tag{13}$$

Since corrosion-induced strain is hard to quantify, Equation (13) cannot be used to establish the effective concrete strength directly. Crack-weakened concrete must be handled indirectly. Here also, an analogy is drawn to sound structures, where lateral strain in rigid-plastic stress fields is not calculated directly. Rather, approximate values are used for the effective concrete strength depending on the internal forces or moments considered and the angle between the imposed cracks and the orientation of the compression field [26]. Since corrosion product expansion usually induces cracks parallel to the bars and hence to the compression field (where bending is concerned), the effective strength of the concrete cover can be defined as [4]:

$$f_{ce} = k_{cp} \cdot \eta_{fc} \cdot f_{ck} \tag{14}$$

where a value of 0.75 is adopted for reduction coefficient k_{cp} [4] and η_{fc} is a factor that introduces the rise in concrete brittleness with rising cylinder strength [4, 24, 26]:

$$\eta_{fc} = \left(\frac{30}{f_{ck}} \right)^{\frac{1}{3}} \leq 1.0 \tag{15}$$

When estimating the lower limit for residual strength, in turn, the concrete cover in the compressive zone of the cross-section analysed should be disregarded, under the assumption that corrosion product expansion induces concrete spalling. In this case also, the reduced area of the reinforcement bar cross-sections must be entered in the strength models.

Similarly, the residual shear capacity is estimated from the (suitably adapted) strength models described in the Model Code [4]. The assumptions adopted are analogous to those defined for estimating the residual bending capacity and used in combination with the updated information on material, geometrical and bond properties. In beam webs, the reinforcement is positioned obliquely to the flow of the compression force and hence to the stress field, and a reduction factor of k_{cp} = 0.55 is used to determine the effective concrete strength as per Equation (14). In columns, in contrast, corrosion-induced cracks are parallel to the flow of compression forces and a k_{cp} reduction factor of 0.75 is assumed to establish the effective concrete cover strength.

The deviation in the results obtained with the aforementioned models from the actual load bearing capacity of corrosion-damaged structural members has not been clearly determined. The impact of strength model uncertainties on reliability is expected to be higher in corrosion-damaged than in sound structures [27]. The quantification of these uncertainties and the estimation of their effect on structural reliability are therefore of cardinal importance when assessing corrosion-damaged RC structures.

3.3.1.5 Resistance Model Uncertainty Variable

General

Defining the model uncertainty, ξ, as a factor to be multiplied by the value obtained with the applied model, R_{mod}, in order to predict the uncertain load carrying capacity, R

$$R = \xi \cdot R_{\text{mod}} \tag{16}$$

the uncertainty associated with a particular resistance model may be obtained by comparing experimental results, r_{exp}, with the values predicted by the applied model, r_{mod}, given the experimental conditions. The random variable representing the model uncertainty, ξ, may be assessed through observations of ξ_i, where

$$\xi_i = \frac{r_{\text{exp},i}}{r_{\text{mod},i}} \tag{17}$$

Model uncertainties defined in this way have mean values close to 1.0 if they are unbiased. Typical coefficients of variation for good models, for example the models for bending resistance of sound reinforced concrete cross-sections, may be in the order of magnitude of 2% to 5%, whereas for models such as for the shear capacity of reinforced concrete members coefficients of variation in the range of 10% to 20% must be expected [28].

For the model uncertainty variable as defined in Relation (16), it is convenient to model the probability distribution function by a log-normal distribution.

Estimation of Model Uncertainties

Based on the experimental results from two studies [29, 30] the parameters for model uncertainty variables for bending and shear resistance models of corrosion-damaged members have been deduced and the obtained results are summarized in this section. The empirical values (where available and nominal values otherwise) for constituent material dimensions and properties were entered in the models from Section 3.3.1.4 to calculate the theoretical load bearing capacities for each beam studied. These theoretical results, together with the load bearing capacity obtained in the respective tests, were then used to determine the probabilistic parameters for model uncertainty variables $\xi_{R,M,corr,sup}$ and $\xi_{R,M,corr,inf}$ for, respectively, the upper and lower bending capacity in RC beams with corroded plain and ribbed reinforcement bars. Only the results in which the theoretical failure mode matched the experimental failure mode were used to derive these probabilistic models. The findings are summarised in Table 1.

Table 1 Probabilistic models for the uncertainties associated with bending and shear strength models adapted from codes [3] and [4] for corrosion-damaged RC beams

Failure mode	Rebar	Limit	Notation	Number of tests	Type	*Mean*	*CoV*
Bending	Plain	Upper	$\square_{R,M,corr,sup}$	6	LN	1.06	0.20
		Lower	$\square_{R,M,corr,inf}$	6		1.18	0.21
	Ribbed	Upper	$\square_{R,M,corr,sup}$	15	LN	0.99	0.10
		Lower	$\square_{R,M,corr,inf}$	12		1.29	0.14
Shear	Ribbed	Upper	$\square_{R,Vs,corr,sup}$	12	LN	0.88	0.18
		Lower	$\square_{R,Vs,corr,inf}$	12		1.00	0.18

The mean value found for model uncertainty variable $\xi_{R,M,corr,inf}$ = 1.29 for RC beams with ribbed steel bars indicated that the model for the lower limit was far too conservative. The model for the upper limit, in turn, with a mean value of 0.99 for variable $\xi_{R,M,corr,sup}$, overestimated the load bearing capacity slightly. The coefficient of variation found for the uncertainty variable in both models, on the order of over 10 %, denoted a lower level of precision than obtained with bending strength models for sound RC beams, whose uncertainties are reported in the literature [28, 31].

The mean value calculated for $\xi_{R,M,corr,inf}$ for RC beams with plain steel bars (1.18) likewise denoted a conservative lower limit, whereas the model for the upper limit, with a mean value of 1.06 for model uncertainty variable $\xi_{R,M,corr,sup}$, underestimated load bearing capacity slightly. The coefficient of variation found for the uncertainty variable in both models for plain steel bars, on the order of 0.2, was somewhat higher than the coefficient found for ribbed bars.

The upper and lower limit values for the theoretical capacity were calculated for each of the tested beams that failed due to shear induced by tensile forces in the web. In these cases also, empirical values for dimensions and material properties were used together with the shear strength models. The theoretical results obtained, together with the load bearing capacity from the respective tests, were then used to deduce the probabilistic parameters for model uncertainty variables $\xi_{R,Vs,corr,sup}$ and $\xi_{R,Vs,corr,inf}$. These variables respectively

represented the upper and lower shear capacity limits in RC beams with corroded reinforcement bars, in keeping with Model Code level of approximation II [4] (Table 1).

The mean value of 0.88 observed for the uncertainty variable for the upper limit, $\xi_{R,Vs,corr,sup}$, denoted non-conservative bias in which the load bearing capacity was overestimated. The model for the lower limit, in turn, with a mean value of 1.00 for model uncertainty variable $\xi_{R,Vs,corr,inf}$, was indicative of a good estimate of shear capacity. The coefficient of variation found for the uncertainty variable, around 0.18, indicated a fairly low level of precision compared to strength models for sound RC beams [31].

Moreover, some of the uncertainties associated with geometric and material variables are implicit in the model uncertainties in Table 1, since not all the information required on the tested beams was available in the respective studies [29, 30]. The uncertainties associated with the models used to determine bending or shear strength in corrosion-damaged beams were therefore presumably overestimated. Hence, further experimental data are needed. Information should be gathered, for example, on deteriorated RC members with larger dimensions than those from the studies cited here [29, 30], extracted from existing structures after exposure to real service conditions and natural corrosion, to improve existing models for calculating the corrosion-damaged structure strength.

3.3.2 Sulphate Attack

All sulphates are potentially harmful to concrete. Sulphate attack is the result of the exposure of a concrete product or structure to an excessive amount of sulphate from internal or external sources. External sulphate attack is most common and typically occurs when water containing dissolved sulphates penetrates the concrete.

Sulphate attack is one of the phenomena that may cause gradual but severe damage to concrete structures. Sulphate attack on concrete is primarily attributed to sodium, magnesium and calcium sulphate salts present in the soil, ground water or sea water which react with calcium hydroxide and, if enough water is present, result in expansion and irregular concrete cracking that can lead to a progressive loss of strength and mass. The severity of sulphate attack depends on water penetration, the type and concentration of the sulphate salt, the salt development mechanism (i.e., crystallisation) and the chemistry of the binder present in the concrete. Sulphate attack on concrete leads to the conversion of cement hydration products to ettringite, gypsum, and other possible phase changes. Accompanying the formation of gypsum and ettringite is a volume increase. The volume expansion within concrete cannot only reduce the porosity of concrete but may also cause damage and cracking of the concrete leading eventually to a strength loss [32].

Concrete has been reported [33] to undergo sulphate attack at sulphate concentrations of about 0.2 % in ground water. Magnesium sulphate may be more aggressive than sodium sulphate and the three key chemical reactions between sulphate ions and hardened cement pastes are: ettringite recrystallisation, calcium sulfoaluminate (ettringite) formation, and decalcification of the main cementitious phase (calcium silicate hydrate).

The concrete structures most exposed to attack by sulphates present in the soil and ground water include footings, foundation walls, retaining walls, piers, culverts, piles, pipes, and surface slabs. Structures in environments where sea spray is common are more resistant to sulphate attack because of the presence of chlorides. In the presence of calcium hydroxide, chloride ions will react with hydrated calcium aluminate to form single- and tricalcium chloroaluminates which will block the formation of ettringite [32]. Dense, low permeability concretes containing low tricalcium aluminate content cement are most resistant to sulphate attack.

Thaumasite formation constitutes a rare form of sulphate attack. Thaumasite is formed during sulphate attack at low temperatures (0º C to 5º C). Formation of thaumasite occurs as a result of the reaction between the calcium silicates in the cement, calcium carbonate from limestone aggregates or fillers, and sulphates, usually from external sources. Thaumasite could also form from ettringite with which it could enter into a solid solution. Thaumasite related deterioration has been identified in historic buildings in cold climates, where hydraulic cement-based mixtures with high w/c ratio were used for restoration purposes [34].

The modelling of degradation of concrete elements due to external sulphate attack involves, on the one hand, the transport of sulphate ions and the chemical reaction between penetrating sulphates and hardened cement pastes and, on the other hand, the loss of strength of the binder matrix and the volumetric expansion leading to cracking. The chemo-transport mechanism is usually modelled using a diffusion-reaction approach. Fick's second law is assumed e. g. in [35] for modelling the diffusion of the sulphate ions. The mechanical consequences of the degradation processes are, in this case, calculated on the basis of the relevant reaction products formed (i.e. ettringite and/or gypsum).

Models for the mechanical behaviour of concrete exposed to external sulphate attack have been developed on the basis of different approaches. A critical review of the most pertinent models proposed in the literature is presented in [36]. According to [36] authors have followed various paths to develop microstructure-based models that can reliably predict the behaviour of hydrated cement systems subjected to sulphate attack. Models derived from these various approaches may be divided into three categories: empirical models, mechanistic (or phenomenological) models, and numerical models. The study concludes that although many empirical and mechanistic models tend to yield fairly reliable results, only numerical models were found to have to ability to capture the complex nature of the degradation mechanisms.

More complex numerical models for the mechanical degradation of concrete due to external sulphate attack have been proposed in recent publications such as [37]. The coupled chemo-mechanical model proposed in [37] is based in the original proposal presented in [35] and explicitly accounts for the mesostructure of concrete. The model, previously developed for purely mechanical analyses, is extended to the domain of chemo-mechanical coupled problems and is based on a nonlinear fracture mechanics crack propagation procedure.

3.3.3 Alkali-Aggregate Reaction

Concrete expansion and cracking with a concomitant loss of strength, stiffness, and durability, may be induced by the calcium alkali-silicate gel forming when alkaline ions present in portland cement react with calcium and hydroxyl ions and certain siliceous constituents in aggregates [38]. Due to the attraction between the polar water molecules and the alkali-silicate ions, this gel takes up pore solution water and expands, altering the concrete. Alkali-silica reactions (ASR) are governed by aggregate reactivity (i.e., amount and grain size of reactive aggregate), the alkali and calcium concentrations in the concrete pore water, cement content (i.e., alkali content), and the presence of water. The most reactive forms of aggregate are strained quartz, amorphous silica, cryptocrystalline quartz, chalcedony, and chert.

The effects of ASR on the engineering properties of concrete can seldom be generalized, since both the expansion rate and total expansion depend on many variables: the reactive aggregate, cement type, cement content, constraint, and environment. At expansive strain values of 0.5 to 1.5 %, compressive strength may decline by 40 to 60 %, tensile strength by as much as 65 to 80 %, and elastic modulus by 60 to 80 % [38].

The consequences of the phenomenon of alkali-aggregate reaction (AAR) on the serviceability and load bearing capacity of RC structures has been subject of a large number of studies world-wide [13]. The evaluation of the current and future condition of existing structures damaged by AAR requires that a wide range of effects be considered. These may include combinations of the following:

- Overall expansions of structural members
- Differential movements between structural members
- Global and local differential movements within structural members
- Effect of stress conditions on expansion
- Effect of the restraint by reinforcement in reducing expansion
- Increased tensile strains in reinforcement
- Pre-stress induced within concrete by the restraining effect of the reinforcement
- Reduced compressive and tensile strength of the concrete
- Reduced elastic modulus of the concrete

The assessment of structures affected by AAR is a complex task which includes calculations of the structural effects, a thorough investigation of the steel detailing and the prediction and evaluation of the effects of future expansions. The determination of the current level of expansion of an affected element is difficult and normally can only be done indirectly. The mechanical properties of concrete in RC elements affected by AAR should be obtained from extracted cores.

The prediction of the future expansion behaviour is another important aspect in the assessment of the structures affected by AAR. This includes an assessment of the likelihood, rate and duration of any future expansion. Further information about the evaluation of the present condition and the prediction of the effects of future expansion can be found in [13]. However, according to [4], at present no suitable predictive analytical or numerical method exists for the durability modelling of the concrete behaviour affected by AAR.

4 STEEL STRUCTURES

4.1 Overview

Corrosion and fatigue crack growth are the two major deterioration mechanisms of steel structures. Corrosion causes loss of material leading to smaller net sections and, consequently, increases the stress level for a given load. The geometric parameters, such as moment of inertia or radius of gyration, decrease due to the reduction in the section area. Buckling capacity of structural members can be critically affected by the reduction in the metal thickness [39].

A steel component under the effect of a cyclic action, such as wind, waves, traffic or mechanical vibrations, would deteriorate due to the propagation of fatigue cracks. The fatigue loading is characterised by the number of stress cycles and the magnitude of stress range for each cycle [8]. Since the fatigue crack propagation modelling involves a number of parameters, which are subject to uncertainties, it should be treated in a probabilistic sense.

In this chapter, only the modelling of degradation of steel structures due to the corrosion is treated, while the modelling of the fatigue crack propagation is treated in Chapter 4.

4.2 Corrosion

4.2.1 General

The types of corrosion that may occur in practice in steel are: general or uniform, pitting, crevice and galvanic. A number of possible different factors and environmental conditions may influence corrosion. Therefore, it is clear that there is a very wide range of corrosion possibilities, particularly when intertwined with the influences of protective paints and cathodic protection.

The so-called ‘‘uniform’’ or ‘‘general’’ corrosion is the nominal loss of material from a plate or a piece of metal. It is commonly defined as the average depth of penetration of corrosion on one side, measured in mm. Usually it is calculated from weight loss measurements obtained from exposing flat plate coupons to an aggressive environment. ‘‘Uniform’’ corrosion loss of material is of the most interest for overall degradation of structural strength, as in plates and structural members. Pitting corrosion is important for containment, such as for pressure vessels and to a lesser extent for offshore platforms. Crevice corrosion is relevant mainly to stainless steels at fittings, bolts, etc. Galvanic corrosion caused by differences in material properties around the heat affected zone of welds may be important for strength considerations of local details and for overall strength of stiffened plates. It has been shown that the material properties themselves, such as elastic modulus and yield strength are not influenced by the corrosion of adjacent material [40].

For steel structures corrosion can be idealized as either ‘uniform’ (or general) corrosion in which the corrosion loss over a surface is sufficiently similar with location to allow this assumption to be made, and ‘pitting’ corrosion. Modelling of the evolution of the corrosion of steel structures in atmospheric and marine environment is presented in the following sections.

4.2.2 Corrosion of Steel in Atmospheric Environment

The characteristics of atmospheric corrosion of steel have been described in several references, e. g. [41]. According to [41], atmospheric corrosion in general is the result of the conjoint action of two factors: oxygen and moisture (water in liquid form). If one of these factors is missing, corrosion does not occur. The atmospheric corrosion increases strongly if the air is polluted by smoke gases, or aggressive salts, as in the vicinity of chimneys and marine environments. The atmospheric corrosion is therefore particularly strong in industrial and coastal areas. The corrosion is, furthermore, much higher if the metal surface is covered by solid particles, such as dust, dirt, and soot, because moisture and salts are then retained for longer time.

The degree to which steel will rust in the atmosphere is directly related to the so-called time of wetness, defined as the average number of hours per year during which the steel is wet. The longer a steel surface remains wet, the more it will corrode. The time of wetness varies with the climatic conditions at the site. It depends on the relative humidity of the atmosphere, as well as the duration and frequency of fog, dew, rain, and snowfall. Direct precipitation is not needed to wet a steel surface. Moisture can be deposited also by capillary action of the porous rust coating, and adsorption by corrosion products or salt deposits on the steel surface [42].

Because of its implication for structural strength, only the modelling of the loss of material due to ‘uniform’ corrosion is considered in this chapter. Although the referenced publications include some theoretical aspects and experimental data which are valid for corrosion in marine atmospheres, only the long-term corrosion of structural steel in rural,

urban and industrial atmospheres, without any marine component, is considered in this section, while marine atmospheres are treated in the following one.

For structural reliability assessments using probability theory [43], the structural strength at any point in time t can be represented by $R(t)$ with probability density function $f_R(r,t)$ where $R(t)$ is a random variable and r a discrete, deterministic value as in the usual notation. Corrosion will affect $R(t)$. A model for corrosion loss of material will represent the amount of corrosion to be expected after a given period of exposure time together with the associated uncertainty [40].

Corrosion loss models may be categorized as those based on some level of theoretical input and those that are essentially empirical [44]. To the first category belong e. g. those proposed in [45, 46]. They both consider metal oxidation to be the controlling mechanism and that this reduces with time as the layers of corrosion products build up.

Is widely accepted that the long-term atmospheric corrosion of steel may be described by an equation of the type [47],

$$c(t) = A \cdot t^b \tag{19}$$

where $c(t)$ is the metal loss after t years, and A and b are empirical constants. The exponent b is usually less than unity and depends on the local atmosphere and exposure conditions of the metal. In the special case when $b = 1$, A represents the mean corrosion rate for 1-year exposure. By plotting the atmospheric corrosion of steel against the exposure time in log–log coordinates, straight lines of variable slopes are obviously obtained. The accuracy of the Equation (19) and its reliability to predict long-term corrosion have been demonstrated by [45, 46], among others authors.

The exponential law, expressed by Equation (19), with b close to 0.5, can result from an ideal diffusion-controlled mechanism when all the corrosion products remain on the metal surface. This situation seems to occur in slightly polluted inland atmospheres. On the other hand, b values of more than 0.5 arise due to acceleration of the diffusion process (e.g. as a result of rust detachment by erosion, dissolution, flaking, cracking, etc.) [45]. This situation is typical of marine atmospheres, even those with low chloride contents. Conversely, b values of less than 0.5 result from a decrease in the diffusion coefficient with time through recrystallisation, agglomeration, compaction, etc. of the rust layer [47]. According to [44] calibration to field data invariably shows that b varies between about 0.3 and 0.8 and that both A and b are highly sensitive to small changes in data or to additional data.

A mixed linear–exponential model, according to which a plot of corrosion against time would consist of an initial parabolic portion followed by a straight line, has been proposed by authors such as [48]. The authors compared both models on the basis of atmospheric corrosion data for weathering steels reported in the United States, concluding that the experimental data fitted the power-linear model better than the power model.

Several experimental studies have been developed in different countries to obtain the variation in the steel corrosion rate with the exposure time in different atmospheres, and to determine how the corrosion data obtained at tests fitted the Equation (19). By the way of example, the results obtained for the long-term atmospheric corrosion in Spain [47], showed that for rural and urban atmospheres, in which the corrosion is low, the corrosion data deviate from the power function expressed by Equation (19), while for the others atmospheres considered (industrial, mild marine and severe marine) follow the power law.

4.2.3 Corrosion of Steel in Marine Environment

As is described in [42], marine environments are generally the most aggressive towards steel. Among the reasons for this are the salt spray produced by nearby bodies of water and the nightly temperature drops near the shore. Moisture retained by salt on steel surfaces, coupled with the presence of chloride ions, causes all structural steels to corrode severely. In the absence of direct sunlight, a salt-covered surface can stay moist most of the day. Other factors affecting corrosion in marine environments are wave action at the surf line, prevailing wind direction, shoreline topography, and relative humidity. Depending on the exposure environment, marine corrosion may be divided into the following categories:

- Immersion
- Splash/tidal zone
- Atmospheric;
- (Semi-)enclosed space

A metallic structure erected in water is subject to considerable corrosion especially in those environments where the water has any degree of salinity as in marine environments. The part of structure subjected to the intermittent or the frequent contact of water, either due to splashing or by tidal changes, experiences the most severe corrosion. This zone which is normally referred to as the splash zone, is defined as that existing between the highest point reached by the water due to either wave action or splashing, to the lowest point reached by air. Corrosion in this zone is usually the most severe and difficult to protect [49]. Because of its importance for the assessment of aging infrastructures, this section deals only with the corrosion structural steels in splash and tidal zones.

Because corrosion is a function of many variables, many of an uncertain nature, a probabilistic model to describe expected corrosion is appropriate. Structural reliability estimation for steel structures exposed to marine environments is of interest for offshores, harbour-side and coastal facilities. In such estimation, the rate, extent, localization and variability of expected corrosion is important, particularly for the splash and tidal zones, where cathodic or coating protection systems are often employed, although this not always the case and, for a variety of reasons, they may not be wholly effective. Therefore, corrosion may threaten safety and performance through loss of material [40]. The reliability assessment should consider the time deterioration of the strength of the structure as well as the probabilistic nature of marine corrosion.

Despite some quite extensive, long-term experimental test programs, the prediction of the likely corrosion loss of material is still rather simplistic, even though the various factors that can influence marine corrosion are known. However, precise understanding of the mechanics of marine corrosion has been slow to develop and mathematical formulation has been largely neglected [50].

In [51] are described the results from efforts made to develop non-linear and piecewise linear phenomenological models for the probabilistic description of the corrosion of steel under tidal conditions. Further, using the available data, the authors have shown that the uncertainty in the expected deterioration effect for tidal zone corrosion in marine environments can be described by a time dependent log normal distribution, as well as for immersion corrosion.

In [52] field data for the corrosion of structural steel in the tidal and near-shore marine atmospheric zones taken from a variety of sources are re-examined and then interpreted in the light of a phenomenological model recently proposed for marine immersion corrosion loss

[53]. Because all factors influencing corrosion have inherent uncertainties, this model is probabilistic, composed of a mean-value function and a zero-mean uncertainty function. The mean-value function is nonlinear in time, controlled by distinctly different phases- kinetic, diffusion, and anaerobic. According to the author, the model predicts a relatively short initial period followed by oxidation as the governing corrosion process. Eventually, this is overtaken by anaerobic corrosion. The author point out that the data trends for tidal and atmospheric corrosion show a surprising degree of consistency with the model, suggesting that longer-term marine corrosion involves some degree of participation of anaerobic corrosion processes. This conclusion is supported by the gradual transition from immersion through tidal to atmospheric corrosion behaviour evident in the data available in the literature.

5 MASONRY STRUCTURES

5.1 Overview

Environmental action and its effects, discussed in Section 2, trigger mechanisms in masonry that lead to its decay and weathering. As mentioned in that section, the various forms of decay can be classified into two main groups, depending on whether the transformations induced in the material are physical or chemical.

Physical processes cause masonry disintegration or breakage without affecting its chemical or mineralogical composition. The stone cracks in these processes, i.e., it crumbles into smaller size materials, favouring subsequent erosion and transport. The rock undergoes physical but no chemical change. These processes are the result of environmental conditions such as the presence of water, temperature variations, wind or soluble salts.

Chemical decay, in which reactions transform rock chemistry, depend on the composition of the minerals present, water transport and the dissolved chemicals. Water is the most important agent in chemical deterioration because its great bipolarity makes it a highly effective solvent and because its molecules are able to form hydrogen bonds [54].

Sections 5.2 and 5.3 summarise the key mechanisms involved in physical and chemical decay in masonry infrastructures. A brief description is given of the models presently in place for some of these agents, such as variations in temperature and humidity, as well as salt crystallisation. Damage induced by foundation settlement is an important issue in masonry structures. However, this topic is not addressed in the following since the present chapter is mainly dedicated to physical and chemical degradation mechanisms of infrastructure systems. Information on settlement induced damage is given in Handbook 2, within the framework of the case study on the Medicean aqueduct in Pisa. Damages induced by humans, on the other hand, are briefly mentioned in the present section.

5.2 Physical Processes

5.2.1 Wind Action

Wind-blown particles impacting masonry cause abrasion wear. The degree of decay depends in part on the kinetic energy of the particles, in turn dependent upon mass and velocity, and in part on the nature of the impacted material. The size of the particles in suspension is governed by wind velocity, a climatic variable that varies with infrastructure location and exposure.

The most common effects of wind action are surface decay and honeycombing. In combination with soluble salt action, wind can constitute a very powerful destructive agent [54].

5.2.2 Freeze-Thaw

Water seeping into pores and cracks in the stone is very harmful if it freezes, for the shift to a solid state entails an increase in volume of 9 %. This rise in volume, together with crystallisation, can induce further cracking or even disintegration. The effect is accentuated when the decline to freezing temperatures occurs abruptly.

The smaller the pores in rock, the lower is the temperature needed, given the difficulty in forming the minimum number of ice crystals required for the full volume to freeze. Other thermodynamic theories have been put forward to explain the lower temperatures needed for water to freeze under these conditions, however. Moreover, contraction of the rocky material itself at lower temperatures would intensify the effects of the abrupt freezing of the water in its pores [54].

Figure 3 Examples of masonry damaged by freeze-thaw effect

5.2.3 Temperature or Moisture Variations

5.2.3.1 Description

Abrupt temperature change may cause the detachment and dropping away of masonry surface layers, a development governed by two mechanisms that may occur simultaneously. On the one hand, the low thermal conductivity of these materials may give rise to differential expansion between the warmer outer and cooler inner layers of the stone. The resulting tangential stress may exceed the material's ultimate limit stress, inducing spalling and flaking (cracks parallel to the surface). On the other, the differences in expansion coefficients among the constituent minerals, including the pore solution, generate internal stress that may lead to similar damage. Moreover, minerals such as carbonates, the main component of limestone, are highly anisotropic.

Another consequence of temperature variations in masonry infrastructures is parallel cleavage, separating parts of the ashlar that were originally a single block. Such cleavage adopts different forms and sizes and depending on the intensity is classified as scaling or spalling. Scaling entails the peeling away of flakes of the same nature as the rock as a result of changes in temperature or freeze-thaw cycles. Spalling involves the peeling away of large, usually stiff flakes up to several millimetres thick, of the same material as the ashlar [54].

Figure 4 Example of cleavage in masonry [54]

Cleavage is generally induced by differential expansion on the ashlar surface due to solar radiation, but it may also be the result of wet-dry or freeze-thaw cycles, crystallisation or hydration pressure, or confinement of salts immediately underneath the ashlar surface.

Changes in saturation cause rock to swell [55]. Water ingress takes place through pores and cracks by means of capillary action, generating a meniscus whose radius grows with the amount of water absorbed and causing the rock to swell. If clay is present in the pores or cracks or if the rock itself contains clay minerals, this effect is greatly intensified by the repulsion forces between two adjacent layers of clay.

5.2.3.2 Modelling

The physical and chemical processes involved in moisture behaviour are complex, rendering a full understanding of its dynamics elusive. Moreover, the signs of decay in masonry may vary depending on its structural and material properties. A number of studies have shown that insufficient knowledge of damp has often led to inadequate remedial measures, accelerating masonry deterioration. Wet-dry cycles play a key role in water movement inside construction materials [56]. Water induces decay through silicate mineral hydrolysis, carbonate dissolution and the formation of gypsum crust. It is also instrumental as the vehicle for mobilising the dissolved salts involved in salt weathering, a key cause of decay in many masonry structures and monoliths [57].

Rising damp is the common term for the gradual upward movement of moisture in the lower parts of walls and other ground-supported structures [58]. Such capillary movement governs moisture transfer to and from masonry structures. Water can reach and dampen a masonry structure through the pore system by the capillary suction of ground moisture, rain or the condensation of the humidity in the air. Evaporation, through which water exits the

structure, is impacted by many parameters, including the capacity of the structure to store moisture and the microclimatic conditions prevailing at the site [56].

As noted in [56], while a host of computer-based tools designed to understand hydrothermal behaviour in masonry are described in the literature, many entail considerable mathematical sophistication, complex measuring instruments and a great deal of information and data, all of which are obstacles to their practical application. Process essentials can be analysed with simpler models, however, that cover key dynamics and their physical interrelationships with no need to use computer-based numerical methods.

One popular group of simplified models is based on the sharp front (SF) theory, which ignores the rather fuzzy boundary between wet and dry regions in a structure, replacing it with a notional sharp boundary. A so-called lumped element model, it calculates how the boundary moves across the structure in the presence of an external supply of water that may subsequently evaporate [56]. Sharp front (SF) moisture dynamics theory informs an analytical model of capillary rise in porous materials affected by both gravitational drainage and evaporation. The model calculates water movement in monoliths (single blocks of stone) and masonry structures, establishing several important scaling relations [58]. In SF models the height of rising damp is clearly defined as the position of the interface between the wet and dry zones. Water evaporates from the wet vertical surface [57]. Such models are also useful for analysing geometrical complications and composite structures comprising different materials [56].

The model described in [58] was modified in [57] by considering evaporation to be uniformly distributed across the wet region. The predictions of this latter model have proven to be consistent with empirical observations. Furthermore, the influence of precipitation and transpiration can be taken into account [57].

Joint moisture and heat transfer is instrumental in most of the aforementioned deterioration mechanisms in brick, concrete and limestone masonry. Several studies have been published on modelling moisture transport across masonry structures and the mechanical stress induced. The mathematical model proposed in [59] addressed hydro-thermo-mechanical factors as well as water vapour transfer across layered masonry structures. It was designed to evaluate hydro-thermal-induced stress in such structures under normal conditions. Relative humidity has been chosen as an appropriate continuous state variable in the model. It also considers the high nonlinearity of heat and moisture transfer across masonry materials as well as of the hydric properties.

The mathematical model developed was applied in non-linear finite element (FE) analysis to calculate the hydro-thermal stress exerted on the layered pore structure in masonry. Mechanical, thermo-mechanical and thermo-hydro-mechanical simulations were performed separately for moisture and heat transfer, assuming steady-state conditions, to determine the contribution of each. Lastly, since moisture and heat migration across porous masonry are time-dependent, transient thermo-hydro-mechanical nonlinear FE analysis was also performed. The numerical study confirmed that hydric effects are the major factor in thermo-hydro-mechanical degradation [59].

5.2.4 Salt Crystallization

5.2.4.1 Description

One of the main causes of masonry stone weathering is soluble salt crystallisation. This mechanism, which may ultimately prompt complete stone decay and destruction, is impacted by the nature of the substrate, climatic conditions and salt-system characteristics.

Therefore, textural anisotropy, pore size and distribution in the stone, temperature and relative humidity and saline system solution chemistry govern the processes taking place [60].

Soluble salts have a dual effect on stone. Firstly, they set off chemical decay by reacting with its components and secondly and much more importantly, they induce mechanical disintegration due to crystallisation-related processes.

Figure 5 Examples of salt crystallization effect

The soluble salts present in a stone's pore system may have several origins [54]:

1. The stone may be attacked by the pollutants in an acid atmosphere, such as calcium carbonates that are transformed into gypsum in carbonated rock or hydrolysed feldspar that in acid media releases alkalis and calcium that combine with sulphate ions to produce soluble salts.
2. The original stone may contain salts in its composition.
3. Materials used in previous interventions, such as portland cement, may contain a certain amount of soluble salts. In other cases, scantly suitable substances may have been applied to the rock, such as acid or alkaline cleaners, consolidants or water repellents, with compounds that subsequently form soluble salts.
4. Groundwater rising due to capillary attraction may carry soluble salts from fertilisers, de-icing salts, pesticides and so on.

Disintegration is triggered when the soluble salts present in a stone's pore solution crystallise as the water evaporates. If evaporation takes place on the surface, superficial decay appears in the form of efflorescence, crusts or patinas, depending on the nature of the salt. If the surface dries and crystallisation proceeds inward, more intense decay is induced by the formation of cryptofluorescence and hardened crusts. Cryptofluorescence generally appears on surfaces exposed to the wind, where evaporation is greater [54].

Salt cryptofluorescence may lead to substantial mechanical decay if the state of hydration is altered. Increases in humidity may provoke severe changes in volume, exerting strong pressure on pore walls that cause failure-induced relaxation. If reiterated, this process may ultimately lead to crumbling under a hardened surface crust, which may eventually fall away.

On occasion, masonry elements may disappear altogether under the effect of advanced salt crystallisation-mediated crumbling or sugaring, combined with other mechanisms such as freeze-thaw, differential thermal expansion, swelling due to differences in hydration states or cement dissolution (Figure 6).

Figure 6 Examples of advanced stages of crumbling and the loss of masonry elements [54]

Another common type of decay associated with salt crystallisation and wind action is honeycombing. In infrastructures with exposed areas, the drying front is very likely to be beneath the surface due to wind action. Moreover, if drainage is not in good condition, which is often the case in bridges, water may seep from the extrados to the intrados across the filler, usually via masonry joints, which are the most severely affected areas. The interface between intra-stone water and the drying front is where the material begins to decay and honeycombing to possibly form. Once the process has begun, the swirls caused by the wind inside the cavities accelerate the damage. The most harmful salts are sodium and magnesium sulphates [54].

5.2.4.2 Modelling Salt Weathering and Its Effects

As noted earlier, salt weathering is a major cause of masonry infrastructure deterioration. Although salt damage has been intensively researched for several decades, the mechanisms and factors that control the formation of salt crystals in porous media and the development of damage by crystal growth are poorly understood. For many years, scant progress was made in the development of realistic models for combined moisture and ion transport, primarily for want of adequate and reliable experimental data [61]. That notwithstanding, some models that jointly calculating moisture and ion transport in porous materials do take salt crystallisation into consideration, as well as the effect of salt precipitation on intra-material transport (see [62]).

A number of salt damage mechanisms and the respective theoretical models proposed to explain the widespread damage observed in field and laboratory studies of the problem were reviewed in [63]. Salt weathering was reported to be primarily due to the build-up of crystal linear growth pressure (generally known as crystallization pressure) against pore walls when salt nucleation and growth take place in a confined space. The review covers models proposed from 1949 to 1997.

The severe consequences of salt crystallisation-induced deterioration prompted the series of experiments described in [63], designed to ascertain the mechanisms involved in salt damage, identify the controlling parameters and acquire a fuller understanding of salt weathering. The ultimate aim is to develop new engineering and heritage conservation approaches to mitigate the problem. Micro- and macroscale experiments were also conducted to study the damage caused by two typical salts (sodium chloride, NaCl, and sodium sulphate, Na_2SO_4) in porous stone at high and low relative humidity.

The experimental results obtained in [63] were compared to the findings delivered by two of the models reviewed. According to the authors, the model for salt damage proposed in [64] did not appear to explain their experimental observations. The original model for crystallisation pressure proposed in [65] and supplemented in [66] did provide a good fit to the empirical data, however. The authors further noted that a model able to explain the structure and rate of some types of natural weathering as well as the modes and rates of decay in stone structures affected by salt weathering is instrumental to the design of conservation procedures for the protection of the cultural heritage.

A more recent mathematical model for mechanical stress attributable to crystallisation pressure was developed and discussed in [67]. Based on crystal growth kinetics for different pore sizes, and the assumption that nucleation and crystal growth take place in large and small pores simultaneously, the model calculates the distribution of salt crystals in different size pores and the resulting surface load on crystals under stress. Crystallisation pressure is calculated as a function of solution supersaturation. Average mechanical stress is then found with a uniform cross-sectional force in a given volume. The crystal growth model proposed was used to calculate the mechanical stress induced by potassium nitrate and sodium sulphate crystallisation in brick samples. A comparison of the analytical to the experimental results showed that the model was able to approximate the strain on the material induced by salt crystallisation in the selected bricks. The authors nonetheless identified a need to improve the model to better predict mechanical stress during drying, for instance, or to simulate the re-allocation of salt crystals from stressed to stress-free pores.

In another vein, the particularly random nature of the occurrence of critical salt crystallisation attacks suggests that masonry decay might profitably be broached probabilistically. Further to that conceit, decay was modelled in [68] from two stochastic approaches:

1. In one, deterioration $L(t,l)$ over time is defined as a stochastic process affecting random variable l, which represents the loss of surface material. The reliability function $R(t)$ is defined as the probability that the system does not fail during time t and $\overline{R}(t)$ is assumed to be the probability that the system will exceed a given significant damage threshold $\overline{L}$ in time t.
2. In the other, deterioration is seen as weakened performance that can be defined as a change in system service state. When the system fails it shifts from current service state i to another state k characterised by a lower level of performance. Deterioration may therefore be defined as system *transition* through different service states due to a discrete number of attacks over time continuum t. Deterioration $L(t)$ over time is defined as a stochastic process affecting random variable τ_i, where τ_i is the service "lifetime" of the system, defined as the time the system remains in a given performance state i before moving on to the next performance state.

The two stochastic approaches were compared in [68]. The authors reported that success depended on the type of the problem studied and the importance for satisfactory modelling of an understanding of the physics of the events analysed. The findings also showed that both approaches were able to predict, in probabilistic terms, the magnitude of the expected salt crystallisation damage over time and the time needed to reach a given damage level.

5.2.5 Damage due to Human Activities

In spite of the importance of the abovementioned degradation mechanisms, most causes of deterioration of masonry structures are human-made. Indeed, often new openings are created for different purposes, thus reducing the load bearing cross-sections. Furthermore, in many cases such interventions are not compliant with the rules of good engineering practice. In other cases, damage is induced by improper use of the construction work or by accidents. Figure 7 shows a typical example from the Medicean aqueduct in Pisa (Handbook 2), an arch damaged by vehicle impact.

Figure 7 Typical damage pattern from vehicle impact on the Medicean aqueduct in Pisa

5.3 Chemical Processes

5.3.1 Dissolution

This is the most important mechanism and affects all manner of masonry, for carbonates, the cement matrix in sandstone, as well as gypsum, to name a few, are all subject to dissolution.

Rain water may dissolve the CO_2 in the atmosphere to form carbonic acid, enhancing its ability to dissolve the calcium carbonate that constitutes the cement matrix in sandstone by transforming it to soluble bicarbonate [54]. This obviously has an adverse impact on bridges built with that type of masonry.

The cement matrix in the sandstone and conglomerate stone used to build some masonry bridges may lose its binding properties, normally due to dissolution, and consequently the cohesion among ashlar constituents. This development causes substantial damage and may even entail the detachment of stone material. The loss of the inner cement matrix due to the ingress of moisture through stone pores weakens the mechanical strength of the masonry material, which ultimately converts into loose sand [54].

5.3.2 Carbonation

The mechanisms involved in carbonation are not generally very harmful in themselves, although they may be symptomatic of other more damaging processes. Lime and sandstone have calcareous cement matrices whose main component is calcium carbonate in the form of calcite. This makes them particularly vulnerable to carbonation.

The chemical conditions that may be present in carbonate rock alteration include:

1. Dissolution and precipitation of the carbonates in the rock with declining temperatures and rising pressure
2. Presence of calcium together with the rise in CO_2 pressure in water and declining temperatures
3. Acidity of water, which conditions the SO_4^{2-} attack

The conclusion to be drawn from the foregoing is that carbonate rock is particularly vulnerable to cold acid rain with a high CO_2 content.

5.3.3 Biological Actions

Degradation of masonry may also be induced through biological organisms growing on the surface of walls. Algae, fungi, lichens, bacteria and other organisms of small dimensions may penetrate the pores where they produce and deposit organic acids with effects which are comparable to those from atmospheric pollution. Roots of larger plants may penetrate the mortar joints, where their growth generates pressures which may lead to the disintegration of the joints themselves. Even birds may seriously damage masonry walls through repeated picking with their beaks.

Figure 8 Examples of masonry walls infested by vegetation

6 INSPECTION AND MAINTENANCE

6.1 Introduction

Infrastructures are ideally designed to ensure an economical operation throughout the anticipated service life in compliance with given requirements and acceptance criteria. Such acceptance criteria are typically related to the safety of persons and risk to environment. Deterioration processes such as abrasion, corrosion or crack growth, will always be present to some degree and depending on the adopted design philosophy in terms of degradation

allowance and protective measures the deterioration processes may reduce the performance of the system beyond what is acceptable. In order to ensure that the given acceptance criteria are fulfilled throughout the service life of the infrastructures it may thus be necessary to control the development of deterioration and if required to install corrective maintenance measures. In usual practical applications inspection is the most relevant and effective means of deterioration control [69]. In accordance with [6] a detailed assessment of an existing structure should be initiated when the result of a preliminary inspection indicates that there is uncertainty in the actions, action effects or properties of the structure, or that structural deterioration due to time-dependent actions (e. g. corrosion, fatigue) or structural damage by accidental actions have been detected.

Even though inspections may be used as an effective mean for controlling the degradation of the considered infrastructure and therefore imply a potential benefit, they may have a considerable impact on the operation of the infrastructure, including direct and indirect economic consequences. For this reason it is necessary to plan the inspections in such a way that a balance is achieved between the expected benefit and the corresponding economic consequences implied by the inspections themselves. A risk-based approach for inspection and maintenance planning is briefly described in the following section.

6.2 Risk-Based Inspection Planning

The degradation processes acting on a given infrastructure are usually subject to substantial uncertainty originating in the loading and exposure conditions as well as the material characteristics of its components. On the other hand, inspection and maintenance actions are also subject to large uncertainties. The quality of inspections normally is quantified in terms of their ability to detect and quantify the dimension of the defects of consideration. For these reasons, when inspection planning for infrastructures is considered it is important to take all these uncertainties into consideration, as they will strongly influence the performance of the infrastructure [70].

Inspections may be used as a tool to reduce the uncertainty in the predicted deterioration process and/or as a mean of identifying deterioration before it becomes critical. Traditional inspection planning approaches are based in prescriptive rules and leave little possibility to adapt the inspection effort to the actual condition of the components of an infrastructure or the importance of the component for the operation of the infrastructure [69]. During the last two decades a large amount of attention has been paid to the development of optimal plans for inspection and maintenance of structures. On the basis of the classical decision theory [71, 72] and efficient reliability methods [73, 74], reliability-based and risk-based approaches for inspection planning have been developed.

Risk-based approaches for inspection planning provide a rational basis for adapting the inspection to the condition of the components of an infastructure and for prioritising inspection efforts in accordance with the importance of the individual components and the different deterioration mechanisms. Risk-based inpection planning approach gives a rational and cost efficient decision framework for determining:

- What to inspect
- How to inspect
- Where to inspect
- How often to inspect
- Actions to be taken depending of the inspection results

and at the same time ensuring and documenting that requirements to the safety of personnel and environment are fulfilled [71].

The number of decision variables and dependencies underlying an inspection planning problem is rather large and calls for a systematic treatment. Therefore, the inspection and maintenance planning is formulated as a problem where the overall service life costs are minimized.

The inspection decision problem may be represented as shown in Figure 3. The basis for determination of the cost optimal inspection and maintenance strategy is pre-posterior analyses from classical decision theory [71, 72]. In the most general case the parameters defining the inspection plan are collected in the vector:

$$\mathbf{i} = (\Delta\mathbf{t},\ \mathbf{l},\ \mathbf{r})^T \tag{20}$$

where

$\Delta\mathbf{t} = (\Delta t_1, \ldots \Delta t_N)^T$ are the time intervals between the inspections;
N is the number of inspections;
$\mathbf{l} = (l(t_1),\ \ldots,\ l(t_N))^T$ are the locations to inspect;
$\mathbf{r} = (r_1,\ \ldots,\ r_N)^T$ is the reliability (quality) of inspections.

The inspection results are uncertain due to the fact that they depend not only on the uncertain performance of the inspection itself but also on the uncertain state of degradation. The uncertain inspection results (Figure 9) are modelled by the random vector $\mathbf{S} = (S(t_1),\ \ldots,\ S(t_N))^T$ in which the individual components refer to the results obtained from the inspections at the different locations $\mathbf{l}(t_i)$. Furthermore $d(s)$ is a decision rule defining the repair action to take depending on the inspection result and ϑ is the realization of the uncertainties influencing the state of the system [75].

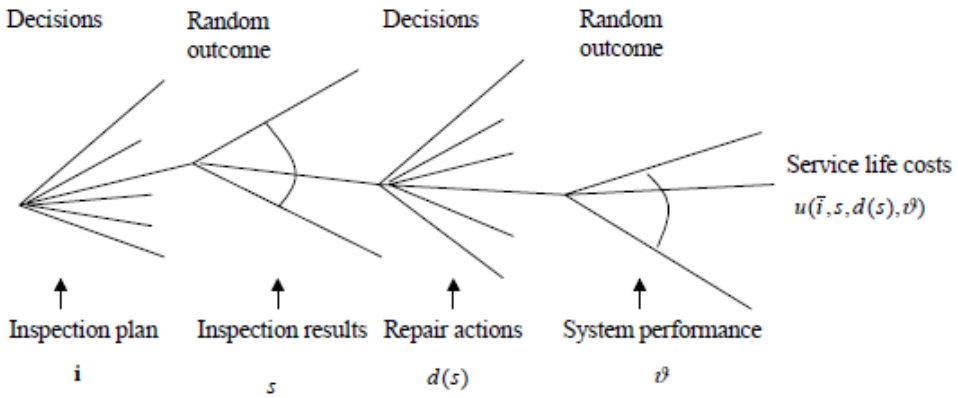


Figure 9 Inspection planning decision tree [75]

The utility associated with the inspection plan and the repair decision rule is denoted $u(\mathbf{i}, \mathbf{S}, d(\mathbf{S}), \vartheta)$ and the optimal inspection may be determined as the plan which maximizes the expected utility. Usually the expected utility may be readily associated with the service life costs and the optimization problem reformulated as a minimization problem, in which the total expected costs are minimized. If the the total expected costs are divided into inspection, repair and failure costs and a constraint related to the minimum level of service life reliability is added the optimisation problem is [75]:

$$\min_{\mathbf{i},d} C_T(\mathbf{i},d) = \left(C_I(\mathbf{i},d) + C_R(\mathbf{i},d) + C_F(\mathbf{i},d)\right), \quad \beta(T,\mathbf{i},d) \geq \beta_{\min} \tag{21}$$

where

$C_T(\mathbf{i},d)$ total expected costs;

$C_I(\mathbf{i},d)$ expected inspection costs;

$C_R(\mathbf{i},d)$ expected repair costs;

$C_F(\mathbf{i},d)$ expected failure costs;

$\beta(T)$ generalised reliability index defined by $\beta(T) = -\Phi^{-1}(P_F(\mathrm{T}))$;

$P(T)$ system failure probability in a specified reference period *T*, such as one year or the service life;

$\beta_{\min}$ minimum reliability index in the specified reference period *T*.

The optimization problem expressed by (21) is numerically very difficult to solve due to the large number of variables involved and uncertainties that should be considered (see Chapters 5 and 8). Furthermore, sometimes it can be difficult to obtain the data to model the costs needed in a cost optimal inspection planning. The numerical effort in determination of an optimal inspection plan can be significantly reduced if a simplified inspection planning is performed using a minimum reliability level (normally code specified) and estimates of the reliability as a function of time, taking inspections into account [76]. The description of the simplified procedure for inspection planning is beyond the scope and will not be dealt with in this handbook. For more information about this procedure see e. g. [77] and [78].

7 CONCLUDING REMARKS

In this chapter, the modelling of relevant degradation processes that affect RC, steel and masonry infrastructures is presented. Depending on the nature of the changes that take place in the material under the effect of environmental actions, deterioration processes may be classified in physical and chemical. The following concluding remarks are in order:

- Modelling the deterioration process involves significant uncertainties associated with the environmental conditions, material properties, predictive models, material testing and inspection methods. Therefore, deterioration processes acting on aging infrastructures should be described in probabilistic terms.
- Corrosion of reinforcing steel in concrete due to the action of chlorides and atmospheric carbon dioxides is generally considered as the main degradation process for RC infrastructures. The modelling of time-dependence of the corrosion process is described using a two-phase model, including the initiation and the propagation phase.
- The main consequences of reinforcement corrosion are: reduction in bar cross section, change of reinforcing bar material properties; concrete cover deterioration and bond deterioration.
- The modelling of another important chemical degradation mechanisms affecting RC infrastructures such as sulfate attacks and alkali-aggregate-reactions have been also briefly described, as well as the modelling of physical deterioration processes such as freeze-thaw attack and abrasion.
- 'Uniform' corrosion loss of material is of the most interest for overall degradation of strength of structural steel. Modelling of the loss of material due to 'uniform' corrosion in atmospheric environment is briefly described.

- Marine environments are, generally, the most aggressive towards steel atructures. Corrosion in splash/tidal zone is usually the most severe and difficult to protect. Despite some quite extensive, long-term experimental test programs, precise understanding of the mechanics of marine corrosion has been slow to develop and mathematical formulation has been largely neglected. However, some probabilistic models to describe deterioration effect for tidal zone corrosion of steel structures in marine environments can be found in the literature.
- Among the chemical deterioration processes affecting masonry infrastructures, the phenomena associated with cement, carbonate and gypsum dissolution, as well as carbonation are the most relevant. Thaw freezing, temperature and moisture variations and salt crystallization are the most important physical deterioration processes that affect masonry infrastructures. Some advances in modelling of these two latter mechanisms are briefly described.
- Inspections may be used as an effective tool for reducing the uncertainties in the degradation process and/or as a mean of identifying deterioration before it becomes critical. The general formulation of a risk-based approach for inspection and maintenance planning has been briefly described. Risk-based inpection planning approach gives a rational and cost efficient decision framework for determining what to inspect; how to inspect; where to inspect; how often to inspect and the actions to be taken depending of the inspection results. However, practical implementation of risk-based approach for inspection planning requires numerical operations to be simplified.

REFERENCES

[1] Val D. (2012) Environmental effects, JCSS Probabilistic Model Code Draft document

[2] ISO 13823:2008, General principles on the design of structures for durability, International Organization for Standardization ISO, Geneva, Switzerland (2008).

[3] EN 1992-1-1, Design of concrete structures – General rules and rules for buildings, European Committee for Standardization, CEN, Brussels (2004).

[4] Fib, Model code for concrete structures 2010, Fédération Internationale du Béton, fib, Lausanne, Verlag Ernst & Sohn, Berlin (2013).

[5] EN 1993-1-1, Design of steel structures – General rules and rules for buildings, European Committee for Standardization, CEN, Brussels (2005).

[6] ISO 13822:2001(E), Bases for design of structures − Assessment of existing structures, International Organization for Standardization, Geneva, Switzerland (2001).

[7] ISO 2394:1998 (E) General principles on reliability for structures, International Organization for Standardization, Geneva, Switzerland, (1998).

[8] JCSS, Probabilistic model code (2001) http://www.jcss.byg.dtu.dk.

[9] Basheer L., Kropp J. and Cleland D. J. Assessment of the durability of concrete from its permeation properties: a review (2001), *Construction and Building Materials*, Vol. 15, pp. 93-103.

[10] Fagerlund G. The critical degree of saturation method of assessing the freeze/thaw resistance of concrete (1977), *Materials and Structures*, Vol. 10, No 58.

[11] Fib, Code-type models for structural behaviour of concrete: Background of the constitutive relations and material models in the fib model code for concrete structures 2010, Féderation Internationale du Béton, Bulletin 70, Laussane, Switzerland (2013).

[12] Tuutti K. (1982) Corrosion of steel in concrete, Swedish Cement and Concrete Institute, CBI, Nº 4-82, Stockholm.

[13] DURACRETE, Probabilistic performance based durability design of concrete structures – Modelling of degradation, The European Union – Brite EuRam III, Contract BRPR-CT95-0132, Project BE95-1347, Document BE95-1347/R4-5 (1998).
[14] Bakker R. F. M. (1994) Prediction of service life reinforcement in concrete under different climatic conditions at given cover. Corrosion and Protection of steel in concrete, Internacional Conference, Sheffield, UK, Swany R.N. Ed.
[15] CEB Task Group V (1996).
[16] Fib, Model code for service life design, Féderation Internationale du Béton, Bulletin 34, Laussane, Switzerland (2006).
[17] DARTS, Durable and reliable tunnel structures: Deterioration modelling, European Commission, Growths 2000, Contract GIRD-CT-2000-00467, Project GrDI-25633 (2004).
[18] DURACRETE, Probabilistic performance based durability design of concrete structures – Statistical quantification of the variables in limit state functions, The European Union – Brite EuRam III, Contract BRPR-CT95-0132, Project BE95-1347, Document BE95-1347/R9 (2000).
[19] CONTECVET, A validated user's manual for assessing the residual service life of concrete structures, EC Innovation Program IN30902I (2001).
[20] Croce P. and Holický M., Eds. (2013) Operational Methods for the assessment of existing structures, Pisa and Prague.
[21] Andrade C., Alonso C., Garcia D. and Rodríguez J. (1991) Remaining lifetime of reinforced concrete structures: Effect of corrosion on the mechanical properties of the steel. Int. conf. on life prediction of corrodible structures, NACE, Cambridge, UK.
[22] García M. D. (1995) Aportaciones al comportamiento resistente de estructuras de hormigón armado afectadas por la corrosión de sus armaduras. Tésis Doctoral. ETS Arquitectura, UPM.
[23] Stewart M. G. (2009) Mechanical behaviour of pitting corrosion of flexural and shear reinforcement and its effect on structural reliability of corroding RC beams, *Structural Safety*, 31(1), pp. 19-30.
[24] Muttoni A., Schwartz J. and Thürlimann B. (1997) Design of concrete structures with stress fields, Birkhäuser Verlag, Berlin.
[25] Prieto M., Tanner P., Andrade C. and Fernández M. (2013) Experimental and numerical study of bond response in structural concrete with corroded steel bars, IABSE Conference on assessment, upgrading and refurbishment of infrastructures, Rotterdam.
[26] Muttoni A., Fernández Ruiz M. et Kostic N. (2011) Champs de contraintes et méthode des bielles-et-tirants, Laboratoire de construction en béton IBETON-ENAC, École Polytechnique Fédérale de Lausanne, EPFL.
[27] Tanner P., Lara C. and Prieto M. (2011) Semi-probabilistic models for the assessment of existing concrete structures. In: Applications of Statistics and Probability in Civil Engineering, Taylor & Francis Group, London.
[28] Schneider J. (1994) Sicherheit und Zuverlässigkeit im Bauwesen – Grundwissen für Ingenieure. Verlag der Fachvereine AG, Zürich und Teubner Verlag, Stuttgart.
[29] Rodriguez J., Ortega L. M. and Casal J. (1995) Relation between corrosion and load bearing capacity on concrete beams. Brite Euram Project BREU-CT-0591.
[30] Benchmark des poutres de la Rance, Revue européenne de génie civil, Volume 11, No. 1-2, Paris (2007).
[31] Tanner P., Lara C. and Hingorani R (2007). Seguridad estructural. Una lucha con incertidumbres, *Hormigón y Acero* No. 245, Madrid, pp. 59-78.

[32] Zhang M., Chen J., Lv Y., Wang D. and Ye J. (2013) Study on the expansion of concrete under attack of sulphate and sulphate-chloride ions, *Construction and Building Materials*, Vol. 39, pp. 26-32.
[33] THE CONCRETE SOCIETY, Diagnosis of Deterioration in Concrete Structures – Identification of Defects, Evaluation and Development of Remedial Action, Technical Report No. 54, , Century House, Berkshire, United Kingdom (2000).
[34] Santhanam M., Cohen M. D. and Olek J. (2001) Sulphate attack research – whither now?" *Cement and Concrete Research*, Vol. 31, pp. 845-851.
[35] Tixier R. and Mobasher B. (2003) Modelling of damage in cement-based materials subjected to external sulfate attack. I: Formulation, *ASCE Journal of Materials in Civil Engineering*, Vol. 15, No. 14, pp. 305-313.
[36] Skalny J., Marchand J. and Odler I. (2002) Sulphate Attack on Concrete. Spon Press, London, New York.
[37] Idiart A., López C. M. and Carol I. (2011) Chemo mechanical analysis of concrete cracking and degradation due to external sulphate attack: A meso-scale model, *Cement and Concrete Composites*, Vol. 33, pp. 411-423.
[38] Swamy R. N. (1989) Structural implications of alkali silica reaction. Proc. Of 8th International Conference on Alkali-Aggregate Reaction, Kyoto, Japan, Elsevier Applied Science Ltd., London, United Kingdom.
[39] Mori Y. (2005) Reliability-based service life prediction and durability in structural safety assessment. In: Structural Safety and its Quality Assurance, Structural Engineering Institute of ASCE.
[40] Melchers R. E. (2003) Probabilistic model for marine corrosion of steel for structural reliability assessment, *ASCE Journal of Structural Engineering*, Vol. 129, No. 11, pp. 1484-1493.
[41] Wranglen G. (1972) An introduction to corrosion and protection of metals, Institut för Metallskydd, Fack 10041, Stockholm.
[42] Albrecht P. (2003) Atmospheric corrosion resistance of structural steels, *ASCE Journal of Materials in Civil Engineering*, Vol. 15, No. 1, pp. 2-24.
[43] Melchers R. E. (1999) Structural reliability analysis and prediction, 2nd ed., John Wiley and Sons, New York.
[44] Melchers R. E. (2008) Development of new applied models for steel corrosion in marine applications including shipping. In: Ship and Offshores Structures, 3:2, pp. 135-144, Taylor and Francis.
[45] Benarie M. and Lipfert F. L (1986) A general corrosion function in terms of atmospheric pollutant concentrations and rain pH", *Atmospheric Environment*, Vol. 20, No. 10, pp. 1947-1958.
[46] Feliu S., Morcillo M. and Feliu Jr. S. (1993) The prediction of atmospheric corrosion from meteorological and pollution parameters - 1. annual corrosion", *Corrosion Science,* Vol. 34, No. 3, pp. 403-422.
[47] De La Fuente D., Díaz I., Simancas J., Chico B., and Morcillo M. (2011) Long-term atmospheric corrosion of mild steel, *Corrosion Science,* Vol. 53, pp. 604-617.
[48] Mccuen R. H. and Albrecht P. (1993) Composite modelling of atmospheric corrosion penetration data. In: Application of Accelerated Corrosion Tests to Service Life Prediction of Materials, ASTM STP1194, American Society for Testing and Materials, Philadelphia, pp. 65–101.
[49] Shaw C. W. and Smith G. R. (1973) Corrosion preventing apparatus and method, US Patent No. 3719049.
[50] Melchers R. E. (1999) Corrosion uncertainty modelling for steel structures, *Journal of Constructional Steel Research*, Vol. 52, pp. 3-19.

[51] Melchers R. E. (1995) Probabilistic modelling of marine corrosion of steel specimens, Proceedings of The Fifth International Offshore and Polar Engineering Conference, Hague, The Netherlands.
[52] Melchers R. E. (2007) Transition from marine immersion to coastal atmospheric corrosion for structural steels, *Corrosion*, Vol. 63, Issue 6, pp 500-514.
[53] Melchers R. E. (2003) Modelling of marine immersion corrosion for mild and low-alloy steels-part 1: Phenomenological model, *Corrosion*, Vol. 59, Issue 4, pp 319-334.
[54] Jerez E., León J. y Martin Caro J. A. (2007) Inspección y diagnosis de puentes ferroviarios de fábrica, ADIF, Administrador de Infraestructuras Ferroviarias.
[55] García de Miguel J. M. (1999) Piedras naturales, prefabricados y otros materiales, Procedimiento y Técnicas Constructivas del Patrimonio - Master de restauración y rehabilitación del patrimonio, Madrid.
[56] D'Agostino D. (2013) Moisture dynamics in an historical masonry structure: The Cathedral of Lecce (South Italy), *Building and Environment*, Vol. 63, pp. 122-133.
[57] Hall C., Hamilton A., Hoff W. D, Viles H. A. and Eklund J. A. (2011) Moisture dynamics in walls: response to micro-environment and climate change, Proceedings of The Royal Society A, Vol. 467, pp. 194–211.
[58] Hall C. and Hoff W. D (2007) Rising damp: capillary rise dynamics in walls, Proceedings of the Royal Society A, Vol. 463, pp. 1871–1884.
[59] Ramézani H. and Jeong J. (2011) Environmentally motivated modelling of hydro-thermally induced stresses in the layered limestone masonry structures: Physical motivation and numerical modeling, *Acta Mechanica*, Vol. 220, pp. 107–137.
[60] Lopez-Acevedo V., Viedma C., Gonzalez V. and La Iglesia A. (1997) Salt crystallization in porous construction materials II. Mass transport and crystallization processes. *Journal of Crystal Growth*, Vol. 182, pp. 103-110.
[61] Pel L., Huinink H. and Kopinga K. (2003) Salt transport and crystallization in porous building materials, *Magnetic Resonance Imaging*, Vol. 21, pp. 317–320.
[62] Nguyen T. Q., Petkovic J., Dangla P. and Baroghel-Bouny V. (2008) Modelling of coupled ion and moisture transport in porous building materials. *Construction and Building Materials,* Vol. 22, pp. 2185-2195.
[63] Rodriguez-Navarro C. and Doehne E. (1999) Salt weathering: influence of evaporation rate, supersaturation and crystallization pattern. *Earth Surface Process and Landforms*, Vol. 24, pp. 191-209.
[64] Wellman H. W. and Wilson A. T. (1965) Salt weathering, a neglected geological erosive agent in coastal and arid environment. *Nature*, Vol. 205, pp. 1097-1098.
[65] Correns C. W. (1949) Growth and dissolution of crystals under linear pressure (1949) *Discussions of the Faraday Society*, Vol. 5, pp. 267-271.
[66] Weyl P. K. (1959) Pressure solution and the force of crystallization - A phenomenological theory. *Journal of Geophysical Research*, Vol. 64, pp 2001-2025.
[67] Espinosa R. M., Franke L. and Deckelmann G. (2008) Model for the mechanical stress due to the salt crystallization in porous materials, *Construction and Building Materials*, Vol. 22, pp. 1350-1367.
[68] Garavaglia E., Lubelli B. and Binda L. (May 2002) Two different stochastic approaches modelling the decay process of masonry surfaces over time, *Materials and Structures*, Vol. 35, pp. 246-256.
[69] Faber M. H. (2002) Risk based inspection: The framework, *Structural Engineering International*, Vol. 12, Nº. 3, pp. 186-194.
[70] Sorensen J. D. and Faber M. H. (2002) Codified risk-based inspection planning, *Structural Engineering International*, Vol. 12, Nº 3, pp. 195-199.

[71] Raiffa H. and Schlaifer R. (1961) Applied statistical decision theory, Harvard University Press, Cambridge University Press, Cambridge, Mass.

[72] Benjamin J. R. and Cornell A. C. (1971) Probability statistics and decision for civil engineers, McGraw-Hill Book Company.

[73] Faber M. H., Kroon I. B. and Sorensen J. D. (1996) Sensitivities in structural maintenance planning. *Reliability Engineering and System Safety*, Vol. 51, Nº 3, Elsevier, pp. 317-329.

[74] Madsen H. O., Krenk S. and Lind N. C. (1986) Methods of Structural Safety, Prentice-Hall, Englewood Cliffs, NJ.

[75] Faber M. H. (2006) Risk and Safety in Civil, Surveying and Enviromental Engineering - Lecture Notes, ETH Zürich, Switzerland. http://www.ibk.ethz.ch/emeritus/fa/education/lecture_notes/Script_Risk_Safety_april_06_web.pdf.

[76] Sorensen J. D., Faber M. H. and Kroon I. B. (1995) Optimal reliability-based planning of experiments for POD curves. Structural Reliability Theory, Paper Nº 154.

[77] Sorensen J. D., Faber M. H., Thoft-Christensen P. and Rackwitz R. (1991) Modelling in optimal inspection and repair. Proceedings of OMAE, Stavanger, Norway, pp. 281-288.

[78] Pedersen C., Nielsen J. A., Riber J. P., Madsen H. O. and Krenk S. (1992) Reliability based inspection planning for the Tyra Field", Proceedings of OMAE, Calgary, Canada.

CHAPTER 4: FATIGUE

Pietro Croce[1], Maria Luisa Beconcini[1], Carlos Lara[2], Daniele Pellegrini[1], Peter Tanner[2]

[1]Department of Civil and Industrial Engineering – Structural Division – Univ. of Pisa, Italy
[2]Eduardo Torroja Institute for Construction Sciences, IETcc-CSIC, Madrid, Spain

Summary

Fatigue is one the most relevant cause of failure of infrastructures or part of them, like bridges, cranes and machinery, offshore structures and so on. A lot of existing structures, mainly built before the 1980, have been designed using inappropriate fatigue classification of details or underestimating the aggressiveness of actual load spectra, so that they are subject to high fatigue failure risk, as demonstrated by fatigue cracks detected even recently in several structures. Fatigue assessment and evaluation of residual fatigue life is then a key topic in verification of existing infrastructure, also in view of designing repair interventions of fatigue damaged details and planning of future maintenance and inspection programmes.

1 INTRODUCTION

Fatigue is one the most relevant cause of failure of infrastructures or part of them, like bridges, cranes and machinery, offshore structures and so on.

As known, *fatigue is the progressive, localised and permanent structural change occurring in a material subjected to conditions that produce fluctuating stresses and strains at some point or points and that may culminate in cracks or complete fracture after a sufficient number of fluctuations* [1] [2].

In engineering structures, fatigue is induced by actions and loads varying with time and/or space and/or by random vibrations. Thus fatigue can be originated by natural events, like waves, wind and so on, or by loads deriving from the normal use of the structure itself, as it happens in bridges, under the action of lorries or trains crossing the bridges themselves.

Unfortunately, millions of existing structures, mainly built before the 1980, have been designed using inappropriate fatigue classification of details or underestimating the aggressiveness of actual load spectra, so that they are subject to high fatigue failure risk, as demonstrated by fatigue cracks detected even recently in several structures.

On the other hand, in the last 30 years, the fatigue behaviour of typical details has been studied so extensively that quite sound knowledge is now available in designing new structures.

Fatigue assessment and evaluation of residual fatigue life is then a key topic in verification of existing infrastructure, also in view of designing repair interventions of fatigue damaged details and planning of future maintenance and inspection programmes. An exhaustive discussion of this relevant subject is outside the scope of the present paper, therefore in the following only some significant aspect is sketched.

First of all, it must be highlighted that the evaluation of the actual fatigue damage requires the knowledge that is the a posteriori reconstruction of the load cycles which acted on the structure during its past life. As for new structures, fatigue loads can be given in terms of

load spectrum, expressing the load variation or the number of recurrences of each load level during the design life of the structure; load spectrum can be expressed by means an appropriate function, graph, histogram or table, deduced from recorded data or by comparison with similar structures. Whenever, as it happens for bridge, the actual load spectrum results so complicated that cannot be directly employed for fatigue checks, it can be replaced by some conventional load spectrum, aimed to reproduce the fatigue induced by the real one.

The evaluation of conventional load spectra is particularly thorny, because it requires considering the actions from the resistance point of view also. In fact, fatigue depends on the nature of the varying actions and loads, and additionally on structural material details, through the shape and the properties of the relevant mechanical fatigue models. When constant amplitude endurance (fatigue) limit $\Delta\sigma_D$ does not exist, like concrete, steel reinforcement or pre-stressing steel in r.c. details, conventional equivalent load spectra, reproducing the actual fatigue damage, are generally adopted, on the contrary, for structural steel or aluminium details, characterized by constant amplitude fatigue limit, it needs to distinguish between equivalent load spectra, reproducing the actual fatigue damage, and frequent load spectra, reproducing the maximum load range significant for fatigue, accordingly as fatigue verifications require cumulative damage computations or boundless fatigue life assessments [3].

In some kind of structures, like offshore structures or cranes and machinery, fatigue load spectra can be often obtained applying the powerful methods of the stochastic process theory, so that the link between the actions and the effects can be expressed by simple formulae, but these methods cannot be applied to bridges, as road or railway traffic loads induce broad band stress histories.

2 FATIGUE MODELS

In the classical approach two kind of fatigue phenomena are generally distinguished, depending on the level of stresses and on the extension of plastic area.

When the notional stress, that is disregarding the stress concentrations, are in the elastic field and the extension of the plasticized zone is very small if compared with the typical dimensions of the detail, being limited, for instance, in the vicinity of stress concentrators or at the apex of the cracks, the number of cycles to failure is high and the phenomenon is identified as *true fatigue* or *high-cycle fatigue*; on the contrary, when dimensions of plasticized zone are comparable with the typical dimensions of the detail, the number of cycles to failure is low and the phenomenon is identified as *low-cycle fatigue*.

Two types of fatigue models (summarized in the following) are generally used for the description of fatigue failures caused by fluctuating actions:

– *S-N* model based on experiments
– Fracture mechanics model

2.1 S-N Approach

An *S-N* curve or Wöhler curve is a relation between the stress range under constant amplitude loading and the number of stress cycles to failure.

Generally, the fatigue curves are reported in the $\log(\Delta\sigma)$-$\log N$ plane, commonly called *S-N* plane.

The complete Wöhler curve (Figure 1) intersects the *S*-axis for N=0,5 (a semi-cycle) at $\Delta\sigma_R$, corresponding to the static resistance of the material. The part of the curve identified by

$N<10^4$ corresponds to the *low-cycle fatigue* interval; the part of the curve identified by $10^5<N<10^6\div10^7$ corresponds to the *high-cycle fatigue* interval; the part of the curve identified by $N>10^6\div10^7$ corresponds to the constant amplitude limit, if any; while the interval $10^4<N<10^5$ is the transition zone. The extension of rounded knee of the curve corresponding to $N=10^6\div10^7$ is usually very small, and can be disregarded.

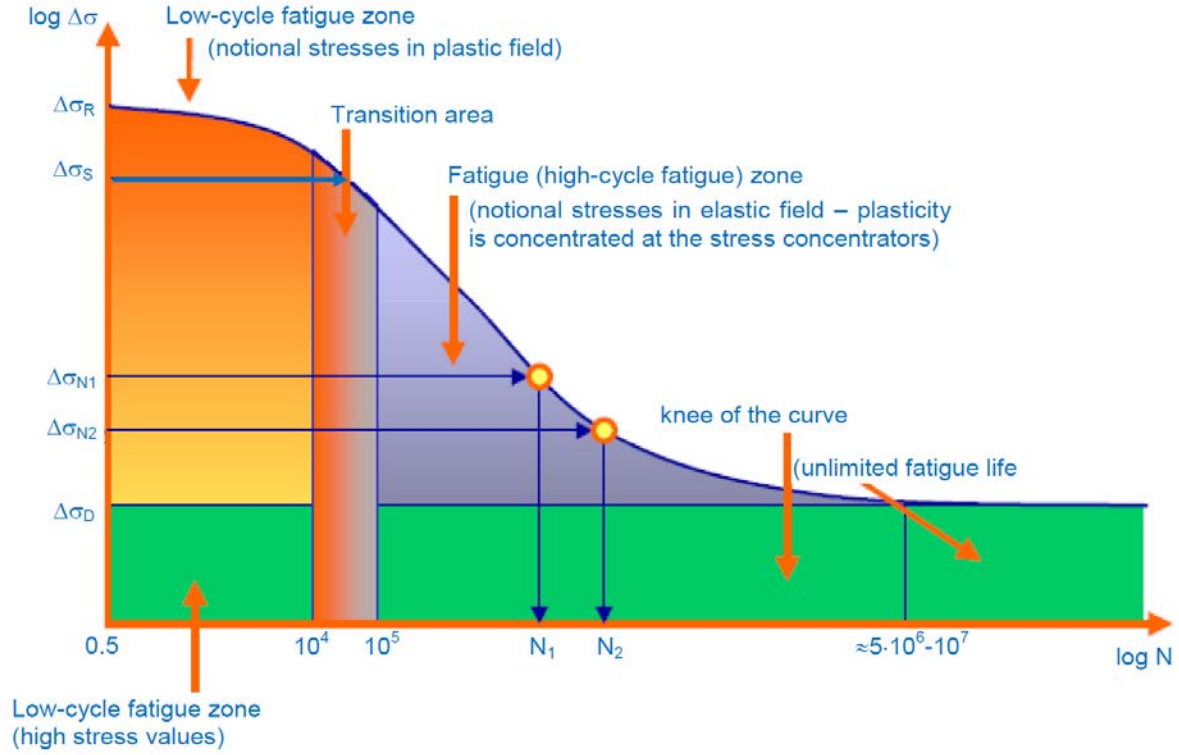


Figure 1 Complete Wöhler curve

In the *high-cycle fatigue* interval, the standard *S-N* curve is linear and it can be expressed in the form:

$$N\,\Delta\sigma^m = A \tag{1}$$

where N is the number of stress cycles to failure at a constant amplitude stress range $\Delta\sigma$, A and m are the material parameters.

In the logarithmic *S-N* chart Wöhler curve pertaining to steel details, characterised by endurance limit, are represented by a bilinear curve, characterised by a sloping branch of constant slope, m=3 (Figure 2) plus a horizontal branch, which is adopted for constant amplitude stress histories or for boundless fatigue life assessments; but, when fatigue damage is to be assessed for stress spectra characterized by varying amplitude of stress cycles, trilinear curves are introduced, characterised by two sloping branches, m_1=3 and m_2=5, (Figure 3) to take into account the influence of pre-existing fatigue damage [4] [5]. In this latter case, a cut-off limit $\Delta\sigma_L$ is introduced, corresponding at 10^8 stress cycles, which is the minimum stress range able to produce fatigue damage in a damaged detail.

In the *low-cycle fatigue* interval, instead in terms of constant stress amplitude, the material behaviour is better expressed in terms of constant plastic strain $\Delta\varepsilon_{apl}$, using the expression experimentally derived by Manson [4] and Coffin [5], whose mathematical formulation is similar to (1)

$$N^k\,\Delta\varepsilon_{apl} = C \tag{2}$$

where N is the number of cycle to failure and k and C are material constants. Equation (2) is usually called Manson-Coffin law.

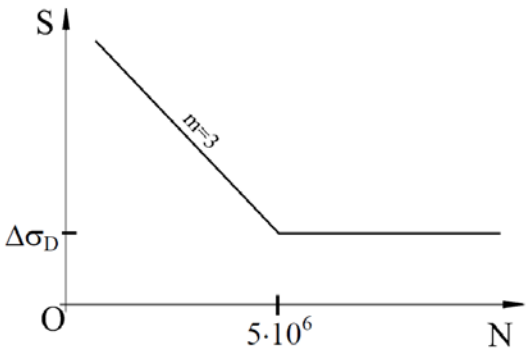


Figure 2 Bi-linear *S-N* curve

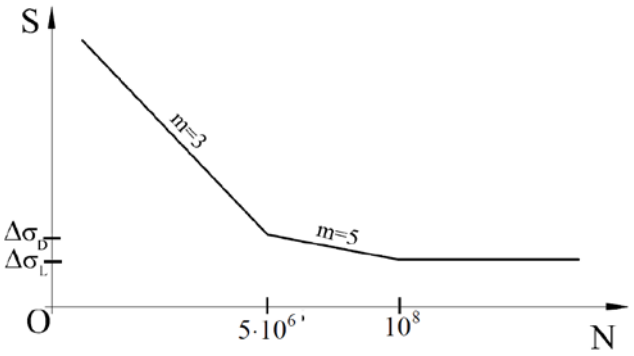


Figure 3 Tri-linear *S-N* curve

Usually, steel details are classified in fatigue using the detail category, which is the characteristic value of the fatigue resistance at $2\cdot10^6$ cycles, allowing to fully define the corresponding bi-linear or tri-linear curve. In Table 1, an example of classification of orthotropic steel deck details is reported, as derived from EN1993-1-9 [6] and EN1993-2 [7].

Unless differently stated, the *S-N* curve of a given structural detail is usually plotted in terms of notional stresses that is disregarding peak stresses, whose theoretical and experimental measure requires ad hoc additional rules. The effects of weld geometry, residual stresses and through thickness stress variation are usually implicitly included in the *S-N* curve, while the effect of other factors such as, plate thickness, environment, weld toe grinding and post-weld heat or mechanical treatments, etc. is accounted through appropriate corrections to the basic *S-N* curve.

In Eurocode EN1999-1-3 [8], bi-linear and tri-linear *S-N* curves are similarly assigned for aluminium structures, but parameters A_1 and m_1, A_2 and m_2, are different: in fact, the slope m_1 can assume the values m_1=3.4 or m_1=4.3, according to the detail under examination, while, as usual, $m_2=m_1+2$.

Other bi-linear models with two sloping lines, characterized by varying parameters A_1 and m_1, A_2 and m_2, $m_2=m_1+2$, are also commonly used for steel reinforcements or for pre-stressing steel details, which does not exhibit fatigue limit [9], [10].

In order to deal with variable amplitude loading, fatigue damage is quantified in terms of Palmgren-Miner's damage summation [11], [12]. Thus, all stress cycles cause a proportional fatigue damage, which is linearly additive.

The total damage D due to a stress spectra composed by k blocks of n_i cycles at stress range $\Delta\sigma_i$, i=1, 2, .., k, is given by:

$$D = \sum_{i=1}^{k} \frac{n_i}{N_i} \tag{3}$$

where N_i is the number of cycles to failure at stress amplitude $\Delta\sigma_i$.

In Equation (3) the material constants can be made explicit, so that it becomes

$$D = \frac{1}{A_1} \sum_{\Delta\sigma_i \geq \Delta\sigma_D} n_i \, \Delta\sigma_i^{m_1} + \frac{1}{A_2} \sum_{\Delta\sigma_D > \Delta\sigma_i > \Delta\sigma_L} n_i \, \Delta\sigma_i^{m_2} \tag{4}$$

Table 1 Fatigue classification of orthotropic steel deck bridge details (EN1993-2)

Detail category	Constructional detail	Description	Requirements
80 71	t≤12mm t>12mm	1) Continuous longitudinal stringer, with additional cutout in cross girder.	1) Assessment based on the direct stress range Δσ in the longitudinal stringer.
80 71	t≤12mm t>12mm	2) Continuous longitudinal stringer, no additional cutout in cross girder.	2) Assessment based on the direct stress range Δσ in the stringer.
36		3) Separate longitudinal stringer each side of the cross girder.	3) Assessment based on the direct stress range Δσ in the stringer.
71		4) Joint in rib, full penetration butt weld with steel backing plate.	4) Assessment based on the direct stress range Δσ in the stringer.
112 90 80	As detail 1, 2, 4 in Table 8.3 As detail 5, 7 in Table 8.3 As detail 9, 11 in Table 8.3	5) Full penetration butt weld in rib, welded from both sides, without backing plate.	5) Assessment based on the direct stress range Δσ in the stringer. Tack welds inside the shape of butt welds.
71		6) Critical section in web of cross girder due to cut outs.	6) Assessment based on stress range in critical section taking account of Vierendeel effects. **NOTE** In case the stress range is determined according to EN 1993-2, 9.4.2.2(3), detail category 112 may be used.
71	$\Delta\sigma = \frac{\Delta M_w}{W_w}$	<u>Weld connecting deck plate to trapezoidal or V-section rib</u> 7) Partial penetration weld with a ≥ t	7) Assessment based on direct stress range from bending in the plate.
50		8) Fillet weld or partial penetration welds out of the range of detail 7)	8) Assessment based on direct stress range from bending in the plate.

Finally, for an ergodic stress process, the scatter may be neglected and the damage can be expressed, for example, by

$$D = \frac{1}{A_1} \sum_{\Delta\sigma_i \geq \Delta\sigma_D} n_i \, \Delta\sigma_i^{m_1} + \frac{1}{A_2} \sum_{\Delta\sigma_D > \Delta\sigma_i > \Delta\sigma_L} n_i \, \Delta\sigma_i^{m_2} = \frac{1}{A_1} E\left[n(t)\right] E\left[\Delta\sigma^m\right] = \frac{1}{A_1} \psi_L(t) \tag{5}$$

where $\psi_L(t)$ may be referred to as the "fatigue loading".

In deterministic applications, Miner's rule stipulates that failure occurs when D is equal to unity but in the Probabilistic Model Code the critical value is also treated as a random variable.

2.2 Fracture Mechanics Approach

An alternative approach is based on the crack propagation laws determined using the fracture mechanics theory.

The general form of the crack propagation law under constant amplitude stress histories is

$$\frac{da}{dn} = f(\sigma, a, C) \tag{6}$$

where a is the crack length, n the number of cycles, da/dn is the crack propagation rate, σ is a representative value of the stress level and C a material constant. An example of a-N curves for increasing values of σ, $\sigma_{i+1} > \sigma_i$, is sketched in Figure 4: when $\sigma < \sigma_1$ the crack propagation rate is nil, while the critical crack length a_i corresponding to unstable propagation reduces as σ increases.

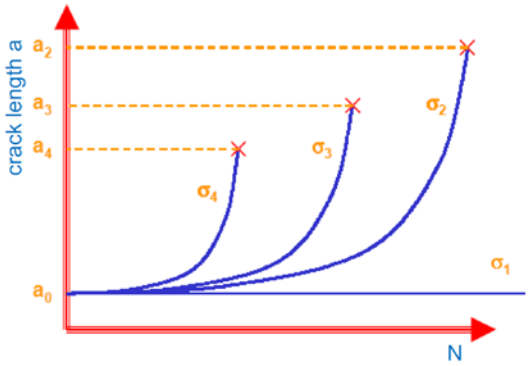


Figure 4 *a-N* curves for increasing values of σ

Generally, the crack propagation rate da/dn is expressed as function of the stress intensity factor range ΔK, which is the variation of the stress intensity factor in the stress cycle.

Recalling that the stress intensity factors are K_n for the three basic cases of stress at the crack tip are expressed by

$$K_j = Y_j \, \sigma_a \sqrt{\pi \, a} \tag{7}$$

where *j=I, II, III* is the basic case number (*I – opening*, *II – in-plane shear*, III – out-of-plane shear, see Figure 5), Y_j is the "geometry factor" and σ_a is the applied stress, derived from a detailed description of the stress system in the vicinity of the crack (but not influenced by it), ΔK_j for the basic cases can be expressed by

$$\Delta K_j = Y_j \, \Delta\sigma_a \sqrt{\pi \, a} \tag{8}$$

so that, it results

$$\Delta K = Y \, \Delta\sigma_a \sqrt{\pi \, a} \tag{9}$$

where ΔK depends on the applied loads, on the crack geometry and on the geometry of the element.

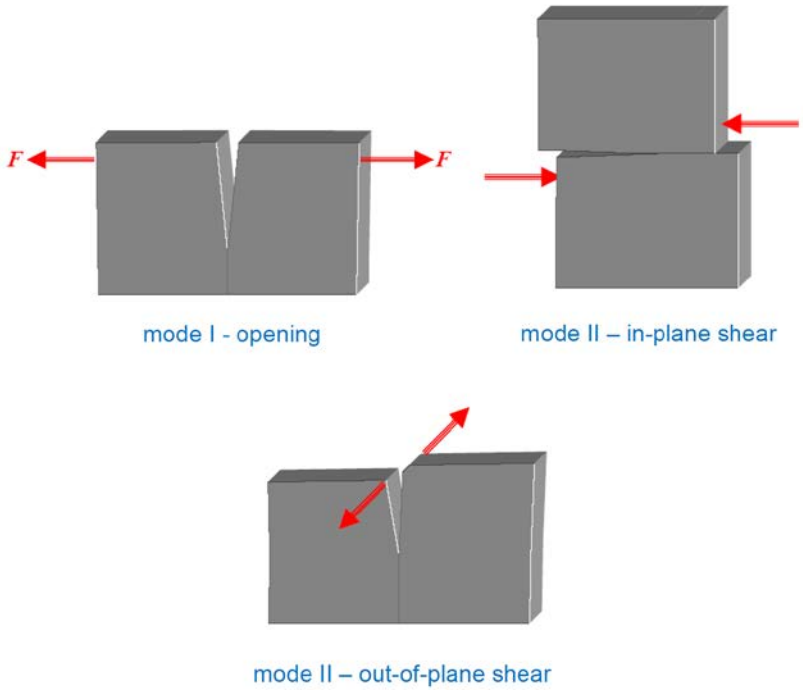


Figure 5 Basic modes of crack propagation

The most common model adopts the *Paris law* [13], which can be written as

$$\frac{da}{dn} = C\left(\Delta K^m\right) \tag{10}$$

being C and m material constants, which can be determined from experiments on simple specimens. The model can be improved by the introduction of a threshold value for ΔK, and, as in the *S-N* case, by considering different C and m parameters corresponding to particular ΔK ranges.

When ΔK can be assumed to be independent on t in a sinusoidal constant amplitude stress history and the critical length of the crack is much higher than the initial crack opening, the Expression (1) can be deduced from the Paris law (10).

Expression of ΔK can be determined by finite element analysis, directly or from the path-independent Rice integral, J [14].

For semi-elliptical cracks in plated structures, which are of interest in many practical applications, two crack dimensions, the depth a and the half-length at the surface c, both of which under fatigue are functions of the loading time t, need to be considered.

For both principal directions of the crack growth a Paris type expression is assumed, which, after integrating the differential equation from an initial defect size of a_0 to a crack size $a(t)$ after time t corresponding to an average number of stress cycles $E(n(t))$, leads to

$$\int_{a_0}^{a(t)} \frac{\sqrt{Q^m}\, dx}{C\, G(x) U_a^m\, Y_m^a \sqrt{\pi\, x^m}} = \sum \Delta\sigma^m = E[n(t)]\, E[\Delta\sigma^m] = \psi_L(t) \tag{11}$$

where $G(x)$ is the threshold correction factor, U_a is the effective stress intensity factor ratio, Q is the elliptic shape factor, and $\psi_L(t)$ (similar to the S-N approach) may be referred to as the "fatigue loading". When the fatigue life of a joint is required, the upper bound of the integral domain $a(t)$ in (11) may be taken as the critical value a_c or simply the plate thickness. Failure due to interaction between brittle fracture and section yielding is catered for through a failure assessment diagram.

3 EXAMPLE

Consider a bridge structure where various nodes are inspected for fatigue cracks. Failure in those cases will happen if the inspection results are considered as satisfactory but the failure event (nevertheless) occurs, assuming that some adequate action is taken if the inspection is not satisfactory. Let the fatigue crack for some selected node grow as indicated in Figure 6.

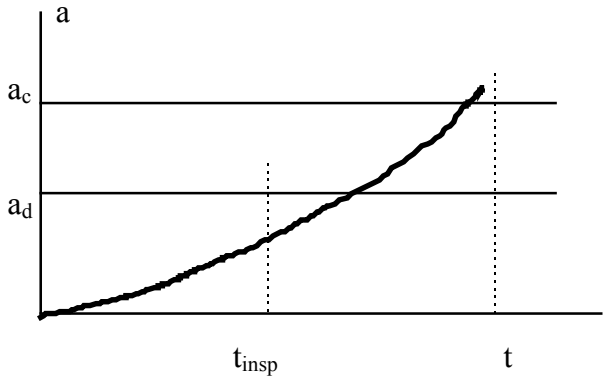


Figure 6 Fatigue failure before time t occurs if at inspection the crack length is smaller than a_d and at time t the crack length is larger than a_{crit}

Fatigue failure will occur as soon as the crack $a(t)$ reaches a random critical length a_{crit} so the failure probability for a period t can be written as:

$$P_F(t) = P\{M_f < 0\} = P\{a_{crit} - a(t) < 0\} \tag{12}$$

Note that a_{crit} is considered as time-independent; if a_{crit} is considered as time dependent (as it is in reality) this equation becomes more complex. Note also that $a(t)$ is an increasing stochastic process as cracks do not get smaller.

Let the reliability be considered as inadequate: that is, $P_F(t)>P_{tar}$, when P_{tar} is some target probability for fatigue failure. For this reason an inspection is planned at some point t_{insp} during the life time. Let the decision rule be that the structure will be repaired if a crack is detected, that is if a crack $a(t_{insp})$ larger than random detection limit a_d is detected.. The probability of failure, given a positive inspection can be written at time t_{insp}. as:

$$P_F(t) = P\{a(t) > a_{crit} \mid -a_m(t_{insp}) < a_m\} \tag{13}$$

If also cost values are attached to inspection and failure, the optimal time of inspection and repair level a_d can be found and decisions regarding the future use of the bridge can be taken as indicated in Figure 7.

Figure 7 Traffic constraints to avoid fatigue failure of metallic bridges

4 CONCLUSIONS

Fatigue assessment and evaluation of residual fatigue life of existing infrastructure have been summarized,

To design repair interventions of fatigue damaged details and planning of future maintenance and inspection programmes it must be pointed out that, since repair of fatigue cracked or damaged joint commonly restores only a part, and often a very low quota, of the original fatigue life, the definition of appropriate repair techniques should be one of the main outcomes of the fatigue assessment. In principle, successful fatigue crack repair techniques should combine satisfactory fatigue performance with simplicity, velocity and easiness of execution, aiming to achieve significant residual fatigue lives of repaired details minimising the total cost of the rehabilitation, which is the sum of the direct costs of repairs and of the indirect costs, due to the out of service or the service restrictions of the structure during the executions of the repairs themselves.

Clearly, cost reduction and repair effectiveness needs claim to lengthen profitably the time interval between two subsequent interventions, so that repairs should concern not only fatigue cracked joints, but also fatigue damaged ones, even if not yet cracked. Moreover, fatigue behaviour of repaired joints is often not known, so that it must be deduced by testing real scale specimens and compared with that of the corresponding virgin (undamaged) joints, also in the light of available literature, so that the fatigue behaviour of repaired details can be suitably identified.

REFERENCES

[1] ASTME 1150-87, Standard definitions of terms relating to fatigue. 1995 Annual Book of ASTM Standards. Vol. 03.01. ASTM. 753-762 (1995).

[2] ASM Handbook, *Fatigue and Fracture. Vol. 19*. Material, Parks: ASM (1996).

[3] Croce P. (2002) Background to fatigue load models for Eurocode 1: Part 2 Traffic loads, *Progress in Structural Engineering and Materials*, 3-2001, pp. 333-345.

[4] Manson S. S. (1953) Behavior of materials under conditions of thermal stress, NACA, TN-2933.

[5] Coffin L. F. (1954) A study of the effects of cyclic thermal stresses on a ductile metal. *Trans. ASME*, Vol. 76, 931-950.

[6] EN1993-1-9, Eurocode 3: Design of steel structures - Part 1-9: Fatigue, CEN, Brussels (2006).

[7] EN1993-2, Eurocode 3 - Design of steel structures - Part 2: Steel Bridges, CEN, Brussels, (2006).

[8] EN1999-1-3, Eurocode 9 - Design of aluminium structures - Part 1-3: Structures susceptible to fatigue, CEN, Brussels (2007).

[9] EN1992-1-1, Eurocode 2: Design of concrete structures - Part 1-1: General rules and rules for buildings, CEN, Brussels (2004).

[10] EN1992-2, Eurocode 2 - Design of concrete structures - Concrete bridges - Design and detailing rules, CEN, Brussels (2005).

[11] Palmgren, A (1924) Die Lebensdauer von Kugellagern, *Z. Vereins Deutscher Ingenieure*; Vol. 68: pp. 339-341.

[12] Miner M. A. (1945) Cumulative damage in fatigue. *Journal of Applied Mechanics*, Vol. 12: pp. A159-A164.

[13] Paris P. C., Erdogan F. (1963) A critical analysis of crack propagation laws. *Journal of Basic Engineering*, Vol. 85D, pp. 528-534.

[14] Rice J. R. (1968) A path independent integral and the approximate analysis of strain concentration by notches and cracks. *Journal of Applied Mechanics*, Vol. 35, pp. 379–386.

CHAPTER 5: PROBABILISTIC RELIABILITY THEORY

Milan Holický[1]

[1]Klokner Institute, Czech Technical University in Prague, Czech Republic

Summary

The basic concepts and procedures of the probabilistic theory of reliability accepted as the basic principles for the development of the partial factor method in the new ISO and CEN documents can be effectively used also for assessing aging infrastructures. Fundamental cases of one and two random variables are supplemented by general methods of probability integration for multivariate cases including FORM, SORM and simulation techniques. Engineering examples are provided to demonstrate practical used of general procedures.

1 INTRODUCTION

1.1 Background Documents

Principles of reliability theory and risk assessment and the common tools applied for analyzing the risk of civil engineering systems described in this text are based on the common concepts presented in the International Standard EN and ISO and other documents [1] to [4] and numerous other publications. General fundamentals including terminology are provided in the JCSS document [4] "Probabilistic Model Code". Elementary treatment of the probabilistic reliability theory and statistical techniques including Bayesian updating of probability distributions can be found in [5, 6]. In addition software products [7] and [8] offers operational tools for probability calculation and processing supplemented by valuable information that can be utilized in practical applications.

1.2 Basic Concepts

The fundamental task of the theory of structural reliability is the analysis of a simple requirement that the action effect E (expressed in a suitable unit) is smaller than the structural resistance R. This requirement accepted as a basis of the European and International Standards [1, 2, 3, 4] can be written in the form of the inequality

$$E < R \qquad (1)$$

The Condition (1) describes a desirable (satisfactory, safe) state of a considered structural component. It is assumed that structural failure occurs when the Condition (1) is not satisfied. Thus, an assumed sharp (unambiguous) distinction between a desirable (safe) and undesirable (failure) state of the structure is given by the equality

$$R - E = 0 \qquad (2)$$

Equation (2) is the fundamental form of the so-called limit state (performance) function [5]. It should be noted, however, that

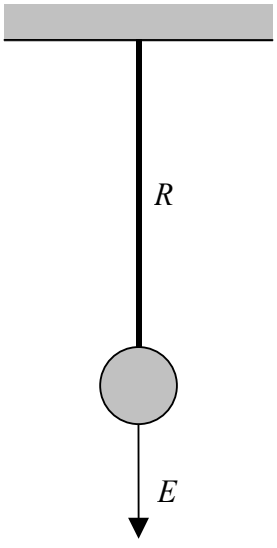


Figure 1 A tie rod

the assumption of a sharp boundary between desirable and undesirable states is a simplification that might not be suitable for all structural members and materials as indicated in the following Example 1.

Example 1

A steel tie rod from Figure 1 has a resistance in axial tension given as $R = \pi d^2 f_y / 4$, where d denotes the diameter of the rod and f_y the yield point. The rod is loaded by a weight $E = V\rho$, where V denotes the volume and ρ the density of the suspended mass. Thus the Condition (1) has the form [5]

$$V\rho < \pi d^2 f_y / 4$$

The limit state function follows from the above inequality as

$$\pi d^2 f_y / 4 - V\rho = 0$$

In this example, the limit state is defined as the state when the stress in the rod reaches the yield point f_y. This simplification is accepted in many common cases, but it may not correspond to the actual failure of the rod, in particular when structural steel with significant ductility and strain hardening is used.

1.3 Random variables

Both the variables E and R are generally random variables and the validity of Inequality (1) cannot be guaranteed absolutely, i.e. with the probability equal to 1. Therefore, it is necessary to accept the fact that the limit state described by Equation (2) may be exceeded and failure may occur with a certain probability. The essential objective of reliability theory is to assess the probability of failure p_f and to find the necessary conditions for its limited magnitude. For the simple condition in the form of Inequality (1), the probability of failure may be formally written as

$$p_f = \mathrm{P}(E > R) \tag{3}$$

The random character of the action effect E and the resistance R, both expressed in terms of a suitable variable (performance indicator) X (i.e. stress, force, bending moment, deflection) is usually described by an appropriate distribution function, i.e. by distribution functions $\Phi_E(x)$, $\Phi_R(x)$ and by corresponding probability density functions $\varphi_E(x)$, $\varphi_R(x)$, where x denotes a general point of the considered variable X used to express both of the variables E and R. Distributions of the variables E and R further depend on appropriate parameters, for example on moment parameters μ_E, σ_E, α_E, μ_R, σ_R and α_R. Let us further assume that E and R are mutually independent (which may be achieved by an appropriate transformation).

Figure 2 shows an example of probability density functions of both the variables E and R and their mutual location. Types of distribution and their parameters indicated in Figure 2 are just indicative information. In particular, the moment parameters (the means and standard deviations) may be considered as relative values related to the resistance mean μ_R (i.e. normalised by μ_R).

Note that the probability density functions $\varphi_E(x)$ and $\varphi_R(x)$ shown in Figure 2 overlap each other and, therefore, it is clear that unfavourable realisations of the variables E and R, denoted by small letters e and r, may occur in such a way that $e > r$, i.e. the load effect is greater than the resistance and failure may occur. Obviously in order to keep the failure probability $p_f = \mathrm{P}(E>R)$ within acceptable limits, the parameters of variables E and R must satisfy certain conditions (concerning the mutual location and variances of both the distributions) depending on their types of distribution.

The desired conditions will certainly include the trivial inequality $\mu_E < \mu_R$ (see Figure 2). Obviously, this "requirement for mutual location" of both the distributions is not sufficient to ensure the specified failure probability p_f. The correct conditions should obviously also include requirements for variances of both variables. This will be clarified in the following discussion of fundamental cases of structural reliability.

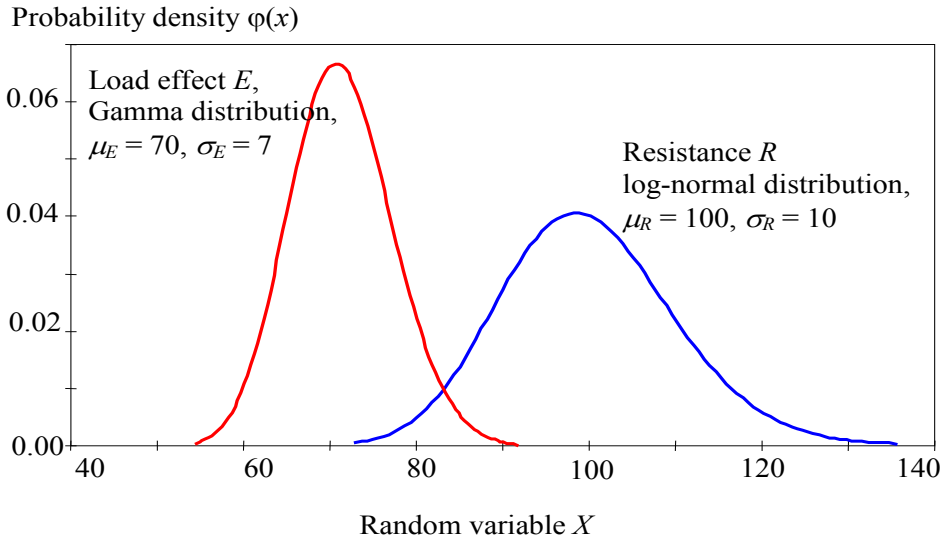


Figure 2 Action effect *E* and resistance *R* as random variables

2 FUNDAMENTAL CASES OF ONE RANDOM VARIABLE

Consider first a special case when one of the variables E and R, say the action effect E, has a very low (negligible) variability comparing to the variability of the resistance R. Then E may be considered a non-random (deterministic) variable, i.e. a variable that always attains a certain fixed value e_0 ($E = e_0$) as indicated in Figure 3. In some cases this assumption may certainly be considered a reasonable approximation. One of these cases may be the loaded tie rod from Example 1, where the weight F of the suspended mass can be determined with sufficient accuracy (i.e. without any significant uncertainty).

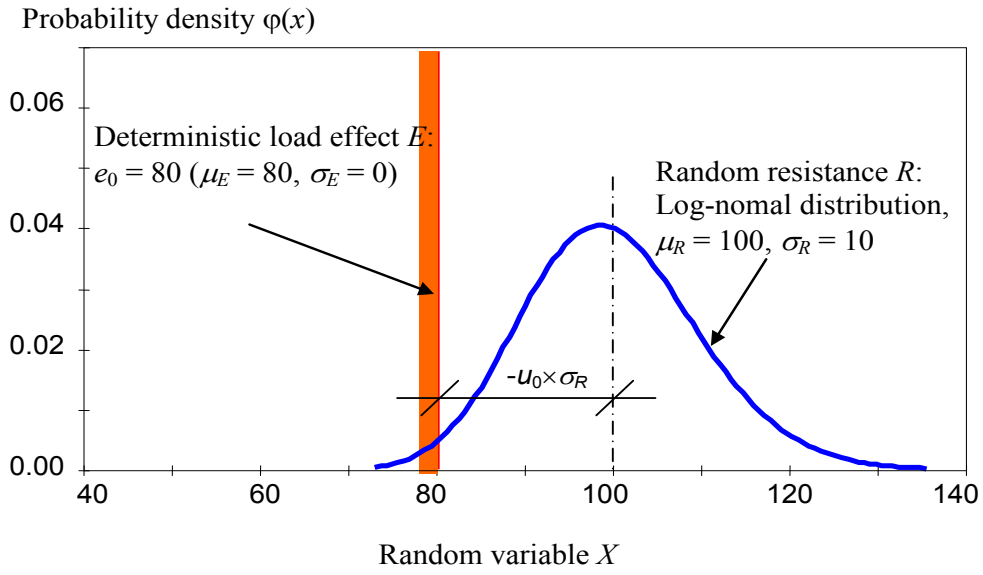


Figure 3 The deterministic effect of actions *E* and random resistance *R*

The special case is illustrated in Figure 3, where the action effect is indicated by a fixed value $e_0 = 80$ ($\mu_E = 80$, $\sigma_E = 0$) and the resistance by the log-normal distribution having the mean $\mu_R = 80$, $\sigma_R = 10$ (all numerical values being normalised dimensionless quantities).

The probability of failure p_f for the special case of the deterministic load effect of actions shown in Figure 3 may be assessed directly from the distribution function $\Phi_R(x)$, similar to the case of a fractile. The value e_0 may be considered as simply a fractile of the resistance R for which the probability p_f may be calculated using known Equations [5].

$$p_f = P(R < e_0) = \Phi_R(e_0) \tag{4}$$

The distribution function $\Phi_R(e_0)$ is usually assessed from tables for a standardised random variable U, for which the value u_0 corresponding to e_0 is computed. It follows from the transformation formula [5]

$$u_0 = (e_0 - \mu_R) / \sigma_R \tag{5}$$

Then the probability of failure is given as

$$p_f = P(R < e_0) = \Phi_{LN,R}(e_0) = \Phi_{LN,U}(u_0) \tag{6}$$

where $\Phi_{LN,U}(u_0)$ is the value of the distribution function of a standardised random variable of the log-normal distribution.

Note that the value $-u_0$ is the distance of the fixed value e_0 of the action effect E from the mean μ_R of the resistance R expressed in the units of the standard deviation σ_R. If the distribution of the resistance R is normal (not log-normal), then the defined distance is called the reliability index β

$$\beta = (\mu_R - e_0) / \sigma_R \tag{7}$$

and the probability of failure may be expressed by the relation

$$p_f = P(R < e_0) = \Phi_U(-\beta) \tag{8}$$

If the resistance R has a different distribution from normal, then the reliability index β is formally defined as a negative value of a standardised random variable corresponding to the failure probability p_f. Thus, in general

$$\beta = -\Phi_U^{-1}(p_f) \tag{9}$$

where $-\Phi_U^{-1}(p_f)$ denotes the inverse distribution function of a standardised normal distribution. As its numerical values are more suitable than the values of the failure probability, the reliability index β defined by Equation (9) is a commonly used measure of structural reliability.

Example 2

Consider that resistance R has a mean $\mu_R = 100$ (expressed in dimensionless units), standard deviation $\sigma_R = 10$ (the coefficient of variation is thus $w = 0.10$). For the deterministic effect of actions it holds that $e_0 = 80$ (see Figure 3). If R has a normal distribution, then the reliability index follows directly from Equation (7)

$$\beta = (100 - 80) / 10 = 2$$

and the probability of failure follows from Relation (8)

$$p_f = P(R < 80) = \Phi_U(-2) = 0.023$$

where $\Phi_u(-2)$ is the value of the distribution function of the standardised normal distribution.

However, if the distribution of R is not normal but log-normal with the lower bound at zero, the skewness $\alpha = 3w + w^3 = 0.301$), then it follows from Equation (5)

$$u_0 = (80 - 100) / 10 = -2$$

The probability of failure p_f is then given as

$$p_f = P(R < 80) = \Phi_{LN,U}(-2) = 0.014$$

where $\Phi_{LN,U}(-2)$ is the distribution function of the standardised random variable U with the log-normal distribution having the lower bound at zero (the skewness $\alpha = 0.301$). The resulting probabilities do not differ significantly but this is because their values are rather high.

If the fixed value of an action's effect decreases to $e_0 = 70$, then for the normal distribution of the resistance R the reliability index is $\beta = 3$ and the probability of failure is

$$p_f = P(R < 70) = \Phi_U(-3) = 0.00135$$

If the distribution of the resistance R is the log-normal distribution with the lower bound at zero, then

$$p_f = P(R < 70) = \Phi_{LN,U}(-3) = 0.00021$$

The reliability index defined by Equation (9) is $\beta = -\Phi_U^{-1}(0.00021) = 3.53$, i.e. greater than the value 3 valid for the normal distribution of the resistance R.

Obviously, when the load effect is only $e_0 = 70$, the resulting failure probabilities are remarkably lower than in the case when $e_0 = 80$. Furthermore, the numerical example also shows that the assumption concerning the type of distribution plays an important role and may be, in some cases, decisive.

3 TWO RANDOM VARIABLES HAVING NORMAL DISTRIBUTION

Assume that both basic variables, the action effect E and resistance R are random variables. It is then more complicated to assess the probability of failure defined by Equation (3). A simple solution can be obtained assuming normal distribution for both E and R. Then, also, the difference

$$G = R - E \tag{10}$$

called the reliability margin, has normal distribution with parameters

$$\mu_G = \mu_R - \mu_E \tag{11}$$

$$\sigma_G^2 = \sigma_R^2 + \sigma_E^2 + 2\rho_{RE}\sigma_R^2\sigma_E^2 \tag{12}$$

Here ρ_{RE} is the coefficient of correlation of R and E. It is often assumed that R and E are mutually independent and $\rho_{RE} = 0$. Equation (3) for the probability of failure p_f can now be modified to

$$p_f = P(E > R) = P(G < 0) = \Phi_G(0) \tag{13}$$

and the whole problem is reduced to the determination of the distribution function $\Phi_G(0)$, which gives the probabilities of the reliability margin G being negative. The distribution function $\Phi_G(0)$ is usually determined using the transformation of the variable G to

the standardised random variable U. Using this equation, the value u_0 corresponding to the value $g = 0$ is given as

$$u_0 = (0 - \mu_G) / \sigma_G = - \mu_G / \sigma_G \tag{14}$$

The probability of failure is then given as

$$p_f = \mathrm{P}(R < E) = \Phi_G(0) = \Phi_U(u_0) \tag{15}$$

The probability density function $\Phi_G(g)$ of the reliability margin G is shown in Figure 4, where the grey area under the curve $\Phi_G(g)$ corresponds to the failure probability p_f.

Assuming that G has the normal distribution, the value $-u_0$ is called the reliability index, which is commonly denoted by the symbol β. In the case of the normal distribution of the reliability margin G, it follows from Equations (11), (12) and (14) that the reliability index β is given by a simple relationship

$$\beta = \mu_G / \sigma_G = \frac{\mu_R - \mu_E}{\sqrt{\sigma_R^2 + \sigma_E^2 + 2\rho_{RE}\sigma_R^2\sigma_E^2}} \tag{16}$$

If the quantities R and E are mutually independent, then the coefficient of correlation ρ_{RE} vanishes (ρ_{RE}=0). Thus, the reliability index β is the distance of the mean μ_G of the reliability margin G from the origin, given in the units of the standard deviation σ_G.

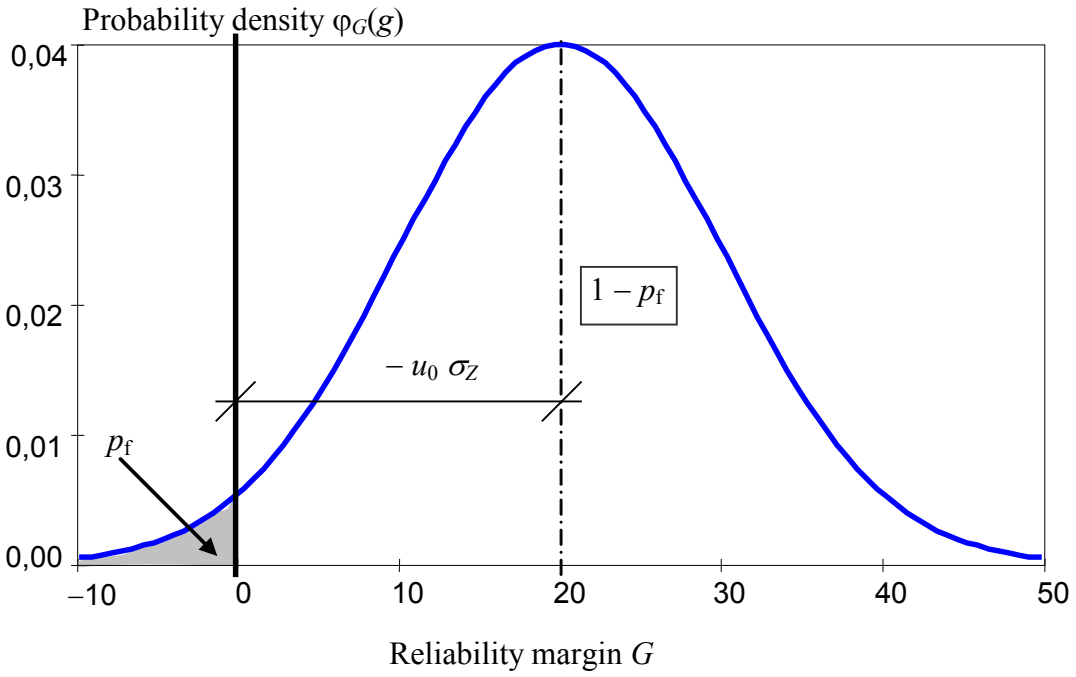


Figure 4 Distribution of the reliability margin *G*

Example 3

Consider again the Example 2, in which the resistance R and the load effect E are mutually independent random variables (ρ_{RE}=0) having normal distribution. The resistance R has the mean μ_R = 100, variance σ_R = 10 (the coefficient of variation is therefore only w = 0.10), and the effect of actions E has the mean μ_E = 80 and σ_E = 8 (all expressed in dimensionless units). It follows from Equation (11) and (12) that

$$\mu_G = 100 - 80 = 20$$

$$\sigma_G^2 = 10^2 + 8^2 = 12.81^2$$

As both basic variables R and E have the normal distribution, the reliability index β follows directly from Equation (7)

$$\beta = 20 / 12.81 = 1.56$$

and the probability of failure follows from relation (8)

$$p_f = \mathrm{P}(G < 0) = \Phi_U(-1.56) = 0.059$$

If the variables E and R are not normal, then the distribution of the reliability margin G is not normal either, and the above-described procedure has then to be modified. In a general case, the numerical integration or transformation of both variables into variables with normal distribution can be used. The transformation into normal distribution is used primarily in software products.

There is, however, an approximate simple procedure that can provide a good first assessment of the failure probability p_f. The reliability margin G may be approximated by three-parameter log-normal distribution. Assume that the distributions of E and R depend on the moment parameters μ_E, σ_E, α_E, μ_R, σ_R and α_R. The mean and the variance of the reliability margin G may be assessed from the previous Equations (11) and (12) which hold for variables with an arbitrary distribution. Assuming mutual independence of E and R, the skewness α_G of the reliability margin G may be estimated using the following approximate relation

$$\alpha_G = \frac{\sigma_R^3 \alpha_R - \sigma_E^3 \alpha_E}{\left(\sigma_R^2 + \sigma_E^2\right)^{3/2}} \tag{17}$$

It is then assumed that the reliability margin G can be described with sufficient accuracy by a log-normal distribution with determined moment parameters μ_G, σ_G and α_G (Equations (11), (12) and (17)). It shows that this approximation offers satisfactory results if the probability of failure is not too small.

Example 4

Consider a tie rod having a rigidity R and bearing a suspended load of a weight E. Let R be a log-normal variable with origin at zero having the parameters (expressed again in relative dimensionless units) μ_R = 100 and σ_R = 10 (and therefore α_R = 0.301), E has the Gumbel distribution with moment parameters μ_E = 50, σ_E = 10 and skewness α_E = 1.14 (the inherent property of the Gumbel distribution).

The moment parameters of the reliability margin are assessed according to Equations (11), (12) and (17)

$$\mu_G = \mu_R - \mu_E = 100 - 50 = 50$$

$$\sigma_G^2 = \sigma_R^2 + \sigma_E^2 = 10^2 + 10^2 = 14.14^2$$

$$\alpha_G = \frac{\sigma_R^3 \alpha_R - \sigma_E^3 \alpha_E}{\left(\sigma_R^2 + \sigma_E^2\right)^{3/2}} = \frac{10^3 \times 0.301 - 10^3 \times 1.14}{\left(10^2 + 10^2\right)^{3/2}} = -0.30$$

For a standardised random variable it follows from Equation (14) that

$$u_0 = -\mu_G / \sigma_G = -50 / 14.14 = -3.54$$

For the log-normal distribution having the skewness $\alpha_G = -0.30$ it holds that

$$p_f = \mathrm{P}(R < E) = \Phi_{\mathrm{LN},U}(-3.54) = 0.00101$$

which corresponds to a reliability index $\beta = 3.09$. A more precise result obtained by the application of the software VaP [8] is $p_f = 0.00189$.

However, when skewness is not taken into account in the assessment of probability, it follows from the standardised normal distribution that

$$p_f = P(R < E) = \Phi_U(-3.54) = 0.00020$$

which differs significantly from the result when the log-normal distribution was considered.

4 TWO RANDOM VARIABLES HAVING GENERAL DISTRIBUTION

The exact solution of the probability of failure p_f defined for the case of two random variables E and R by Equation (3) may be obtained by integration. Figure 5 is used to explain the integration procedure. Let the event A denote the occurrence of an action effect E in the differential interval $\langle x, x+dx\rangle$. The probability of the event A is given as

$$P(A) = P(x < E < x+dx) = \varphi_E(x)\, dx \tag{18}$$

Let us denote B as the event that a resistance R occurs within the interval $\langle -\infty, x\rangle$. Probability of the event B is given by the relation

$$P(B) = P(R < x) = \Phi_R(x) \tag{19}$$

The differential increment of the probability of failure dp_f corresponding to the occurrence of the variable E in the interval $\langle x, x+dx\rangle$ is given by the probability of simultaneous occurrence of the events A and B, i.e. by the probability of their intersection $A \cap B$. According to the principle of multiplication of probabilities, it holds that

$$dp_f = P(A \cap B) = P(A)\, P(B) = P(x < E < x+dx)\, P(R < x) = \varphi_E(x)\, \Phi_R(x)\, dx \tag{20}$$

The above-mentioned assumption of mutual independence of the variables E and R, and thus also of the events A and B, is applied here.

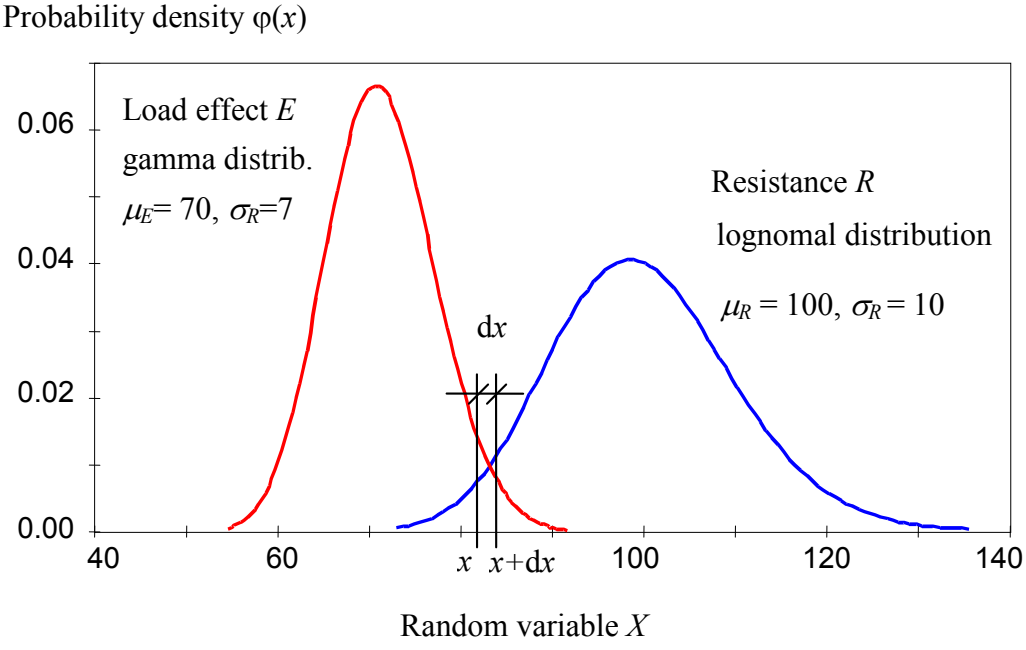


Figure 5 Distribution of variables *E* and *R*

The integration of the differential relationship (20) over the interval in which both the variables E and R occur simultaneously (generally the interval $<-\infty, \infty>$) leads to the relation

$$p_f = \int_{-\infty}^{\infty} \varphi_E(x)\,\Phi_R(x)\,dx \tag{21}$$

The integration of the relation (21) usually has to be carried out numerically or using the simulation methods of Monte Carlo.

Example 5

The action effect E and the resistance R are described by a log-normal distribution with the same parameters as in Example 4 (the Gumbel distribution for E was simply approximated by a log-normal distribution having the same parameters). The approximate solution in Example 4, based on the log-normal distribution with the lower bound at zero, leads to the probability of failure $p_f = P(R < E) = \Phi_{LN,U}(-3.54) = 0.00100$. Numerical integration, according to relation (21) using the programme MATHCAD, leads to a solution $p_f = P(R < E) = 0.00187$, the programme VaP suggests a solution $p_f = P(R < E) = 0.00189$, which can be considered a very good approximation.

Figure 6 shows the variation of the probability of failure p_f with the coefficients of skewness α_E and α_R for the same parameters of variables E and R ($\mu_R = 100$, $\sigma_R = 10$, $\mu_E = 50$ and $\sigma_E = 10$) as in Example 4 and 5. The probability of failure p_f is assessed by direct integration using the programme MATHCAD. It follows from Figure 6 that the probability of failure p_f depends greatly on the skewnesses α_E and α_R (therefore on assumed theoretical models), and in practical conditions, can differ by several orders of magnitude, even when the means and the standard deviations of the variables E and R remain the same.

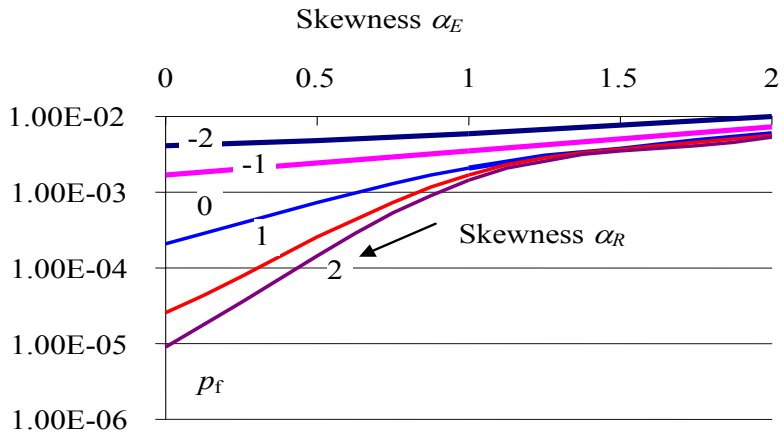


Figure 6 Variation of the probability of failure p_f with the coefficients of skewness α_E and α_R for $\mu_R = 100$, $\mu_R = 10$, $\sigma_E = 50$ and $\sigma_E = 10$.

Thus, it appears that the determination of failure probability in the case of a simple example described by Inequality (1), where only two random variables E and R are involved, is easy only when both the variables are normally distributed. If they have other distributions, the exact solution is more complicated and the resulting values depend significantly on the assumed types of distributions. The approximate solution assuming for E and R a general (three-parameter) log-normal distribution provides a good first estimate failure probability. The obtained values should be, however, verified by more exact procedures considering appropriate theoretical models of E and R.

5 DESIGN POINT IN EUROCODES

Various simplifications are accepted in order to enable practical application of important principles of the theory of reliability and their effective applications in operational documents like Eurocodes. Figure 7, taken from EN 1990 [1], illustrates basic probabilistic principles accepted for developing the partial factor method in Eurocodes. The basic variables E and R considered above are indicated in a two-dimensional graph together with the fundamental limit state function (10).

The horizontal axis shows the ratio R/σ_R, the vertical axis the ratio E/σ_E. It is assumed that the variables E and R are independent and both normally distributed. Consequently the joint probability distribution function can be represented by concentric circles corresponding to different levels of the probability density. The assumption of the normal distribution may be in some cases (see Examples 5 and 6) an unrealistic hypothesis, which might provide an approximation only. However, as described later, the actual distributions of both the basic variables E and R can be transformed at a given point into the normal distributions and, therefore, Figure 7 may be considered as indicating a specific case of variables after such a transformation.

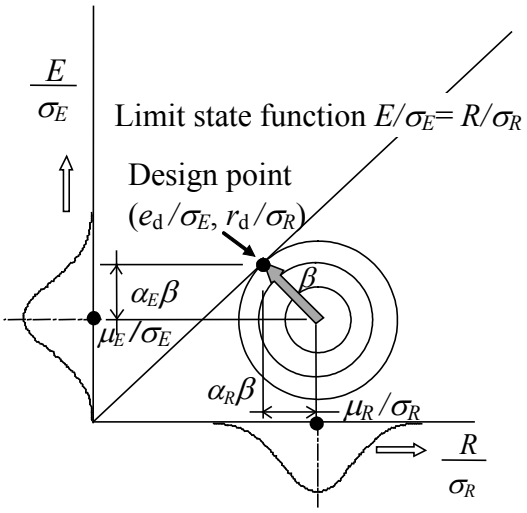


Figure 7 Design point

The safe (desirable) region, where the Condition (1) is satisfied, is located under the failure boundary (under the diagonal of the axes), the failure (undesirable) region lies above the diagonal. The design point (e_d, r_d) can be any point lying on the failure boundary (the diagonal). However, it has been shown that the best option is the point on the limit state function closest to the mean (μ_E, μ_R). Then important properties (like consistency and invariance of the solution under different formulations of the limit state function and choice of basic variables) are assured. Accepting this finding, the design point coordinates follow from Figure 7 in the form

$$e_d = \mu_E - \alpha_E \beta \sigma_E \tag{22}$$

$$r_d = \mu_R - \alpha_R \beta \sigma_R \tag{23}$$

where α_E and α_R denote here the so-called FORM sensitivity factors of the variables E and R, and not the skewness as in the previous sections (such unpleasant ambiguity is accepted in order to use the same notation as the documents CEN [1] and ISO [2,3]). The "minus" signs in Equations (22) and (23) are used consistently with the convention provided for the sensitivity factors α_E and α_R in Eurocode [1]).

It follows from Figure 7 for the weight factors α_E and α_R (direction cosines of the normal to the failure boundary), considering the convention in Equations (22) and (23)

$$\alpha_E = -\sigma_E / \sqrt{\sigma_E^2 + \sigma_R^2} \tag{24}$$

$$\alpha_R = \sigma_R / \sqrt{\sigma_E^2 + \sigma_R^2} \tag{25}$$

In Eurocodes an approximation of these sensitivity factors by fixed values is further accepted

$$\alpha_E = -\sigma_E / \sqrt{\sigma_E^2 + \sigma_R^2} = -0.7 \tag{26}$$

$$\alpha_R = \sigma_R / \sqrt{\sigma_E^2 + \sigma_R^2} = 0.8 \tag{27}$$

while the validity of such an approximation is delimited in EN 1990 [1] by means of a condition for the ratio of the standard deviations in the form

$$0.16 < \sigma_E / \sigma_R < 7.6 \tag{28}$$

When this condition is not satisfied, it is recommended that the weight factor $\alpha = \pm 1.0$ is used for a variable having the greater standard deviation. Let us remark that this simplification is on the safe side as the sum of squared direction cosines should be equal to 1.

The design values e_d and r_d of the variables E and R are thus defined as the fractiles of the normal distribution

$$\mathrm{P}(E > e_d) = \Phi_U(+\alpha_E \beta) = \Phi_U(-0.7\beta) \tag{29}$$

$$\mathrm{P}(R < r_d) = \Phi_U(-\alpha_R \beta) = \Phi_U(-0.8\beta) \tag{30}$$

where $\Phi_U(u)$ denotes a standardised distribution function of normal distribution. If $\beta = 3.8$, then the design values e_d and r_d are fractiles corresponding approximately to probabilities 0.999 and 0.001. Note that in Equation (29) the use of the symmetry of normal distribution is taken into account, i.e. of the relationship $1 - \Phi_U(-\alpha_E \beta) = \Phi_u(+\alpha_E \beta)$.

If the load or resistance models contain several basic variables (other loads, several materials, geometrical data), Equations (29) and (30) are used for the dominating variables (the most significant for the considered structural member) only. For other (non-dominating or accompanying) variables the requirements on the design values are decreased by reducing the reliability index β using factor 0.4, thus

$$\mathrm{P}(E > e_d) = \Phi_U(+0.4\alpha_E \beta) = \Phi_U(-0.28\beta) \tag{31}$$

$$\mathrm{P}(R < r_d) = \Phi_U(-0.4\alpha_R \beta) = \Phi_U(-0.32\beta) \tag{32}$$

If the reliability index $\beta = 3.8$, the design values of non-dominant variables are fractiles corresponding approximately to the probabilities 0.9 and 0.1.

Thus, the design values e_d and r_d are the upper fractiles (for actions) or the lower fractiles (for resistance), corresponding to certain probabilities of being exceeded (actions) or not reached (resistance). For the dominant variables, the probabilities are given by the

distribution function of the normal standardised distribution for values $u = +\alpha_E\beta$ and $-\alpha_R\beta$, in the case of non-dominant variables for reduced values $u = +0.4\alpha_E\beta$ and $-0.4\alpha_R\beta$. These probabilities (for the lower fractile approximately 0.001 for dominant and 0.1 for non-dominant variables) are then used to determine the design values of the basic variables having an arbitrary type of (non-normal) distribution. Note that in the case of upper fractiles (actions) the complementary probabilities (0.999 and 0.9) and appropriate probability distributions are to be considered.

Example 6

The design values e_d and r_d of the variables E and R from Example 4 will be assessed assuming that the reliability index $\beta = 3.8$, $\alpha_E = -0.7$ and $\alpha_R = 0.8$. According to Equation (29), it holds for E that

$$P(E > e_d) = \Phi_U(\alpha_E\beta) = \Phi_U(-2.66) = 0.0039$$

The complementary probability is therefore 0.9961 and we obtain from equation

$$e_d = \mu - (0.45 + 0.78\ln(-\ln(p)))\sigma = 50 - (0.45 + 0.78\times\ln(-\ln(0.9961)))\times 10 = 88.75$$

We remark that when the normal distribution is assumed, we obtain from equation

$$e_d = \mu + u_p\sigma = 50 + 2.66 \times 10 = 76.6$$

According to Equation (5.30), it holds for R

$$P(R < r_d) = \Phi_U(-\alpha_R\beta) = \Phi_U(-3.04) = 0.0012$$

For the log-normal distribution with the mean of 100 (units) and the standard deviation of 10 (units) it follows from equation

$$r_d \cong \mu \exp(u_{\text{norm},p} \times w) = 100 \times \exp(-3.04\times 0.10) = 73.79$$

For the normal distribution we obtain

$$r_d = \mu + u_p\sigma = 50 - 3.04 \times 10 = 69.6$$

Obviously, it holds for the design point that $e_d > r_d$ and the tie rod does not satisfy the Condition (1) (we know from Example 4 that β is only 3.09). In order to satisfy the condition for a reliability index of 3.8, the parameters of the variables E and R would have to be modified.

6 MULTIVARIATE CASE

In previous sections the basic case of two random variables and a linear performance function have been considered. As a rule more basic variables $X_1, X_2, \ldots X_n$ have to be considered. Variables $X_1, X_2, \ldots X_n$ are denoted as the vector $\boldsymbol{X}$ $[X_1, X_2, \ldots X_n]$ and their realisations $x_1, x_2, \ldots, x_n$ as the vector $\boldsymbol{x}$ $[x_1, x_2, \ldots, x_n]$. In the multivariate case the reliability margin (10) may be generalised as

$$G(X_1, X_2, \ldots X_n) = G(\boldsymbol{X}) \tag{33}$$

The safe domain of the basic variables is described by the inequality

$$G(X_1, X_2, \ldots X_n) = G(\boldsymbol{X}) > 0 \tag{34}$$

The unsafe domain of the basic variables is described by the inequality

$$G(X_1, X_2, \ldots X_n) = G(\boldsymbol{X}) < 0 \tag{35}$$

The limit state function is thus given as

$$G(X_1, X_2, \ldots X_n) = G(\boldsymbol{X}) = 0 \tag{36}$$

When a non-linear performance function $G(\boldsymbol{X})$ and more basic variables $\boldsymbol{X}$ ($\boldsymbol{X}$ is a vector of basic variables) are considered, failure probabilities p_f can be generally expressed using the limit state function $G(\boldsymbol{X})$ as

$$P_f = \mathrm{P}(G(\boldsymbol{X}) \le 0) = \int_{G(\boldsymbol{X}) \le 0} \varphi(\boldsymbol{X}) \mathrm{d}\boldsymbol{X} \tag{37}$$

$\varphi(\boldsymbol{X})$ is the joint probability density function of the vector of all the basic variables $\boldsymbol{X}$ and the inequality $G(X) \le 0$ denotes the failure domain (the equality $G(X) = 0$ denotes the limit state and $G(X) \ge 0$ denotes the safe domain). However, such a function may be difficult to find or may be very complicated. The integral in Equation (37) can also be written as multiple integral

$$P_f = \mathrm{P}(G(\boldsymbol{X}) \le 0) = \int_{G(\mathbf{X}) \le 0} \varphi_{X1}(x_1)\varphi_{X2}(x_2)\ldots\varphi_{Xn}(x_n)\,\mathrm{d}x_1\,\mathrm{d}x_2\ldots\mathrm{d}x_n \tag{38}$$

The integral in Equation (37) or (38) indicates how the probability p_f can be determined provided that the joint probability density function $\varphi(\boldsymbol{X})$ and densities $\varphi_{Xi}(x_i)$ are known. In some special cases the integration indicated in Equations (37) and (38) can be done analytically, in some other cases, when the number of basic variables is small (up to 5), various types of numerical integration may be effectively applied.

In general (see ISO 2394 [2]), the failure probability p_f may be computed using:

- Exact analytical integration
- Numerical integration methods
- Approximate analytical methods (FORM, SORM, methods of moments)
- Simulation methods
- A combination of these methods

Exact analytical methods can be applied only in exceptional academic cases. Numerical integration can be applied much more frequently. The most popular computational procedures to determine the failure probability constitute approximate analytical methods. In complicated cases simulation methods or their combination with approximate analytical methods are commonly applied. Most of the commercially available software products include approximate analytical methods and various types of simulation methods.

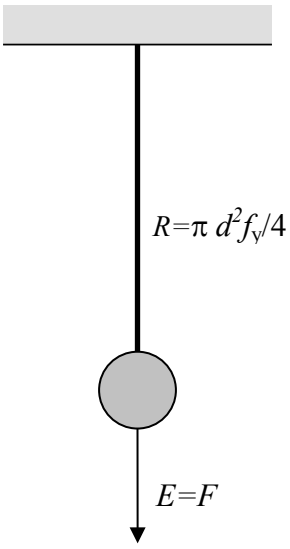


Figure 8 A tie rod

Example 7

To illustrate the general concepts described above consider again the tie rod described in Example 1 (see Figure 8). The resistance of the rod is given by a non-linear $R = \pi\, d^2 f_y/4$, where d denotes the diameter of the rod (in general a random variable), f_y is the yield point (a random variable of considerable deviations). Assume that the rod is carrying a deterministic mass $E = F$. The reliability margin (33) has therefore the form

$$G(\boldsymbol{X}) = \mathrm{g}(d, f_y, F) = \pi\, d^2 f_y/4 - F > 0$$

The limit state function is given as

$$G(\boldsymbol{X}) = g(d, f_y, F) = \pi\, d^2 f_y/4 - F = 0$$

In addition to constants there are three basic variables entering the limit state function: d, f_y and F. Note that the limit state is defined as reaching the yield point f_y, which is a commonly accepted simplification that may not be adequate for some types of construction steel.

The limit state function may be difficult to show graphically in the case of more than two basic variables. For given forces F = 100 and 50 kN the limit state function is shown as $G(\boldsymbol{X}) = 0$ in Figure 9, where $G(\boldsymbol{X}) > 0$ corresponds to the safe domain and $G(\boldsymbol{X}) < 0$ the unsafe domain. The limit state function is a non-linear smooth curve. Figure 9 shows also the means of d and f_y (30 mm and 290 MPa) and the design points, which are derived using the FORM method (described below) and assuming normal distribution of d and f_y having standard deviations 3 mm and 25 MPa respectively.

In accordance with Equations (37) or (38) the failure probabilities p_f can be determined by integration over the unsafe domain, which is located below the curve representing the limit state function (see Figure 9).

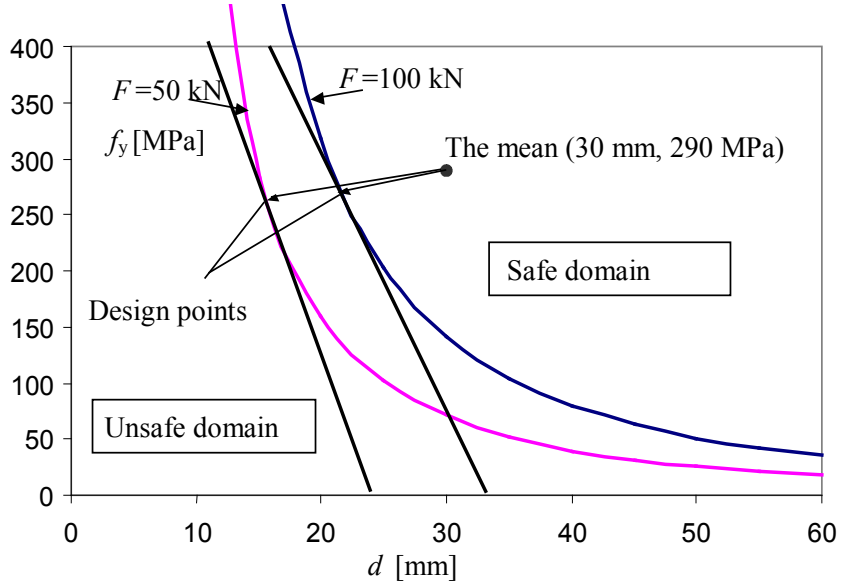


Figure 9 The limit state function and design points for the tie rod

7 FORM AND SORM

The FORM (First Order Reliability Method) is one of the basic and very efficient reliability methods. The FORM method is used as a fundamental procedure by a number of software products for the reliability analysis of structures and systems. It is also mentioned in EN 1990 [1] that the design values are based on the FORM reliability method. Figure7 illustrates this concept for two variables E and R. Considering a multivariate case when basic variables are described by a vector $\boldsymbol{X}$ [X_1, X_2, ... X_n], the main steps of the FORM method can be summarised as follows:

– The basic variables $\boldsymbol{X}$ are transformed into a space of standardised normal variables $\boldsymbol{U}$, and the performance function $G(\boldsymbol{X}) = 0$ transformed into $G'(\boldsymbol{U}) = 0$ (Figure 10).

– The failure surface $G'(U) = 0$ is approximated at a chosen given point by a tangent hyperplane (using Taylor expansion).
– The design point, i.e. the point on the surface $G'(U) = 0$ closest to the origin, is found by iteration (see Figure 10).
- The reliability index β is determined as the distance of the design point from the origin (see Figure 10) and then the failure probability P_f is given as $P_f = \Phi(-\beta)$.

The FORM method can be refined by approximating the failure surface $G'(U) = 0$ by a quadratic surface. Such a method is called The Second Order Reliability Method (SORM). In literature on structural reliability a number of other improvements and additional modifications may be found [3].

The first step, transformation of the original variable $\boldsymbol{X}$ into a space of standardised normal variables $\boldsymbol{U}$, is illustrated in Figure 10 (a) showing the original basic variables R and E and Figure 10 (b) showing the transformed variables U_R and U_E. The transformation into the equivalent normal variables at a given point $\boldsymbol{x}^*$ is based on two conditions:

– Equal distribution functions:

$$\Phi_X(x^*) = \Phi_U\left(\frac{x^* - \mu_X^e}{\sigma_X^e}\right) \tag{39}$$

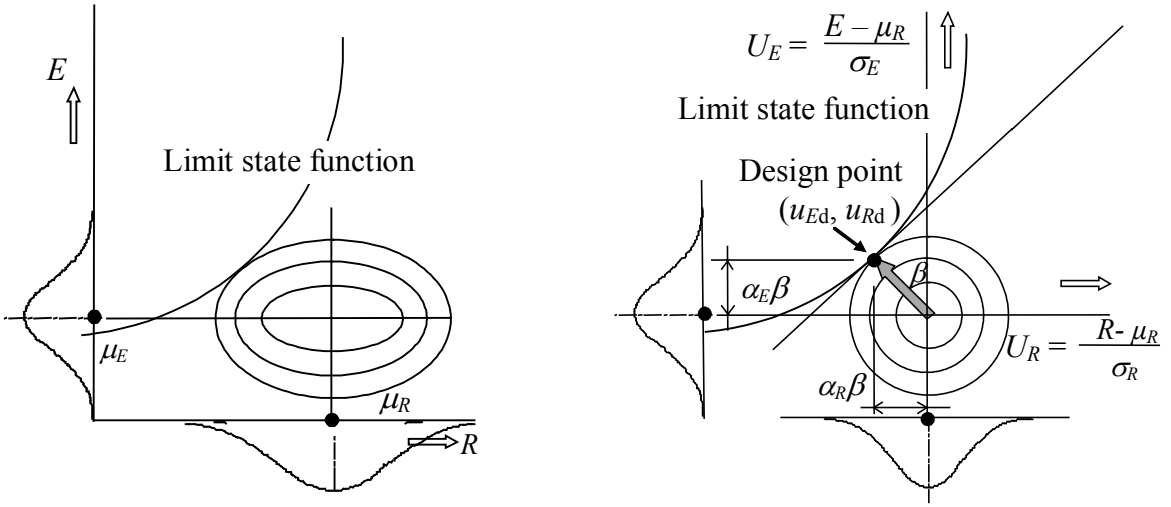
– Equal probability increments:

$$\varphi_X(x^*) = \frac{1}{\sigma_X^e}\varphi_U\left(\frac{x^* - \mu_X^e}{\sigma_X^e}\right) \tag{40}$$

The mean and the standard deviation of the equivalent normal distribution follow from Equations (39) and (40)

$$\mu_X^e = x^* - \sigma_X^e\left[\Phi_U^{-1}(\Phi_X(x^*)\right] \tag{41}$$

$$\sigma_X^e = \frac{1}{\varphi_X(x^*)}\varphi_U\left[\frac{x^* - \mu_X^e)}{\sigma_X^e}\right] = \frac{1}{\varphi_X(x^*)}\varphi_U\left[\Phi_U^{-1}(\Phi_X(x^*))\right] \tag{42}$$



(a) Original basic variables R and E. (b) Transformed variables U_R and U_E.

Figure 10 First Order Reliability Method

The whole computation iteration procedure of the FORM method can be summarised in the following ten steps:

1. The limit state function $G(\boldsymbol{X})=0$ is formulated and theoretical models of basic variables $\boldsymbol{X} = \{X_1, X_2, ...X_n\}$ are specified.

2. The initial assessment of the design point $\boldsymbol{x}^* = \{x_1^*, x_2^*, ...x_n^*\}$ is made; for example by the mean values of n – 1 basic variables and the last one is determined from the limit state function $G(\boldsymbol{x}^*) = 0$.

3. At the point $\boldsymbol{x}^* = \{x_1^*, x_2^*, ...x_n^*\}$ equivalent normal distributions are found for all the basic variables using Equations (37) and (38).

4. The transformed design point $\boldsymbol{u}^* = \{u_1^*, u_2^*, ...u_n^*\}$ of the standardised random variables $\boldsymbol{U} = \{U_1, U_2, ...U_n\}$ corresponding to the design point $\boldsymbol{x}^* = \{x_1^*, x_2^*, ...x_n^*\}$ is determined using equation

$$u_i^* = \frac{x_i^* - \mu_{X_i}^e}{\sigma_{X_i}^e} \tag{43}$$

5. Partial derivatives denoted as a vector $\boldsymbol{D}$ of the limit state function in respect of the standardised variables $\boldsymbol{U} = \{U_1, U_2, ...U_n\}$ are evaluated at the design point

$$\boldsymbol{D} = \begin{bmatrix} D_1 \\ D_2 \\ \\ \\ D_n \end{bmatrix} \quad \text{where} \quad D_i = \frac{\partial G}{\partial U_i} = \frac{\partial G}{\partial X_i}\frac{\partial X_i}{U_i} = \frac{\partial G}{\partial X_i}\sigma_{X_i}^e \tag{44}$$

For a linear limit state function $a_0 + \sum a_i X_i = 0$ the derivatives are $D_i = a_i$.

6. The reliability index β is estimated as

$$\beta = -\frac{\{D\}^T\{u^*\}}{\sqrt{\{D\}^T\{D\}}} \quad \text{where} \quad \{u^*\} = \begin{Bmatrix} u_1^* \\ u_2^* \\ \\ \\ u_n^* \end{Bmatrix} \tag{45}$$

For a linear limit state function $a_0 + \sum a_i X_i = 0$ the reliability index is given as

$$\beta = \frac{a_0 + \sum a_i \mu_{Xi}^e}{\sqrt{\sum (a_i \sigma_{Xi}^e)^2}} \tag{46}$$

7. Sensitivity factors are determined as

$$\{\alpha\} = \frac{\{D\}}{\sqrt{\{D\}^T\{D\}}} \tag{47}$$

8. A new design point is determined for n – 1 standardised and original basic variables from

$$u_i^* = \alpha_i \beta_i \tag{48}$$

$$x_i^* = \mu_{X_i}^e - u_i^* \sigma_{X_i}^e \tag{49}$$

9. The design value of the remaining basic variable is determined from the limit state function $G(\boldsymbol{x}^*) = 0$.

10. The steps 3 to 9 are repeated until the reliability index β and the design point $\{\boldsymbol{x}^*\}$ have the required accuracy.

Note that different sign conventions are used in literature and software products concerning the FORM method. In particular, the sensitivity factors in Equations (23), (24) and the derivatives in (47) sometimes have opposite signs to those indicated in the above-mentioned equations. The signs of the sensitivity factors and the derivatives of the limit state function used here are consistent with those provided in EN 1990 [1].

Example 8

Consider the tie rod from Example 1 where the resistance R has the log-normal distribution LN(100, 10) with the lower bound at zero (the skewness $\alpha_R = w_R + w_R^3 = 0.301$) and the load E has the Gumbel distribution GUM(50, 10). Using the iteration procedure indicated above the following results may be obtained: the reliability index $\beta = 2.90$, the failure probability $p_f = 0.0019$, the design point $e_d = r_d = 89.8$, and the sensitivity factors $\alpha_R = 0.36$ and $\alpha_E = -0.93$. Almost the same numerical results are obtained when the Gumbel distribution is approximated by the three-parameter log-normal distribution having the skewness $\alpha_E = 1.14$.

Note that if both the basic variables R and E have the normal distribution, then the reliability index $\beta = 3.54$, the failure probability $p_f = 0.0002$, the design point $e_d = r_d = 75$, and the sensitivity factors $\alpha_R = 0.707$ and $\alpha_E = -0.707$.

8 SIMULATION METHODS

Various simulation methods (direct, adaptive and allocated) are very popular and attractive for their simplicity and transparency. All the simulation methods are based on the generation of random variables of given distribution [5]. Available software products (EXCEL, MATHCAD, MATLAB) include special subroutines for the generation of commonly used types of distributions (uniform, normal, log-normal, Gumbel).

Simulation methods have a number of modifications that can be divided into two basic groups:

– Zero-one indicator based methods, which operate in the original space of variables $\boldsymbol{X}$
– Conditional expectation methods, which can be called semi-analytical methods

The first group of the zero-one indicators includes the direct Monte Carlo simulation (when the original probability density is applied), the method of importance sampling (when the original probability density close to the design point is applied) and the adaptive sampling (updated importance sampling). The second group of the conditional expectation consists of

directional simulation (suitable in the case of a union of events) and axis orthogonal simulation (suitable in the case of an intersection of events).

In the following the direct Monte Carlo method is described briefly. Information concerning more sophisticated simulation methods are available in a number of specialised references.

Simulation of a random variable X having an arbitrary distribution $\Phi_X(x)$ may be in general carried out provided that a generator of random numbers having the uniform distribution in the interval <0, 1> is available. If z_j denotes realisation of a random Z having the uniform distribution in the interval <0, 1>, then the corresponding realisation x_j of the variable X can be obtained using the inverse of the distribution function $\Phi_X^{-1}(z)$, which has the definition domain interval <0,1>. Realisations x_i of the random variable X can be therefore obtained from the relationship

$$x_j = \Phi_X^{-1}(z_j) \tag{50}$$

Using Equation (50), realisations x_{ij} of all basic variables X_i can be generated and then it is verified whether a combination of obtained realisations leads to a failure or not. A failure occurs if

$$G(x_{1i},x_{2i},x_{3i},...) < 0 \tag{51}$$

If the number of all realisations is n and the number of realisations which comply with Inequality (51) is n_f, the failure probability p_f may be assessed using the classical definition of probability based on the ratio

$$p_f = \frac{n_f}{n} \tag{52}$$

Obviously, the assessment of the probability p_f is more accurate when the number of realisations n is sufficiently large. A general rule for the specification of the number n is relatively simple. If the expected failure probability is about 10^{-5}, i.e. from the number of realisations 10^5 on average just one should lead to a failure. Then n should be about two orders greater, thus $n > 10^7$. Note that the coefficient of variation w_{pf} of the failure probability can be estimated using formula

$$w_{pf} = (1-p_f)^{0.5}(n\,p_f)^{-0.5} \tag{53}$$

If $p_f = 10^{-5}$ and $n = 10^7$, then it follows from (53) that the coefficient of variation is $w_{pf} = 0.10$, which is considered a reasonable accuracy. Clearly, to realise $n = 10^7$ generations of all the basic variables is a time-consuming, cumbersome procedure. That is why a number of modifications of the direct Monte Carlo have been developed (zero-one indicator-based methods or conditional expectation methods, methods of Latin Hypercube Sampling, or their combination with FORM). These modifications significantly improve the assessment and decrease the number of required realisations. A detailed description of these methods is available in specialised literature and in manuals to the software products COMREL [7] and VaP [8].

Example 9

Consider Example 4 describing a tie rod R carrying the load E. Assume that R has the log-normal distribution with the lower bound at zero, $\mu_R = 100$ and $\sigma_R = 10$ ($\alpha_R = 0.301$), the load effect E has the Gumbel distribution having the parameters $\mu_E = 50$ and $\sigma_E = 10$ ($\alpha_E = 1.14$). It is known from Example 8 that E may be approximated by the three-parameter log-

normal distribution. As the failure probability is expected by the value $p_f = 1.9\ 10^{-3}$ assessed in Example 4 and 8, the number of trials should be at least about 10^5. The following table shows the results obtained by different methods.

Method used for determining p_f	B	p_f
Second Moment approximation – Example 4	3.54	0.00020
Third Moment approximation – Example 4	3.09	0.00101
FORM – Example 8	2.90	0.00189
Crude Monte Carlo (100000 trials)	2.90	0.00188
Numerical integration	2.90	0.00187

It follows from the above table that the second moment approximation provides only a very rough estimate. The third moment approximation seems to provide a much better result. Obviously the most consistent results are obtained by FORM, crude Monte Carlo simulation, and numerical integration. There is practically no difference between the results obtained by these methods, and $\beta = 2.9$ and failure probability $p_f = 0.0019$ can be considered as a sufficiently accurate estimate of the reliability level.

9 TARGET RELIABILITY LEVEL

Probabilistic reliability methods are based on the comparison of the failure probability p_f with its target value p_t or the reliability index β with its target value β_t. It is generally required to design the structure in such a way that the determined values of the basic variables are close to the target values specified in appropriate code provisions, thus

$$p_f \approx p_t, \text{ or } \beta \approx \beta_t \tag{54}$$

The target values of the reliability index β_t given in EN 1990 [1] were derived mainly from a number of previous reliability studies of structural members made from different materials. However, it should be mentioned that the obtained reliability indices depend on many factors (the type of component, loading conditions and structural materials) and, consequently, have a great scatter. It is known that the results of any reliability study are significantly dependent on the assumed theoretical models used to describe the basic variables. Unfortunately these models are not unified and have not been used systematically. Still, the recommended values of the reliability index may be considered as reasonable average values of the reliability level characterising the existing structures.

Another possibility for specifying the target reliability index or the target failure probability is the minimum requirement for human safety from the individual or social point of view when the expected number of fatalities is taken into account. This approach is briefly described in ISO 2394 [2]. Without going into detail it starts from an accepted lethal accident rate of 10^{-6} per year, which corresponds to the reliability index $\beta_{t,1} = 4.7$. This value corresponds to the target reliability index accepted in EN 1990 [1] for an ultimate limit state per one year.

The reliability index for a period of n years may be then calculated from the following approximate equation

$$\Phi(\beta_{t,n}) = [\Phi(\beta_{t,1})]^n \tag{55}$$

from which the approximate value $\beta_{t,50} = 3.8$ may be obtained from $\beta_{t,1} = 4.7$.

Here $\Phi(.)$ denotes the distribution function of a standardised normal distribution. Figure 1 shows the variation of β_n with β_1 for $n = 5$, 25, 50 and 100. Note that if the reference period T_1 is one year, then n indicates the number of years of the reference period T_n ($n = T_n$).

Figure 11 confirms the data indicated in EN 1990 [1]. For example if the target reliability level of a structure is specified by $\beta_{50} = 3.8$ for the design working life $T = 50$ years, then it could be verified using the reference period $T_1 = 1$ year and $\beta_1 = 4.7$. When, however, the same reliability index 3.8 is specified for a structure having the design working life $T_n = 25$ years only, thus $\beta_{25} = 3.8$, then the reliability of this structure could be verified using the alternative reference period $T_1 = 1$ year and the reliability index $\beta_1 = 4.5$, similarly when $\beta_5 = 3.8$ then $\beta_1 = 4.2$ (see Figure 11).

It should be emphasised that both values $\beta_{t,1} = 4.7$ and $\beta_{t,50} = 3.8$ correspond to the same reliability level, but to different reference periods considered for the assessment of the design values of some actions (1 and 50 years). The reference period may, or may not, coincide with the design working life.

A completely different question is the specification of the reliability index for a construction works of a limited design working life. A practical illustration of the numerical calculation of the reliability index is shown in Example 10.

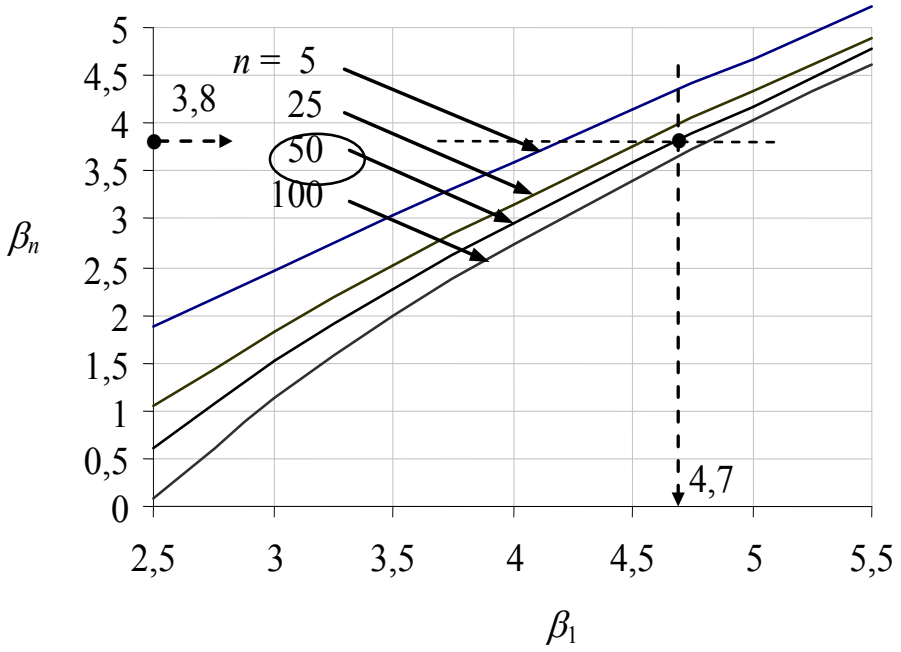


Figure 11 Variation of β_n with β_1 for n = 5, 25, 50 and 100

It should be noted that the actual frequency of failure may be dependent on many factors not considered in partial factor design and, consequently, the reliability index β may not correspond to actual frequency of structural failure.

Example 10

Consider an agricultural structure with moderate consequences of failure and a limited design working life of 25 years. In that case it may be reasonable to specify $\beta_{t,1} < 4.7$, say $\beta_{t,1} = 4.2$. Using now the Equation (55), it can be found that for the design working life $n = 25$ years

$$\Phi(3.4) = [\Phi(4.2)]^{25}$$

and thus $\beta_{t,1} = 4.2$ corresponds to $\beta_{t,25} = 3.4$. Note that using the same Equation (55) for $n = 50$ years, it follows that $\beta_{50} = 3.2$. The correct interpretation of this finding is as follows: if the input data (in particular actions) are related to 1 year and the design calculations are done for this period, then $\beta_1 = 4.2$ should be considered; if the input data are related to 25 years, then $\beta_{t,1} = 3.4$ should be considered in the design verification.

10 PROBABILISTIC OPTIMISATION

Probabilistic optimisation is another way of estimating an adequate reliability level. In general, with increasing reliability level the cost of a structure increases and unfavourable consequences due to structural failure decrease. The basic idea of probabilistic optimisation is to find such a reliability level which would minimize the total cost. To illustrate this concept a simple objective function describing the total cost is considered in the following. Assume that the total expected cost C_{tot} of a structural member can be approximated by the objective function

$$C_{\text{tot}} = C_0 + C_P\, P + C_{\text{f}}\, p_{\text{f}}(P) \tag{56}$$

Here C_0 denotes the initial cost, which is assumed to be independent of the decisive parameter P, the product $C_P\, P$ is the additional cost of the member caused by the chosen magnitude of the parameter P (C_P denotes the cost per unit parameter P), the product $C_{\text{f}}\, p_{\text{f}}(P)$ is the expected cost caused by the failure of the structural member. The necessary condition for the minimum of the total cost is given by the derivative

$$\partial C_{\text{tot}} / \partial P = C_P + C_{\text{f}}\, \partial p_{\text{f}}(P) / \partial P = 0 \tag{57}$$

Thus the optimum P (if it exists) may be determined from the condition

$$\partial p_{\text{f}}(P) / \partial P = - C_P / C_{\text{f}} \tag{58}$$

Equation (58) can be used in a computer program to determine the optimum value of the parameter P. However, it is often less difficult and more transparent to compute the total cost C_{tot} using Equation (56) or a similar (more sophisticated) objective function.

Example 11

Consider the reinforced concrete slab described in the Appendix to Chapter 1. Consider the fundamental objective function (56), in which the reinforcement area A is considered the decisive parameter P. Figure 12 shows the total cost of the slab as a function of the reinforcement area A assuming the partial costs C_A = 100, C_{f} =1000 of some financial units. The results shown in this example were obtained using the programme COMREL [7] of the software product STRUREL [8].

It follows from Figure 12 that under the given assumptions the optimum reinforcement area seems to be close to A = 0.0009 m^2, which is the value slightly lower than the design area obtained by the partial factor method in the Appendix to Chapter 1. Figure 12 also shows the variation of the reliability index β with the reinforcement area A. Note that the optimum reinforcement area A = 0.0009 m^2 leads approximately to the reliability index β = 4.2 (see Figure 12).

Obviously the optimum area A is dependent on the assumed partial costs C_A and C_{f}. Figure 13 shows the optimum reinforcement areas A for selected partial costs C_A and C_{f} and the reliability index β that is independent of partial costs C_A and C_{f}. The cost per unit reinforcement area C_A is constant (100 units) and only the cost of failure varies from 10 to 100000 units.

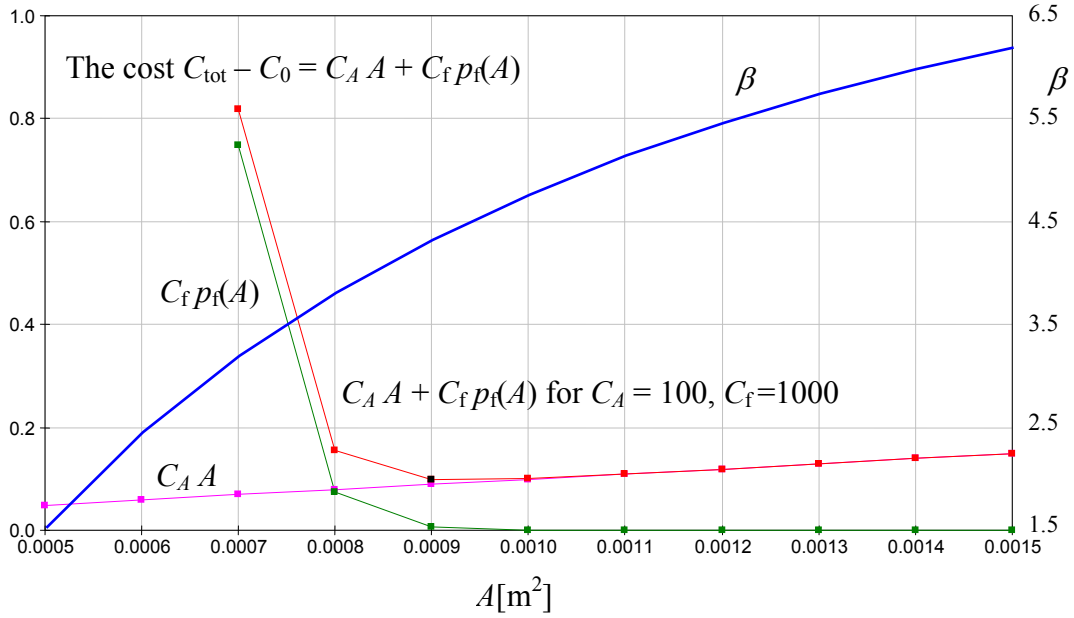


Figure 12 Variation of the total cost of the slab and the reliability index β with the reinforcement area A assuming the partial costs $C_A = 100$, C_f=1000 [units]

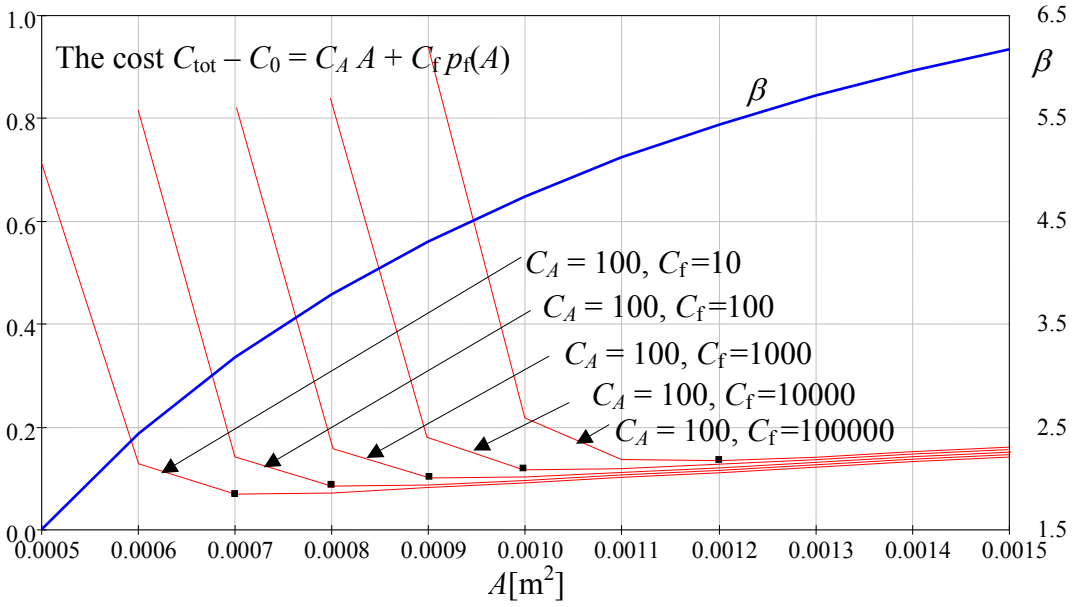


Figure 13 Variation of the total cost and the reliability index β with the reinforcement area A for selected partial costs C_A and C_f

It follows from Figure 13 that with increasing cost of failure C_f the optimum reinforcement area A and corresponding reliability index increase. It should be noted that for a very low cost C_f (less than or equal to C_A) the reliability level may be rather low (less than 3.5) and may become unacceptable. The reliability requirements on the maximum failure probability may then turn out to be decisive and the optimal reinforcement area A cannot be used.

11 UPDATING OF PROBABILITY DISTRIBUTION

If the prior probabilities distribution is described by a continuous probability density function $\varphi(x)$ of a random variable X and likelihood derived from newly available data by conditional density function $\varphi(I|x)$ where I denotes outcomes of additional investigation, then *a posteriori* (updated) probability density $\varphi(x|I)$ may be derived [5,6] as

$$\varphi(x \mid I) = \frac{\varphi(x)\varphi(I \mid x)}{\int \varphi(x)\varphi(I \mid x)\mathrm{d}x} \tag{59}$$

Note that the likelihood $\varphi(I|x)$ is the conditional probability density function describing the probability that the outcome of the updating investigation I (information obtained from I) is due to the occurrence of x. Formulae (59) can be used for the updating of distribution functions when additional experimental investigations are used for assessing new or existing structures.

Example 12

Assume that a variable X has a normal prior distribution with probability density function $\varphi(x)$ having the mean μ and the standard deviation σ. Additional investigation indicated that the likelihood function $\varphi(I|x)$ is described by a general three-parameter log-normal distribution having the same standard deviation σ but the mean equal to $\mu + 0.5\ \sigma$ and the skewness $\alpha = 1$. Using a numerical integration of it follows from Formulae (59) that the updated distribution $\varphi(x|I)$ has the following parameters

$$\mu_{x|I} = \mu + 0.18\ \sigma$$
$$\sigma_{x|I} = 0.64\ \sigma$$
$$\alpha_{x|I} = 0.39$$

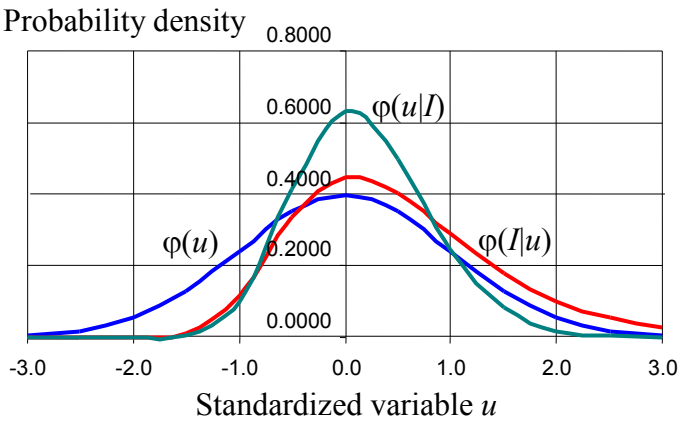


Figure 14 Prior probability density $\varphi(x)$, likelihood $\varphi(I|x)$ and the updated probability density function $\varphi(x|I)$

Figure 14 shows the prior probability density $\varphi(u)$, likelihood $\varphi(I|u)$ and the updated probability density function $\varphi(u|I)$ using standardized random variable U.

It follows from Figure 14 that the updated distribution has considerably lower variability than the prior distribution. Obviously updating of probability distributions may be extremely effective when assessing characteristic values of the resistance variables using additional tests.

12 CONCLUDING REMARKS

The theory of structural reliability becomes a powerful tool when used for the development of new standards or, alternatively, for the direct verification of both new and existing structures. Recently revised national and international standards for structural design are systematically based on probabilistic concepts, mathematical statistics and the theory of structural reliability. This approach has also been used by the European committee for standardization (CEN) in developing the new European standards for structural design, called Eurocodes [1], and by the international standard organisation (ISO) in developing recent international standards [2,3].

REFERENCES

[1] EN 1990 Basis of design. European Committee for Standardisation. CEN (2002).
[2] ISO 2394, General principles on reliability for structures (1998).
[3] ISO 13822 Assessment of existing structures. (2002).
[4] Probabilistic model code, Parts 1 to 4, Basis of design, Load and resistance models, Examples, JCSS (2001-2002).
[5] Holický M. (2009) Reliability analysis for structural design, SUN MeDIA Stellenbosch, ZA.
[6] Holický M. (2013) Introduction to probability and statistics for engineers, Springer.
[7] STRUREL, Structural Reliability System (includes programmes STATREL, COMREL, SYSREL and NASREL – COMREL version 8.00 of 2003). RCP Consulting software, Munich, Germany.
[8] VaP, Variable Processor 2.0. Petschacher Consulting, Feldkirchen, Austria (2003).

CHAPTER 6: RISK ANALYSIS

Dimitris Diamantidis[1], Raffaele Boccara[2] and Milan Holický[3]

[1]OTH Regensburg, Germany
[2]OTH Regensburg, Germany, and University of Pisa, Italy
[3]Klokner Institute, Czech Technical University in Prague, Czech Republic

Summary

Risk assessment plays an essential role in the risk management of infrastructure systems. This chapter provides the general formulation of risk depending on two parameters: hazard probability and associated consequences. Structures and infrastructure systems are exposed to a multitude of possible hazards. In the case of civil engineering structures, these hazards include both, those from the environment (wind, temperature, snow, avalanches, rock falls, ground effects, water and ground water, chemical or physical attacks etc.) and those from human activities (usage, chemical or physical attacks, fire, explosion). Methods and tools for the calculation of risk are provided and illustrated.

1 INTRODUCTION

1.1 Background Documents

Relevant background documents are given also in the introductory Chapter 1. Methods of risk assessment and analysis are more and more frequently applied in various technical systems (see for example [3], [4], [6]) including chemical and nuclear plants, road tunnels, bridges and offshore structures. Available national and international documents ([16], [17], [18], [19], [20], [21]) try to harmonise general methodical principles and terminology that can be also applied in the risk assessment of infrastructure and its systems. Several other literature sources are provided in this Handbook 1.

1.2 Scope

Risk assessment is indispensable in order to identify hazards, assess vulnerabilities of the infrastructure and evaluate the associated impact (consequences) on assets, infrastructures or systems taking into account the probability of the occurrence of the identified hazards. This is a critical element that differentiates a risk assessment from a typical impact assessment methodology. There are a significant number of risk assessment methodologies applicable to risk-based evaluation of infrastructure. However, there is a differentiation of risk assessment methodologies based on the scope of the methodology, the audience to which it is addressed (policy makers, decision makers, research institutes, practicing engineers) and the degree of sophistication.

The scope of this chapter is to summarize risk analysis methods used in civil engineering and to illustrate them in characteristic examples. Case studies applications are

presented in Handbook 2. Risk analysis methods are important for decision making and for optimal management of the aging infrastructure.

2 RISK FORMULATION

2.1 General Framework

The risk combined with a hazard is a combination of the probability of occurrence of this hazard and the consequences in case that it occurs, as mentioned in the introduction paragraph of this chapter. Therefore the most sophisticated methods for risk analysis take all consequences of a failure directly into account. The basic relation for this kind of risk calculation can be represented as

$$R = \iint_{\dots} \int C(X_1, X_2, \dots, X_\mathrm{n}) f\,(X_1, X_2, \dots, X_\mathrm{n}) \mathrm{d}X_1 \mathrm{d}X_2 \dots \mathrm{d}X_\mathrm{n} \tag{1a}$$

where R is the risk, C is the consequences and f is the joint probability density function of the basic random variables. This integral is difficult to be solved especially under the consideration of time effects described in Chapter 2. Therefore, the simplest function relating the two constituents of risk is often use by multiplying the probability of failure, p_f by the consequences C:

$$R = p_\mathrm{f}\, C \tag{1b}$$

The calculation of the failure probability has been discussed in the previous Chapter 5. The consequences may be expressed in monetary units or in terms of injured or dead per event, or by some other indicator. Since fixed deterministic values are used for considering the consequence of a possible failure, the application of Equation (1b) is formally simple. As in all this kind of studies, the main problem is the adoption of reasonable quantities for the constituent of risk, particularly for the possible damage. But as long as the results are interpreted in a comparative way, extremely useful information is obtained for optimization problems.

Independently on the representation of risk, difficulties exist with very small probabilities or frequencies of occurrence of events with very large consequences. In such cases the product rule is no longer applicable since, from a mathematical point of view, zero times infinity may take any value. For this and other reasons, risks have to be considered in more detail. Terms like acceptable risks, voluntary and involuntary risks, individual and collective risks or residual risk are clearly to be distinguished. It should also be observed that the subjective perception of risk often differs from what could be called objective risk.

2.2 Categorization of Consequences

One of the main steps of a risk analysis is the quantification of the ‘cost of failure’. A systematic procedure to describe and if possible quantify consequences is required. In general, consequences resulting from civil structures failure may be divided into four main categories:

– Human
– Economic
– Environmental
– Social

Table 1 presents a list, certainly not exhaustive, for the purpose of undertaking risk assessment of major structural systems. As described above the consequences can be direct or indirect; note that according to the system under investigation the same consequence can be direct or indirect.

Table 1 Categorization of consequences [1]

Type	Direct	Indirect
Human	Injuries Fatalities	Injuries Fatalities Psychological damage
Economic	Repair of initial damage Replacement/ repair of contents Rescue costs Clean up costs	Replacement/ repair of structures/ contents Rescue costs Clean up costs Collateral damage to surroundings Loss of functionality/ production/ business Temporary relocation Traffic delay/management costs Regional economic effects Investigations/ compensations Infrastructure inter-dependency costs
Environmental	CO_2 emissions Energy use Pollutant releases	CO_2 emissions Energy use Pollutant releases Environmental clean-up/ reversibility
Social		Loss of reputation Erosion of public confidence Undue changes in professional practice.

It is evident that the level and sophistication of the various analysis types increases considerably as the range and extent of considered consequences widens. Reference [1] suggests advanced structural analysis, considering a multitude of non-linear material and geometric effects, when a particular failure scenario needs to be taken beyond initial damage and member failure.

These consequences can be measured in terms of damaged, destroyed, expended or lost assets and utilities such as raw materials, goods, services and lives. They may also include intangibles, either from a practical or theoretical standpoint especially in the case of social consequences and long-term environmental influences. Where possible the consequences should be described in monetary units, though this is not easy to achieve, and may not be desirable or, indeed, universally acceptable.

Sources for the quantification of consequences from structural failures can be found in literature and some examples are collected by [7]. The information includes:

– Natural hazards loss estimation manuals (e.g. [8])
– Benchmark studies (e.g. [9])

– Reports analyzing past failures (e.g.[10] and [11])
– Industry and regulatory authorities guidelines (e.g. [12] and [13])
– Insurance reviews (e.g. [14])
– General literature

Several expressions for the consequences can be found. For example [15] introduce, in case of earthquake, an expression for the number of fatalities, K_s, in a particular building

$$K_s = K_m M_2 M_3 (M_4 + K_5) \quad (2)$$

where

K_m average number of people in the building;
M_2 factor between zero and one related to the building occupancy cycle;
M_3 factor that accounts the fraction of people that would be trapped in a collapse following an earthquake;
M_4 factor related to immediate mortality rates in earthquake event;
M_5 factor related to post-collapse mortality rates in earthquake event.

There exist other models, which for example follow an event tree philosophy, and the associated conditional probabilities have also been developed for regional loss estimation [14]. This study provides an empirical expression for K_s in the case of a bridge collapse:

$$K_s = N_C F \quad (3)$$

where

N_C the commuter population;
F usage factor which depends on the assumed time of the accident (similar to M_2). The values range is 0.02 during peak times and 0.01 otherwise.

Economic consequences models are, on a whole, available for a wide variety of building and bridge structures, especially with respect to repair/reconstruction costs, typically linked to a damage severity index. An important distinction between structural and non-structural costs is often made, though data for the latter are more difficult to collect and categorize. Some characteristic case studies will illustrate these aspects in Handbook 2.

2.3 Simple Example of Risk Calculation

Consider for example that the collapse of a structure has a probability of failure of 10^{-3} per year in at site (a) and 10^{-2} per year at site (b). The designer has to decide where to install this structure. Analyzing the situation the designer computes that the economic loses of the failure are quantifies in 2 000 000 € in site (a) and 1 500 000 € in site (b). It is assumed that all the other costs are equal in both cases. The designer now can use the risk definition to decide where to install the structure. In Table 2 the used data are summarized.

Table 2: Summarized data of the example of risk calculation

Site	p_f [/year]	Failure costs [€]	Risk [€/year]
(a)	10^{-3}	2 000 000	2 000
(b)	10^{-2}	1 500 000	15 000

For this simple calculation the basic definition of risk was used (see Equation 1.b). As summarized in the table the risk of scenario (a) is 2 000 €/year and on scenario (b) 15 000 €/year so is immediate that the designer will choose scenario (a).

3 RISK ANALYSIS TOOLS

3.1 Basic Aspects

A typical risk analysis scenario approach is developed into three main steps [2]:

- Identification of hazards (see for example Chapter 2)
- Evaluation of the hazards and determination of the risk corresponding to the different scenarios
- Planning of safety measures if needed

The first step, the qualitative identification of the hazard is the most important and difficult step of the approach. Indeed, once the potential hazards and combination of hazards are recognized, usually it is relatively strait forward to adopt appropriate measures to overcome their consequences. Various techniques exist that may help the engineer to recognize possible hazards. The common characteristic of all the techniques is that they are based on asking questions. If the right questions are asked it is quite simple to identify all the potential hazards, but it is necessary to have imagination and creativity. When searching for potential hazards it is necessary to have, at the beginning stage of the design, a complete idea of the whole construction, use, repair and/or its future replacement or demolition process. For a correct conduction of this aspect it is useful to have experience and to consult specialized literature.

Different strategies of thinking with a view to identifying possible hazards and hazard scenarios are known under various names, for example Hazard Operability Study (HAZOP), What-if Analysis, Failure Mode and Effect Analysis (FMEA) [3]. In daily practice the goal is to recognize all possible hazards related to a particular problem. In order to reach this goal, a combination of different of these strategies, which are listed below is to be applied:

- In *chronological analysis*, the whole process, step by step, has to be established in mind. Typical questions to be asked are: what will occur, where and when?
- In an *utilization analysis* questions concerning the use of the infrastructure are to be raised: what equipment or machines will be used? What influence do they have? What could go wrong?
- In an *influence analysis* it is looked at influences from the natural environment and from human activities. New situations have to be anticipated, which could make initially harmless influences dangerous. Furthermore, it must be looked at individual hazards that alone would be negligible, but in combination become dangerous;
- *Energy analysis* where and in what circumstance potential due to different energies could lead to a hazardous situation? The failure of an energy supply can also constitute a danger.
- *In material analysis* it has to be looked, for example, at the durability, combustibility, toxicity or explosiveness of the used raw materials.
- *Examining interfaces* hazards can be anticipated where different materials are in contact, or where information has to be transmitted, or where responsibilities are not clearly defined.

The morphological thinking is a method developed in [4] and its main points, listed above, are taken from [2].

3.2 Risk Analysis Methodologies

In this paragraph a brief description of the main risk analysis methodologies is illustrated. The main aspects are here taken from [3], [4] and [5].

3.2.1 Hazard and Operability Studies (HAZOP)

The Hazard and Operability Studies (HAZOP) was developed in the early 1970s. HAZOP can be defined as the application of a formal systematic critical examination of the process and engineering intentions of new or existing facilities. This technique is usually performed using a set of guidewords: NO/NOT, MORE/LESS OF, AS WELL AS, PART OF REVERSE, AND OTHER THAN. From these guidewords scenarios that may result in a hazard or an operational problem are identified. Consider the possible flow problems in a process line, the guide word MORE OF will correspond to high flow rate, while that for LESS THAN, low flow rate. The consequences of the hazard and measures to reduce the frequency with which the hazard will occur are then discussed. This technique had gained wide acceptance in process industries as an effective tool for plant safety and operability improvements.

3.2.2 Failure Mode and Effects Analysis (FMEA/FMECA)

This method was developed in the 1950s by reliability engineers to determinate problems that could arise from malfunctions of military systems. Failure mode and effects analysis is a procedure by which each potential failure mode in a system is analyzed to determine its effects on the system and to classify it according to its severity.

When the FMEA is extended by a criticality analysis, the technique is then called Failure Mode and Effects Critically Analysis (FMECA). Failure mode and effects analysis has gained wide acceptance by the aerospace and military industries. In fact, the technique has adapted itself in other form such as measure mode and effects analysis.

3.2.3 Tree Based Techniques

In order to introduce some clarity and completeness in the engineering work, logic trees (fault tree, event tree, cause-consequent chart) are used in risk analysis. The use of such tool is very widespread in risk analysis and implies some important advantages. Influence from the environment and from human activities can easily be considered simultaneously. Logic trees also can contribute to detect the most effective countermeasures. Furthermore, they are easy to understand and therefor very helpful for communication purposes with non-experts.

Fault Trees Analysis

The concept of fault tree analysis (FTA) was originated by “Bell Telephone Laboratories” in 1962 as a technique with which to perform a safety evaluation of the Minutemen Intercontinental Ballistic Missile Launch Control System.

A fault tree can be defined as a logical diagram for the representation of combinations of influences that can lead to an undesired event. When establishing a fault tree, the undesired event constitutes the starting point. Going out from this event the possible causes are to be identified. Possible causes and consequences are to be linked, in a logic way, without

introducing any loops. Every event that is not a consequence of a previous event has to be considered as an independent variable [6].

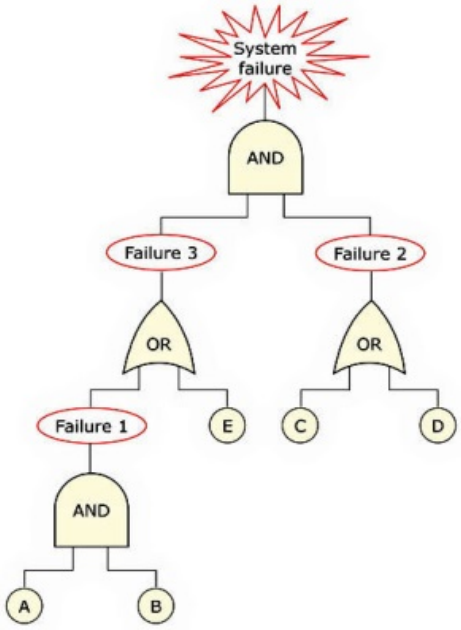


Figure 1 Fault tree example construction

After a failure event, fault trees can be used in order to clarify the cause in a case where they are unknown. The most common application, however, consists in detecting possible causes of undesirable events before they can occur. A typical example for a fault tree construction is shown in Figure 1.

The example of the fault tree shown in Figure 2 describes a failure of a plane frame. In Figure 2, all relevant failure mechanisms are represented. Concerning this example, it must be highlighted that in the *sway* mechanism the vertical load ($G+Q$)has no influence and that in the *vertical* (*beam*) mechanism the horizontal load H has no influence, while the activation of the alternative *combined* mechanisms 1 or 2 obviously depends on the directions of applied loads: in effect, according to the *kinematic theorem* of limit analysis, the actual failure mechanism corresponds to the lower bound of load multipliers corresponding to all possible failure mechanisms.

Once the qualitative point of view is described, the calculation of the probability of occurrence of the undesirable top event should be performed. For fault trees the probabilities have to be calculated depending on the type of the logic gate. If the different components must fail at the same time (the first and the second and…), they constitute a parallel system represented by an AND-Gate. Therefore, the corresponding probability is obtained according to the multiplications rule, by multiplying the probabilities of the different components. On the other hand, if only one component must fail (either the first, or the second, or…), they represent a series system. In that case we are talking about as OR-Gate and the probability of failure is obtained by adding the probabilities of the different components [5].

Fault tree analyses help:

– To gain an understanding of the system
– To document the failure relationships of the system
– To exhaustively identify the causes of a failure

– To assure compliance with requirements or a goal
– To identify any weaknesses in a system
– To prioritize contributors to failure
– To identify effective upgrades to a system
– To optimize operations and processing
– To quantify the failure probability and contributors

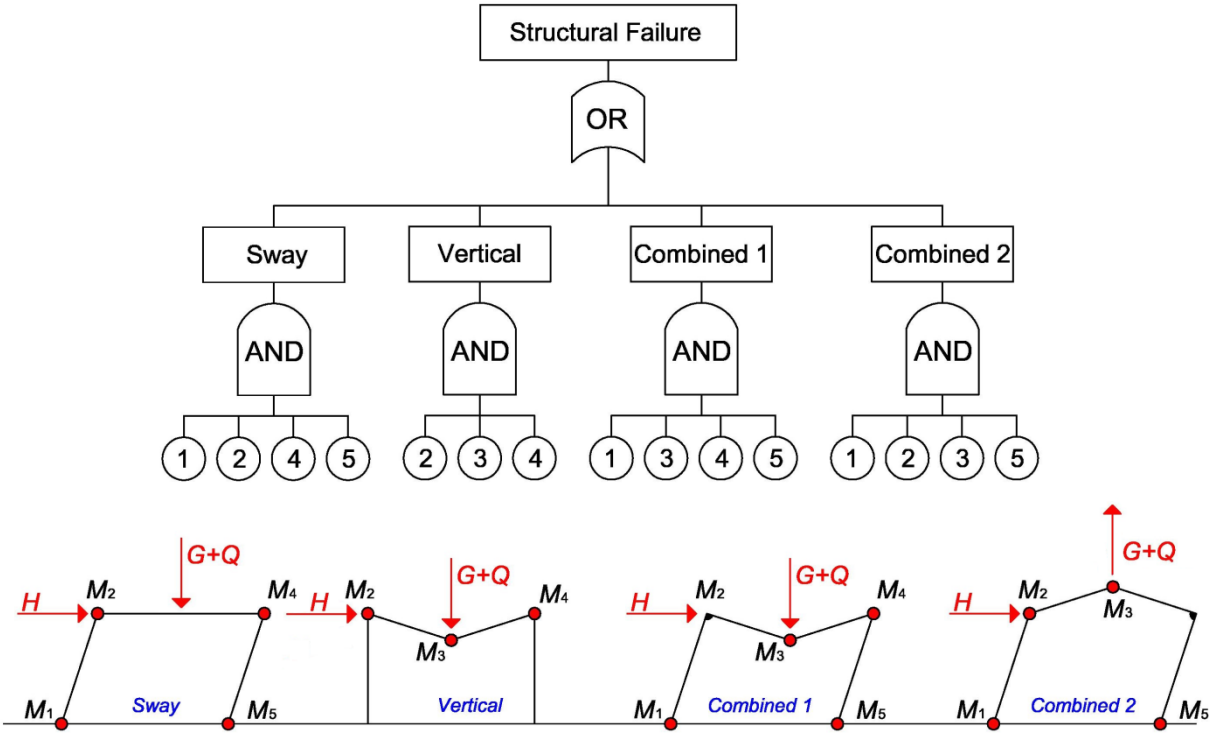


Figure 2 A fault tree describing the failure of a plane frame

Event Tree Analysis

Starting from an initial event an event tree identifies all the possible consequences events. Each path consists of a sequence of events and ends up at the consequence level as shown in Figures 3 and 4. The aim is the establishment of possible consequences of an initial event. In a second step the event tree also can be used for the calculation of probabilities of occurrence of these consequences.

Once the qualitative point of view is described, the calculation of the probability of occurrence of the consequences of a given initiating event should be performed. For event trees as shown above the possible subsequent events following a previous event exclude each other. The sum of the probabilities at a gate of an event tree must be the unity. The calculation of the probability of occurrence of a consequence is very simple since it is obtained by multiplying the probabilities of the different events constituting the path that leads to the considered consequence (conditional probabilities). The main point is the establishment of the probabilities of the different branch events, and very often numerical values are based on subjective estimations [5]. Therefore sensitivity analyses are recommended in order to investigate the influence of the uncertain input parameters. In summary the event tree construction comprises the following steps:

– Identification of a relevant accidental (initial) event
– Identification of the barriers that are designed to deal with the event

– Construction of the event tree
– Description of the (potential) resulting accident sequences
– Determination of the frequency of the accidental event and the (conditional) probabilities of the branches in the event tree
– Calculation of the probabilities/frequencies for the identified consequences
– Compilation and presentation the results from the analysis

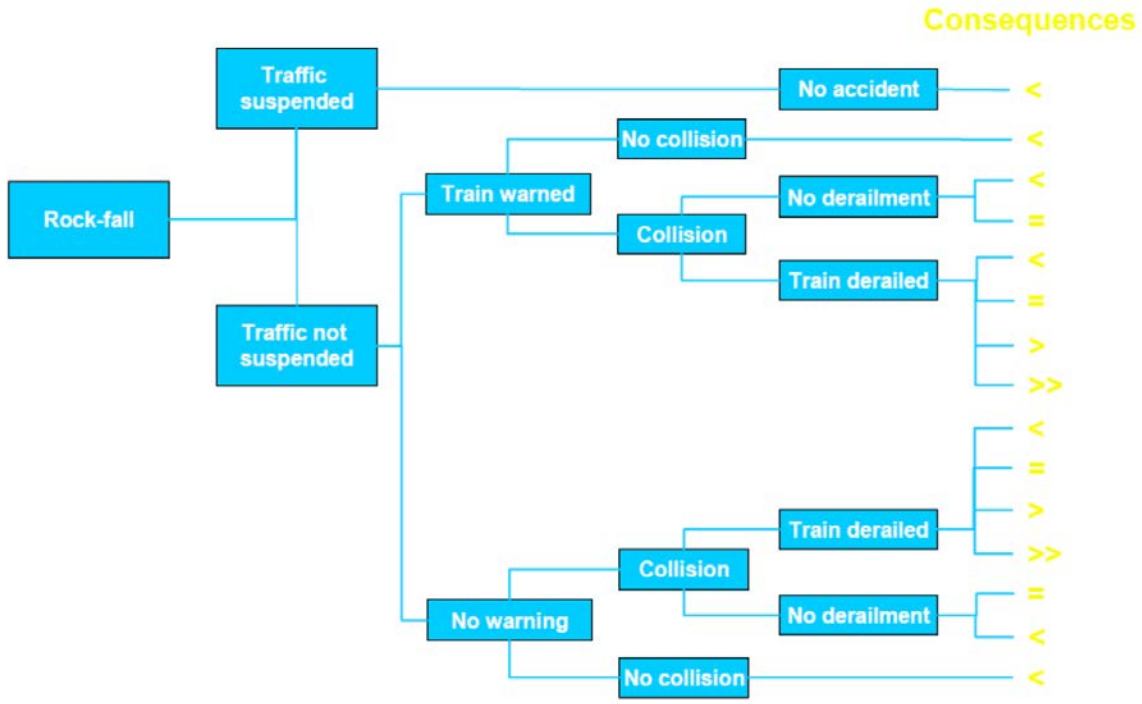


Figure 3 Example of an event tree for a rock fall on railway-line [2]

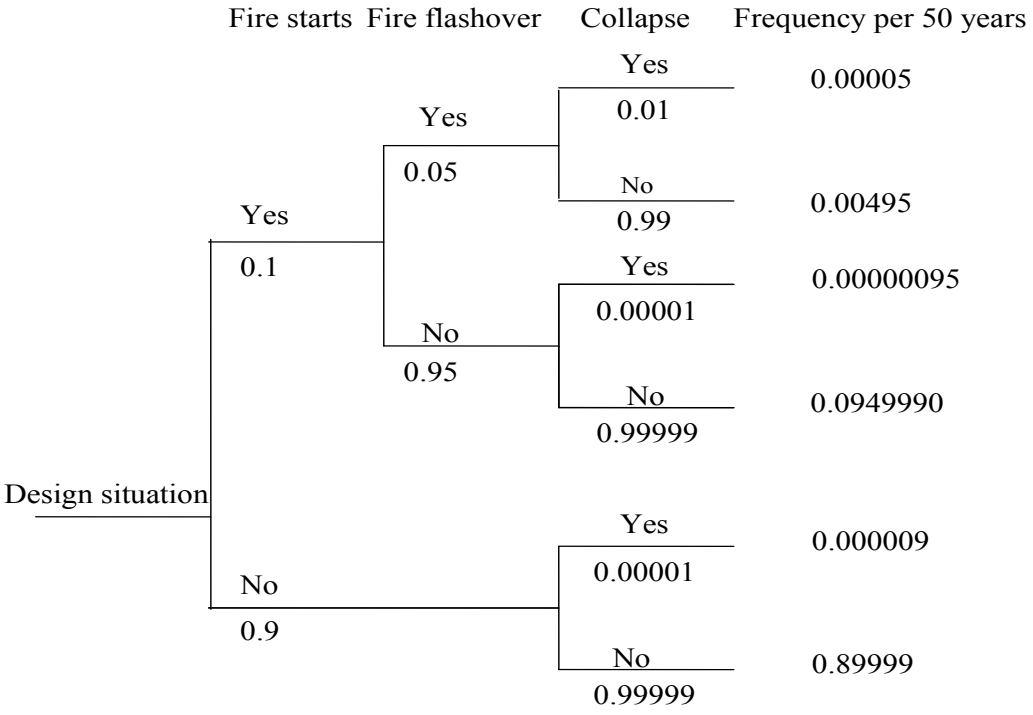


Figure 4 Event tree describing the collapse of a structure under fire design situations

Cause/ Consequence Analysis

The two previous analysis methods can be combined to obtain a Cause/Consequences Analysis (CCA). This technique combines cause analysis, described by fault trees, with consequence analysis, described by event trees, and thus deductive and inductive analysis. The questions are formulated in a way that the answer can be only "yes" or "no". In this way a very compact representation of complex problems can be obtained. With the probability of the various events in the CCA diagram, the probabilities of the various consequences can be calculated, thus establishing the risk level of the system [5].

The simplest form of the cause/consequences consideration is the so-called prior-analysis of the risk (utility), when the basic statistical and probabilistic information is available prior to any decision or activity. The prior analysis is an assessment of the risk associated with different decisions; it is commonly used for comparing the risks corresponding to different decisions. The posterior decision analysis differs from the prior analysis be considering possible changes in the branching probabilities and/or the consequences due to risk reducing measures, risk mitigating measures, and the collection of additional information. The posterior decision analysis may be used to evaluate different additional activities affecting the total risk.

Another important modification of logic trees is known as the pre-posterior decision analysis. The aim of the pre-posterior decision analysis is to identify the optimal decisions with regard to activities that may be performed in the future, for example a planning of risk reducing activities and/or the collection of the new information. An important per-requisite for the pre-posterior decision analysis is the consideration of future actions that may be applied taking into account the results of the planned activities.

3.3 Example of a Simple Decision Analysis

In order to illustrate decision analysis a simple example is presented first. The investigated problem refers to a design situation in which a beam of a truss is analyzed. The beam is subjected to a compression force. The designer has to decide if (a) to keep the existing beam (b) or substitute the beam with a one with a larger area A_1 (c) or even with an area $A_2 > A_1$. Supposing that in the first event (a) the probability of failure is $1.2\ 10^{-2}$ per year; that in the second case (b) 10^{-3} per year and in the third case (c) 10^{-4} per year.

Analyzing the consequences of the failure of the truss structure the designer evaluates the collapse consequences in a monetary quantity i.e. 1.000 000 €, to be the same, for all the scenarios (a), (b) and (c). Analyzing the different cost of realization of the different situations the designer evaluate that case (a) has no investment cost i.e. 0.00 €. Case (b) has investment costs of 10 000 € and case (c) 13 000 €. The small difference between the costs of case (b) and (c) is due only to material costs, all other costs, demolition, disposal and reconstructions are the same. All the previous information and the outcomes are summarized in Table 3 and in the Figures 5 and 6.

Table 3 Data of the Example for the decision analysis of a beam of a truss structure

Scenario	Probability of failure [/year]	Scenario cost [€]	Collapse consequences [€]
a	$1.2\ 10^{-2}$	0	1 000 000
b	10^{-3}	10 000	1 000 000
c	10^{-4}	13 000	1 000 000

At this point the designer has to decide which of the three options is the best choice. The expected costs are calculated in the following way:

$$E[d_0] = \theta_0(d_0)\, C_0(d_0) + \theta_1(d_0)\, C_1(d_0) = (1 - 1.2\ 10^{-2})\ 0.00 + 1.2\ 10^{-2} 10^6 = 12\ 000 \quad (4a)$$

$$E[d_1] = \theta_0(d_1)\, C_0(d_1),\ \theta_1(d_1)\, C_1(d_1) = (1 - 10^{-3}) 10\ 000 + 10^{-3} 1\ 010\ 000 = 11\ 000 \quad (4b)$$

$$E[d_2] = \theta_0(d_2)\, C_0(d_2),\ \theta_1(d_2)\, C_1(d_2) = (1 - 10^{-4}) 13\ 000 + 10^{-4} 1\ 013\ 000 = 13\ 100 \quad (4c)$$

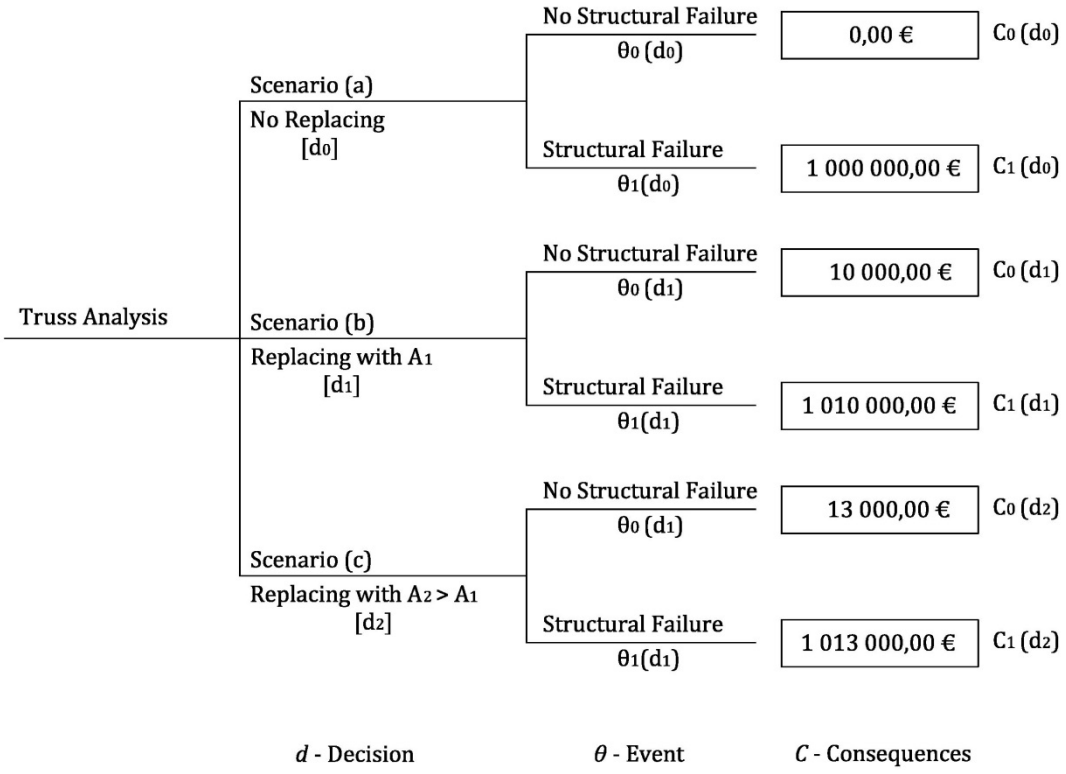


Figure 5 Event tree for the presented example

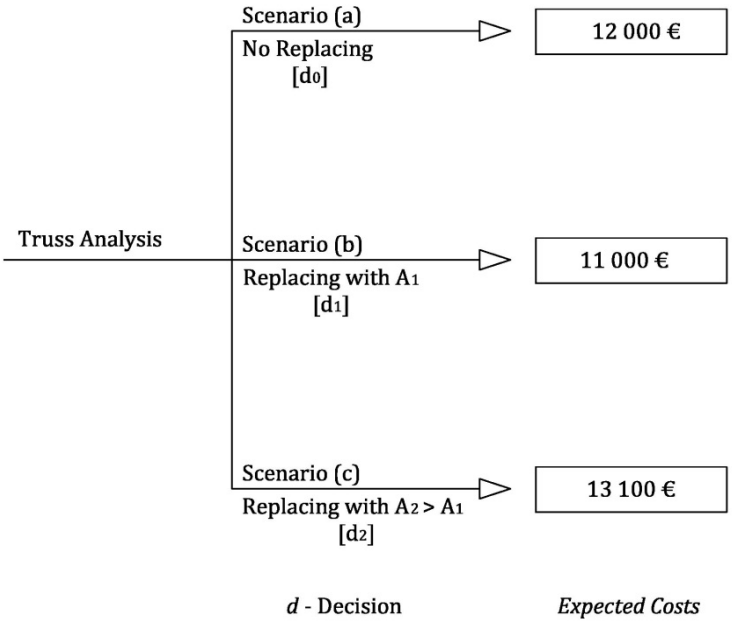


Figure 6 Summary of utilities in the scenarios of the example

It is shown in Figure 6 that the decision d_1 yields the largest expected utility (smallest cost) and consequently the decision that implies replacing the beam with a new one having area A_1 is the 'optimal decision'.

3.4 Bayesian Network

A very effective tool for risk analysis is the Bayesian (belief) causal networks [22]. A simple example of the causal network is shown in Figure 7. The network containing only four chance nodes describes the structural failure under persistent and fire design situations similar to the event tree in Figure 4. Compared with the event tree shown in Figure 4 the network in Figure 7 also includes the effect of sprinklers (node B). Note that the directional arrows in Figure 7 indicate the causal links between interconnected chance nodes.

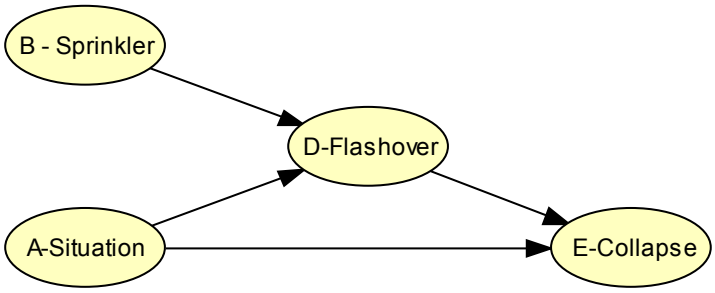


Figure 7 Causal network describing the structural failure under persistent and fire design situations

The collapse of a structure depends on the probability of persistent and fire situations and on the conditional probabilities of the full development of fire, which depend on the capability of sprinklers and on the conditional probability of the structural collapse under the conditions given by parent nodes (for example when fire is fully developed – fire flashover). Obviously the causal network representation seems to be much more effective than the event

tree version. Moreover, each node may have several states. Consequently, the input data are not indicated directly in the graphic representation of the network but are given in the tables of conditional probabilities.

The basic principle of the probability calculation used in the Bayesian networks may be illustrated considering the nodes A, B and D of the network in Figure 7. One child node D (Fire flashover) is dependent on two parent nodes: A (Design situation) and B (Sprinklers). If the parent nodes A and B have the discrete states A_i and B_j, then the probability of the event D_k (a particular state of the node D) is given by the formula

$$P(D_k) = \sum P(D_k| A_iB_j)P(A_i)\ P(B_j) \tag{5}$$

Equation (5) represents a fundamental theoretical tool for analysing the Bayesian network. The input data consist of the probabilities $P(A_i)$ and $P(B_j)$, and the conditional probabilities $P(D_k|A_iB_j)$. These extensive data are based on available statistical evidence, probabilistic analysis or expert assessment (judgement) and are transparently summarised in the tables of conditional probabilities.

Bayesian networks supplemented by decision and utility nodes called influence diagrams provide a powerful tool for risk estimation. In fact, the influence diagram is a generalisation of the cause/consequence-chart discussed above. The main features of this tool are illustrated by the example shown in Figure 8, which is an extension of the fundamental task indicated in Figure 7. Figure 8 shows a simplified influence diagram, which has been developed recently [23] for the risk analysis of buildings under persistent and fire design situations.

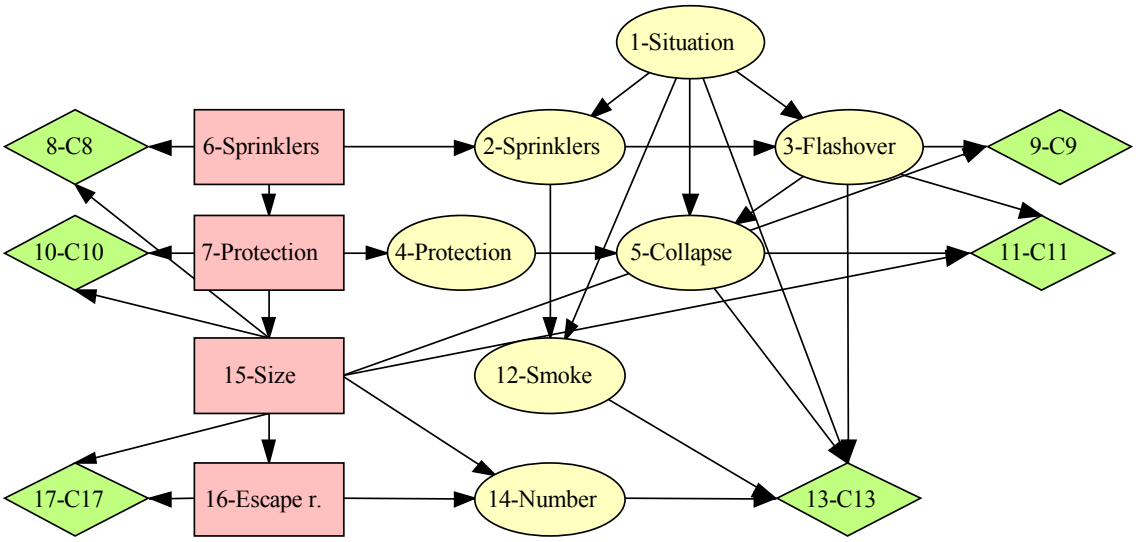


Figure 8 Bayesian network describing a structure under normal and fire design situations

The network consists of seven chance nodes numbered 1, 2, 3, 4, 5, 12 and 14, four decision nodes 6, 7, 15 and 16, and six utility nodes 10, 9, 10, 11, 13 and 17. The utility nodes represent the costs of various fire safety measures (nodes 8, 10, 17), the damage to the building (nodes 9, 11), and injuries (node 13).

Directional arrows indicating the causal links between the parent and child nodes interconnect the chance, decision and utility nodes. All the causal links must be described by

appropriate input data (conditional probabilities or utility units) linked to assumed states of the nodes. For example the utility nodes (except the utility node 13) are directly dependent on the size of the building (node 15). The utility node 13, describing the cost of injury, is affected by the size of the building through the number of endangered persons represented by the chance node 14. These data are often difficult to specify, and an expert assessment has to be made.

The chance nodes 1, 2, 3, 4, 5, 12 and 14 represent alternative random variables having two or more states. The node 1-Situation describes the probability of fire start $p_{\mathrm{fi,s}} = \mathrm{P}(H_2)$ and the complementary probability $1 - p_{\mathrm{fi,s}}$ of the normal situation H_1. The chance node 2-Sprinklers describes the functioning of sprinklers provided that the decision (node 6) is positive; the probability of the active state of the sprinklers given the fire start is assumed to be very high, for example 0.999. The chance node 3-Flashover has two states: the design situation H_3 (the fire design situation without flashover) and H_4 (the fire design situation with flashover when the fire is fully developed).

When sprinklers are installed, the flashover in a compartment of 250 m^2 has the positive state with the conditional probability 0.002; if sprinklers are not installed, then $\mathrm{P}\{H_4|H_2\} = 0.066$. It is assumed that with the probabilities equal to the squares of the above values the fire will flash over the whole building, thus the values 0.000004 and 0.0044 are considered for the chance node 3. The chance node 4-Protection (introduced for formal computational reasons) has identical states to the decision node 7-Protection. The chance node 5-Collapse represents the structural failure that is described by the probability distribution linked to three child nodes (1, 3, 4). This situation can hardly be modelled using a decision tree. Note that the probability of collapse in the case of fire but not flashover may be smaller than in a persistent situation, due to a lower imposed load. The implementation of Bayesian networks in case studies will be shown in Handbook 2.

4 CONCLUSIONS

Risk assessment plays an essential role in the risk management of infrastructure systems. This chapter provides the general formulation of risk depending on two parameters: hazard probability and associated consequences. Structures and infrastructure systems are exposed to a multitude of possible hazards. Methods and tools for the calculation of risk are provided and illustrated in examples and the following conclusions can be drawn:

a) FMEA and FMECA are used in the case of preliminary risk analysis, especially in industry, offshore analysis. On the other hand HAZOP has been widely used in the chemical industries for detailed failure.

b) The tree based methods are mainly used to find cut-sets leading to the undesired events. In fact event tree and fault tree have been widely used to quantify the probabilities of occurrence of accidents and other undesired events leading to the loss of life or economic losses in probabilistic risk assessment.

c) Probabilistic risk analysis methods represent a powerful tool for decision making such as inspection and maintenance planning, selection of appropriate safety measures. The methods will be illustrated in case studies dealing with existing infrastructure presented in Handbook 2.

REFERENCES

[1] Chryssanthopoulos M., Janssens V., Imam B. (2011) Modelling of failure consequences for robustness evaluation. In: IABSE–IASS Symposium: Taller, Longer, Lighter, London.

[2] International Council for Research and Innovation in Building and Construction, CIB Report 259 Rotterdam (2001).

[3] Schneider J. (1997) Introduction to safety and reliability of structures. *Structural Engineering*, IABSE, Zurich.

[4] Stewart M. G. and Melchers R. E. (1997) Probabilistic risk assessment of engineering system. Chapman & Hall, London.

[5] American Society of Safety Engineers, Risk Analysis Methodologies*, TECH 482/535 Class Note* (2005).

[6] Meyna A. (1985) Hanbuch der Sicherheitstechnik, Band 1 C. Hanser Verlag, München.

[7] Janssens V., O'Dwyer D. and Chryssanthopoulos M. (2011) Building failure consequences. In: Proceedings of Final Conference COST Action TU0601 Robustness of Structures, Prague.

[8] Federal Emergency Management Agency (FEMA), HAZUS-MH MR3 Technical Manual, Washington D.C. [available at http://www.fema.gov/plan/prevent/hazus/, accessed on April 30th, 2011].

[9] LessLoss, Earthquake disaster scenario prediction and loss modelling for urban areas, Less Loss Report No. 2007/07, SPENCER, IUSS Press, Pavia (2007).

[10] Faber M. H., Kubler O., Fontana M. and Knobloch M. (2004) Failure consequences and reliability acceptance for exceptional building structures. IBK/ETH report, Zürich.

[11] Xie F. and Levinson D. (January 2009) Evaluating the effects of I-35W Bridge collapse on road-users in the Twin Cities Metropolitan Region, 88th Transportation Research Board Conf., Washington D.C.

[12] Van Essen H. P. et al. (2004) Marginal costs of infrastructure use – Towards a simplified approach, Final Report, CE-publications, Delft [available from *www.ce.nl* , accessed on April 30th, 2011].

[13] Rail safety and standards board (RSSB), Proposals for the Weighting of Major and Minor Injuries, London (2008).

[14] Munich Re, Topics Geo-Natural Catastrophes 2010, Munich (2011).

[15] Coburn A. R. and Spence R. (1992) Factors determining human casualty levels in earthquakes: Mortality prediction in building collapse, Proc. 10th World Conf. Earthquake Eng., Madrid; pp. 5989-5994.

[16] NS 5814, Requirements for risk analysis (1991).

[17] CAN/CSA-Q634-91. Risk analysis requirements and guidelines (1991).

[18] ISO 2394. General principles on reliability for structures (1998).

[19] ISO/IEC Guide 7, Risk management – Vocabulary - Guidelines for use in standards (2002).

[20] ISO/IEC Guide 51, Safety aspects – Guidelines for their inclusion in standards (1999).

[21] ISO 9000, Quality management systems – Fundamentals and vocabulary (2000).

[22] Jensen F.V. (1996) Introduction to Bayesian networks, Aalborg University, Denmark.

[23] Holický M. and Šajtar L. (2005) Risk assessment of road tunnels based on Bayesian network, Advances in Safety and Reliability, ESREL 2005, Taylor & Francis Group, London, pp. 873-879.

CHAPTER 7: RISK ACCEPTANCE CRITERIA

Dimitris Diamantidis[1] and Raffaele Boccara[2]

[1]OTH Regensburg, Germany
[2]OTH Regensburg, Germany, and University of Pisa, Italy

Summary

As described in the previous chapter risk is the parameter which is a combination of the probability of failure and the consequences of failure. In general it is important to distinguish between various types of consequences of failure for example human losses, environmental damage and economic losses. This chapter deals with acceptance of risk in broad terms and in particular with respect to infrastructure systems and structures specifically. Methodologies for the derivation of risk acceptance are described. Guidelines providing target reliability values for structural components and systems are discussed. Risk acceptance criteria for existing structures are also reviewed.

1 INTRODUCTION

1.1 Background Documents

Structural codes traditionally have been concerned foremost with public preventing loss of life or injury. European codes [1], and [2] and international documents [3], [4], [6], [7] and [22] provide general principles and guidance for application of probabilistic methods to civil engineering structures and describe risk and reliability acceptance criteria. A useful summary of such acceptance criteria can be found in [5]. General aspects on risk acceptance are summarized in the introductory Chapter 1.

1.2 Scope

Risk is a major issue related to the future use of infrastructures. Their structural systems have in general a good record and failures are mainly due to natural disasters or human errors or a combination of both. Safety is a property of the structure that can be achieved or assured and is quantified usually by the probability of failure or the associated reliability index. A more complete parameter to approach safety is risk which is a combination of the likelihood and the consequences of structural failure, viewed in a context. Thereby it is important to distinguish between various types of consequences, i.e. human losses, environmental damage and economic losses as mentioned in Chapter 5. Codes and standards traditionally have been concerned foremost with public safety preventing loss of life or injury.

Risk acceptance criteria applicable to the management of infrastructure and its systems/components are presented in this chapter. Target values inherent in the codes and standards are reviewed. Thereby the difference in setting risk acceptance criteria for new and existing structures is highlighted. A “discount” in the safety requirements for existing

structures is namely unavoidable mainly due to economic constraints. The need for periodic review of risk acceptance criteria is emphasized.

2 RISK ACCEPTANCE CRITERIA

2.1 General Principles

The concept of risk acceptance criteria is well established in many industrial sectors. Comparative risk thresholds are established which allow a responsible organization to identify activities which impose an unacceptable level of risk on the participating individuals or society as a whole. Risk acceptance can be defined by two different approaches:

a) *Implicit criteria* often involve safety equivalence with other industrial sectors. This approach is very common because some industrial sectors developed quantitative risk criteria well before others, so there was the possibility to compare calculated risks on this basis. The criteria are thereby implied in prescriptive rules.

b) *Explicit criteria* or risk-based criteria are now applied in many industrial sectors, as they tend to provide either a quantitative decision tool to the regulator or a comparable requirement for the industry when dealing with the certification / approval of a particular structure or system.

Acceptable risk levels cannot be defined in an absolute sense. As it is mentioned above, each individual has his own perception of risk, or expressed in decision theoretical terms, his own "preferences". In order to define what is meant by "acceptable risk levels", a framework for risk acceptability shown in Figure 1(based on [9]) is adopted. It is well known, that some risks are so high that they are unacceptable. Therefore risks should be reduced to a level that is "as low as reasonably practicable" (ALARP) as it is discussed in Section 2.3. In principle, there is also a level of risk that is negligible and needs no further risk reduction effort.

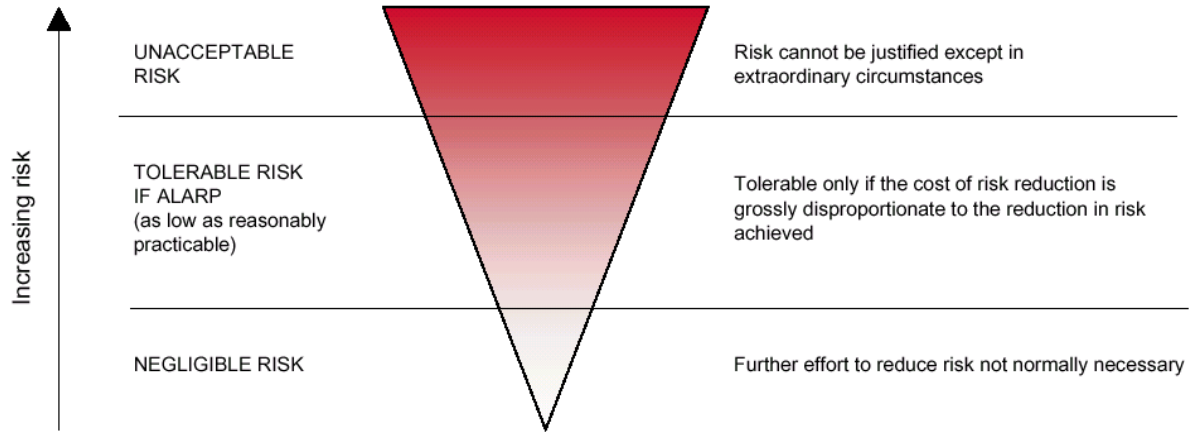


Figure 1 Framework for risk acceptability [9]

In order to define the criteria in Figure 1 in more tangible terms many different aspects need to be taken into account and it is important to incorporate them into a consistent

framework. These observations result in the conclusion that the decision to accept risk is not based on the absolute notion of one acceptable risk level but has some flexibility as the judgement depends on the cost/benefit ratio and the degree of voluntariness. In almost all studies two points of view are chosen:

(a) The point of view of the individual, who decides to undertake an activity weighing the risks against the direct and indirect personal benefits.

(b) The second is the point of view of the society, considering if an activity is acceptable in terms of the risk-benefit trade-off for the total population.

2.2 Individual Risk

Existing statistics about human fatality risk give guidance to set individual risk (probability of a person being killed per year). The average death rate from accidents and other adverse effects including fire and structural failure at work and at home ranges from 10^{-4} per year to more than $5 \cdot 10^{-4}$ per year with long term tendency to smaller values.

The acceptability of the risk from the individual point of view depends on the type of the activity, for example if the activity is voluntary or not. To have more information about the argument there are many scientific papers as [5], [6], [11], [16], [18] and [22].

The annual probability of being harmed describes the risk to an individual due to a hazardous situation. This probability is called the individual risk. With respect to fatality risks, the individual risk is the annual probability of being killed. The individual risk can also be defined as the frequency at which an individual may be expected to sustain a given level of harm from the realization of specified hazards. In general, individual risk is an issue in so-called daily risks, such as risks to workers.

Table 1 gives an overview of personal risks in developed countries. In column (b) the risk is presented as the probability per time unit of being killed when actually doing the activity mentioned in column (a). Such a frequency is called a fatal accident rate, FAR.

Following the original publication [16] the FAR is expressed as the number of fatalities per 100 million hours. Often the FAR is computed for a given activity during 10^8 hours of exposure, roughly equivalent to the total working hours of 1 000 individuals in a 40-year professional life. This is a useful measure in comparisons of the risk of various professions, but it is often misleading as in many cases we are only exposed to the considered activity for a small fraction of the time. The estimated part of time spent on the activity is mentioned in column (c). In column (d) the frequency is presented as the fatality probability per time unit averaged out over the hours doing and not doing the activity. It is expressed as an annual probability. The link between the two risk measures, of course, is the part of time spent on the activity. Formally the relation is given by:

$$p(A\ person\ being\ killed\ in\ one\ year) = \frac{FAR}{10\ 000} \cdot part\ of time \qquad (1)$$

The smallest component of the social acceptance of risk is the personal cost-benefit assessment by the individual. Attempts to model this appraisal procedure quantitatively are not feasible; therefore it is proposed to look at the pattern of preferences revealed in the accident statistics.

Table 1 Fatal accident rates [17]

(a)	(b)	(c)	(d)
Cause of Death	During activity [/10^8 hrs]	Part of time (average)	Annual probability [1/year]
Rock climbing	4000	0.005	1/500
Motorcycle accidents	300	0.01	1/3000
Skiing	130	0.01	1/8000
Workers in high rise building industry	70	0.2	1/700
Deep sea fishing	50	0.2	1/1000
Workers on offshore oil – and gas-rigs	20	0.2	1/2500
Disease average for 40-44 age group	17	1	1/600
Travel by air	15	0.01	1/70000
Travel by car	15	0.05	1/13000
Disease average for 30-40 age group	8	1	1/1200
Coal Mining	8	0.2	1/6000
Travel by train	5	0.05	1/40000
Construction industry	5	0.2	1/10000
Agriculture (employees)	4	0.2	1/12000
Accidents in the home	1.5	0.8	1/9000
Travel by local bus	1	0.05	1/200000
Chemical industry	1	0.2	1/50000
California earthquake	0.2	1	1/50000

The fact that the actual personal risk levels connected to various activities show statistical stability over the years and are approximately equal for the Western countries indicates a consistent pattern of preferences. The probability of losing one's life in normal daily activities such as driving a car or working in a factory appears to be one or two orders of magnitude lower than the overall probability of dying. Only a purely voluntary activity such as mountaineering entails a higher risk.

This observation of public tolerance of greater risks from voluntary than from involuntary activities, taking into account the direct benefit of the activity, may be used as a basis for decisions with regard to the personally acceptable probability of failure in the following way:

$$p_{fi} = \frac{\beta_i \cdot 10^{-4}}{p_{d|fi}} \tag{2a}$$

in which $pd|fi$ denotes the annual probability of being killed in the event of an accident (here structural failure). In this expression the policy factor βi varies with the degree of voluntariness with which an activity i is undertaken and with the benefit perceived. It ranges from 100, in the case of complete freedom of choice like mountaineering, to 0.01 in the case of an imposed risk without any perceived direct benefit (see Table 2 and [18]).

Table 2 Policy factor β_i as a function of voluntariness and benefit

β_i	Voluntariness	Direct benefit	Example
100	Completely voluntary	Direct benefit	Mountaineering
10	Voluntary	Direct benefit	Motor biking
1.0	Neutral	Direct benefit	Car driving
0.1	Involuntary	Some benefit	Factory
0.01	Involuntary	No benefit	LPG-station

Consequently general guidelines for assessment of the target reliabilities with respect to human safety are provided in *ISO 2394:1998* [3]. Based on the concept of individual risk, the annual target failure probability $p_{\mathrm{ft,hs}}$ depends on the conditional probability of an occupant fatality p_1, given the failure of the structure:

$$p_{\mathrm{ft,hs}}\ (\text{per year}) \leq 10^{-6}\ (\text{per year}) / p_1 \tag{2b}$$

Evaluation of available database of structural failures [24] revealed probability of fatality given the structural failure. It is emphasised that those values should be considered as an upper bound on p_1. Given the structural failure, the probability of fatality of an individual occupant (whom risk is being assessed) is obviously lower than the probability of at least one fatality of any occupant.

2.3 Societal Risk

To society as a whole or to a company or institution responsible for a specific activity, the total damage due to hazard is of prime interest as shown in the previous chapter. To comprehend this point of view the notion of collective risk R is introduced

$$R = \sum_{i=1}^{n} p_i C_i \tag{3}$$

where n is the number of all independent and mutually exclusive accident scenarios i, p_i is the probability of occurrence (per year) of scenario i and C_i are the consequences of scenario i. These consequences may for instance be measured in fatalities per year, monetary units or emission of a given substance.

In most practical studies the societal risk of an installation is given in the form of a numerical F-N curve [6]. In these diagrams N represents the number of fatalities and F the frequency of accidents with more than N fatalities, see also [9]. The curves show the relationship between the annual frequency F of accidents with N or more fatalities. Usually these curves are shown in a log-log plot with the frequency F in the ordinate axis (see for example Fig. 2). F-N curves were originally developed for nuclear hazards to illustrate thresholds that reflect societal aversion to multiple fatalities during a single catastrophic event. The graph is subdivided into four areas: *unacceptable risk*; tolerable risk that should be reduced further if practicable according to the *as low as reasonably practicable* (ALARP) principle; *broadly acceptable risk*; and a region of low probability but with the potential for >1 000 fatalities that requires intense scrutiny. From the perspective of potential loss of life from a hazard, development typically is approvable if it can be demonstrated that the risk falls in the ALARP or Broadly Acceptable regions on an F-N curve.

The recommendations of the F-N Curve can be represented also in a so-called risk-acceptability matrix as shown in an example related to road tunnels in Chapter 9. For this purpose qualitative hazard probability levels have been defined as well as hazard severity levels of accidental consequences. The hazard probability levels and the hazard severity levels can be combined to generate a risk classification matrix. The authority is usually responsible for defining the tolerability of the risk combinations contained within the risk classification matrix. This procedure can be seen as a different method for the same purpose of the F-N curve. The case study will be shown in Chapter 9 of this handbook.

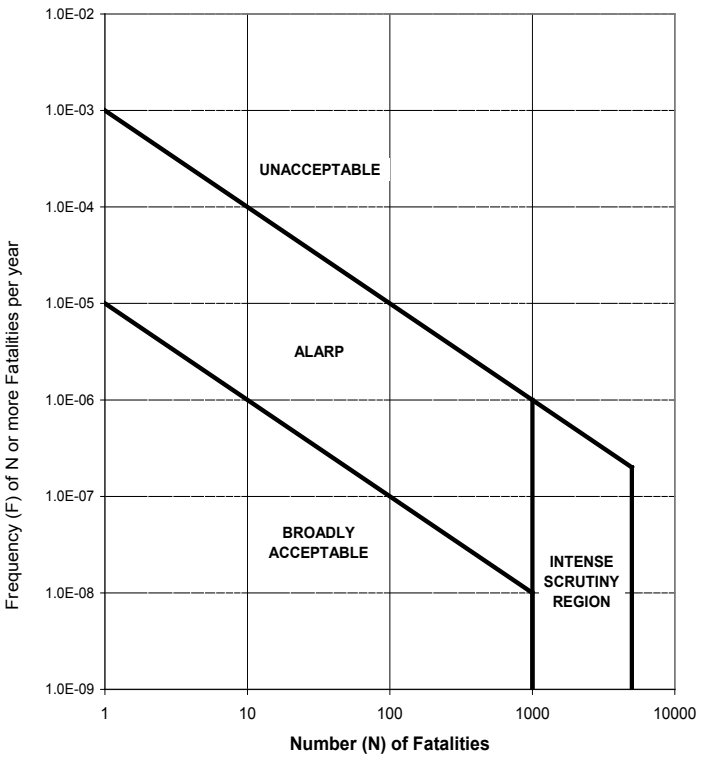


Figure 2 Risk acceptability example in terms of F-N curves

This procedure is especially useful in cases in which limited accidental data are available [10]. Performance objectives are needed in risk-based assessment of infrastructural systems. They should be defined at a global level, that is as an acceptable extent of collapse and acceptable other consequences of a hazard with a specified intensity and associated return period. An example of a performance matrix, expressed in terms of acceptable degrees of damage, which can form the basis for performance-based design with global performance objectives is shown in Table 3 for consequences classes as defined in Eurocode [1] and described in Table 4. Finally acceptance criteria in terms of F-N curves for various European countries are summarized in Figure 3.

Table 3 Performance matrix with acceptable degrees of damage for consequences classes CC1 to CC3 taken from [8]

Event Size	CC1	CC2	CC3
Very Large	Severe	High	Moderate
Large	High	Moderate	Mild
Medium	Moderate	Mild	Mild
Small	Mild	Mild	Mild

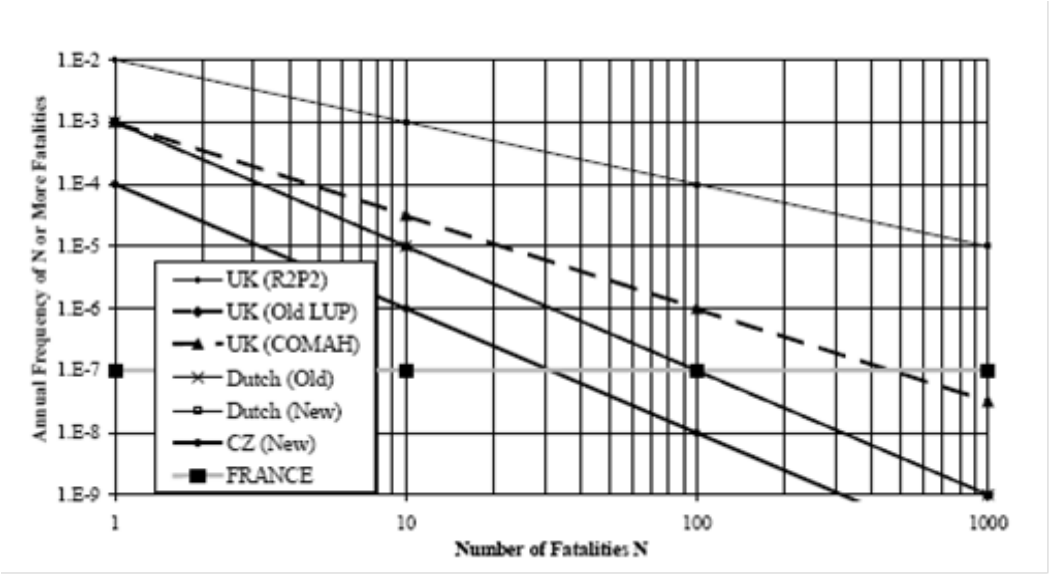


Figure 3 **F-N curves relating expected fatalities (N) from an accidental event and the annual frequency of occurrence (F) of event with no less than N fatalities [11]**

Table 4 Consequences classes in Eurocodes, EN 1990 [1]

Class1	Description	Examples
CC1	Low consequences for loss of human life, and economic, social or environmental consequences are small or negligible	Agricultural buildings, simple solar structures
CC2	Medium consequences for loss of human life, economic, social or environmental consequences are considerable	public buildings where failure consequences are medium
CC3	High consequences for loss of human life, or economic, social or environmental consequences are very great	Grandstands, public buildings where the consequences of failure are high (e.g. concert hall), important bridges

3 OPTIMIZATION BASED ON THE LQI PRINCIPLE

In cost-benefit-analysis, usually, all losses – including fatalities - are expressed in monetary units. For having a method that allows taking the loss of human lives accessible to calculations the Life Quality Index Approach (LQI) was introduced. This method facilitates the development of risk acceptance criteria [14]. The basic idea of the method is to model the preferences of a society quantitatively as a scalar valued Social Indicator comprised by a relationship between the Gross Domestic Product (GDP) pro capita g, the expected life at birth e and the proportion of life spend for earning a living w.

$$LQI = g^w \, e^{1-w} \tag{4}$$

where w in developed countries it is assumed ca. 1/8 [6].

Every risk measure will affect the value of the LQI. The consideration that any investment into life risk reduction should lead to an increase of the LQI leads to the following risk acceptance criteria [14].

$$\frac{\Delta g}{g} + \frac{\Delta e}{e}\frac{1-w}{w} \geq 0 \tag{5}$$

Using LQI in an optimization method, the optimum acceptable Implied Cost of Averting Fatality (ICAF) can then be calculated (8) considering that remaining life of an arbitrary individual equals half of the life expectancy at birth, the number of years saved by averting one fatality is given by (6).

$$\Delta e = \frac{e}{2} \tag{6}$$

$$|\Delta \mathrm{g}|_{max} = \frac{g}{e}\frac{1-w}{w}\Delta e = \frac{g}{2}\frac{1-w}{w} \tag{7}$$

$$ICAF = |\Delta \mathrm{g}|_{max}\,\Delta e = \frac{ge}{4}\frac{1-w}{w} \tag{8}$$

It should be noticed, that the value expressed by Equation (8) is not the value of one life. The human life is beyond price. ICAF is also not the amount of a possible monetary compensation for the relatives of the victim of the occurrence. ICAF is just the monetary value, which society should be willing to invest for saving one life according to its ethical principles.

ICAF values vary from 3 to 4 million Euros for developed countries to approximately 100 000 Euro for developing countries [15]. When conducting a cost-benefit-approach the ICAF value is utilized for evaluating investments; in [10] an expression for the evaluation of each individual safety measure was given as:

$$\frac{(C_{lk}\,\delta(T))}{T} + C_{Ak} < ICAF\; dR_{Hk} + dR_{Ck} \tag{9}$$

where

C_{lk}	investment cost;
C_{Ak}	annual maintenance/operation costs;
T	desired lifetime of measure;
dR_{Hk}	risk reduction due to measure k related to human risk;
dR_{Ck}	risk reduction due to measure k related to economic risk;
δ(T)	discount rate.

When the inequity is satisfied it means that the safety measure is beneficial. Note that the parameters entering the inequity are associated to significant variability and in many cases the risk reduction effect cannot be easily quantified. Probabilistic risk optimization based on the comparison of societal and economic consequences may provide valuable background information for a rational decision concerning effective safety measures applied to infrastructure.

The concept of Life Quality Index LQI and derived notions of the Societal Willingness To Pay SWTP and Societal Value of Statistical Life SVSL are further discussed in case studies. They seem to provide an effective and powerful tool to balance societal and economic aspects.

The acceptability limits (or targets) for probabilities of failure should in summary take into account, implicitly or explicitly, for probabilities of failure, potential losses, amount of investments necessity to improve reliability, and possibly combination of all these factors. The next section summarizes how these aspects have been taken into account in codes.

4 TARGET RELIABILITY

4.1 Target Reliability in Codes

The choice of the target reliability is an important first step for the calibration of structural design codes as well as during the probabilistic design of structures outside the code envelope. The risk acceptance criteria are introduced in pre-codification documents and structural codes as [1], [3] and [4] that will be analyzed in this paragraph.

In the Eurocodes [1], this information is given in terms of target and acceptable failure probabilities and associated reliability indices and used to obtain safety factors for design purpose. The values reflect the possible failure consequences classified similarly as in Table 7; the reference time period, usually 1-year, is used; the target values are valid for component failure and are shown in Table 5. They depend on a specific reference period (1 year and 50 year), without any explicit link to the design working life T_d. It's important to note that the β values given by [1], β_a and β_d, for each reliability class correspond to the same reliability level. Practical application of these values depends on the time period T_a considered in the design, which may be connected with available statistical information concerning time variant actions (e.g. wind, earthquake etc.). For example if for reliability calculations wind action statistical values are associated to a reference period of one year the β_a target value should be compared to the computed one. If however statistical values associated to a reference period of 50 years are used –resulting to higher mean values and lower reliability index - the β_d target value should be compared to the computed one.

Table 5 Reliability classification in accordance with Eurocode [1]

Reliability Class	Consequences for loss of human life, economic, social and environmental	Reliability index β		Example of buildings and civil engineering works
		β_a for T_a=1 y	β_d for T_d=50 y	
RC3 - High	High	5.2	4.3	Bridges, Public buildings
RC2 - Normal	Medium	4.7	3.8	Residential and offices
RC1 - Low	Low	4.2	3.3	Agricultural buildings

The probabilistic model code [4] issued by JCSS provides also target reliabilities for different structural classes as shown in Table 6. The target reliabilities relate to ultimate limit states and to dominant component failures and have been derived based on monetary optimization studies. The target reliabilities in this case are given as a function of the costs of the risk reduction measure and the consequences in case of failure, both defined relative to the initial construction costs of the structure.

Table 6 Target reliabilities related to one year reference period and ultimate limit state according to JCSS [4], [22]

Relative cost of safety measure	Consequences of failure					
	Minor		Moderate		Large	
	β	p_f	β	p_f	β	p_f
Large (A)	3.1	$\sim 10^{-3}$	3.3	$\sim 5\ 10^{-4}$	3.7	$\sim 10^{-4}$
Normal (B)	3.7	$\sim 10^{-4}$	4.2	$\sim 10^{-5}$	4.4	$\sim 5\ 10^{-6}$
Small (C)	4.2	$\sim 10^{-5}$	4.4	$\sim 5\ 10^{-6}$	4.7	$\sim 10^{-6}$

Similar recommendations are given by ISO in [3] and also in [23] for a design working life T_d, called in ISO, 'life time', without specification of any particular value of T_d. Table 7 shows the reliability index differentiated by two factors: the relative costs of safety measures and consequences of failure. The new draft of *ISO 2394* [17] indicates that the target failure probabilities may be selected solely on the basis of an economic optimization only if no risk of loss of human lives is associated with structural failures. Otherwise the marginal life-saving costs principle applies. This may be elaborated using individual or societal risk criteria or the Life Quality Index (LQI) approach. The values corresponding to small consequences of failure are related to serviceability limit states.

Table 7 Target reliability index for design working life T_d given by ISO [3]

Relative cost of safety measure	Consequences of failure			
	Small	Some	Moderate	High
High	0	1.5	2.3	3.1
Moderate	1.3	2.3	3.1	3.8
Low	2.3	3.1	3.8	4.3

It appears that available documents do not provide direct and explicit guidance on how to take into account β values when the design working life time T_d is different, as in many design situations, from the one used to calculate these values. Lifetime acceptable failure probabilities can be approximately obtained according to JCSS [4] by multiplying the annual failure probability value by the working life T_d and by a factor c that accounts for the dependence of different failure events within the year. In many cases, as mentioned in [5] the annual failure events are independent and consequently $c \cong 1$. The variation with time is discussed in more detail in Chapter 5.

4.2 Existing Structures

The risk acceptance criteria for existing structures have been defined in terms of annual probability of failure or target reliability index in various sources. However in case of reassessment and possible strengthening of existing structures there seems to be an agreement that, due to economic reasons in many cases, it is appropriate to consider lower target values compared to new structures [5], [19]. This can be done by using one category lower values in the tables of Section 4.1. For example instead of using for a RC3 structure a target annual value of β=4.3 (see Table 5) a value of 3.8 may be used.

Other procedures for deriving target reliability levels for existing structures can be found in the literature. Minimum safety levels for the evaluation of existing bridges have been

for example developed and incorporated in the Canadian bridge code [20]. The proposed target reliability index β is given as:

$$\beta = 3.5 - (\Delta_C + \Delta_S + \Delta_I + \Delta_R) \geq 2.0 \tag{10}$$

where

- Δ_C adjustment factor for component behavior (0.0 for failure with without warning, 0.25 for failure with little or no warning but retention of post failure capacity, 0.5 for gradual failure with warning);
- Δ_S adjustment factor for system behavior (0.0 if element failure leads to total collapse, 0.25 if element failure probably does not lead to total collapse, 0.5 if element failure leads to local failure only);
- Δ_I adjustment factor for inspection level (-0.25 if component is not inspectable, 0.0 if component is regularly inspected, 0.25 if critical component is inspected by evaluator);
- Δ_R adjustment factor for risk category (0.0 for all traffic categories except supervised overload, 0.5 for supervised overload).

Finally other proposals for target failure probabilities can be found in the literature such as in [21]:

$$p_f = \frac{S \cdot T \cdot A \cdot C_F}{N \cdot W} 10^{-4} \tag{11}$$

where

- T residual service life;
- N number of lives put to danger;
- S social criterion factor (Preservation value);
- A activity factor;
- C_F economical factor (consequences of failure);
- W warning factor.

Equation (11) includes social criteria in a simplified form. The warning factor corresponds to the likelihood that, given failure or recognition of approaching failure, a person at risk will be killed. The economical factor reflects cost-benefit considerations. The proposed Equations (10) and (11) have a similar background and dot not account directly for a lower acceptable failure probability of an existing structure compared to a new one due to higher marginal cost to increase the safety of an existing structure.

5 CONCLUSIONS

Several concepts of determining the acceptable level of risk have been presented herein. In assessing the safety of a system risk acceptability criteria are applied in order to approve its safety. The following concluding remarks are drawn based on the aforementioned considerations:

- Risk acceptance criteria are affected by several factors such as social, technical, administrative, political, legal, economic etc.
- Two basic types of risk criteria are distinguished: individual and societal.

– Risk acceptance criteria are provided in standards and codes as a function of the associated consequences.
– Target reliability levels are in fact defined in many standards reflecting the acceptable reliability of the subject structure/structural member.
– The matter of acceptable risk needs to be examined periodically, to accommodate new information, conditions and regulations.
– Relaxed acceptance risk criteria maybe selected for existing structures due to economic reasons.

REFERENCES

[1] EN 1990 Eurocode – Basis of structural design, European Committee for Standardization (2002).
[2] EN 1991 Eurocode 1: Actions on structures – Part 1-1: General actions – Densities, self-weight, imposed loads for buildings, European Committee for Standardization (2002).
[3] ISO 2394, General principals on reliability for structures (1998).
[4] Joint Committee on Structural Safety (JCSS) Probabilistic model code, Parts 1 to 4, Basis of design, Load and resistance models, JCSS (2001-2002).
[5] Diamantidis D., Vrouwenvelder T. (2012) Acceptance criteria. In: Innovative methods for the assessment of existing structures, CTU in Prague, pp.110-121.
[6] CIB Report: Risk assessment and risk communication in civil engineering, Second draft October 2000, CIB March 2001.
[7] ISO 13824: Bases for design of structures – General principles on risk assessment of systems involving structures (2009).
[8] Diamantidis D., Vogel T. (2011) Designing for robustness. Structural robustness design for practicing engineers. COST TU061, T.D. Gerard Canisius, revised version (September 2011), http://www.cost-tu0601.ethz.ch.
[9] HSE, Reducing risks, protecting people, Discussion document, Health & Safety Executive, London (1999).
[10] Diamantidis D. et al. (2012) Performance assessment of road tunnels based on risk acceptance criteria. In: International Conference PLSE, Hongkong.
[11] Jasanoff S. (1998) The political science or risk perception. *Reliability engineering and system safety*, 59, Elsevier Science Limited, Northern Ireland, pp. 91-99.
[12] Baker J. W., Schubert M., Faber M. H. (2008) On the assessment of robustness, *Structural Safety*, 30; pp. 253-267.
[13] Faber M. H. et al. (2007) Principles of risk assessment of engineered systems. , *Applications of Statistics and Probability in Civil Engineering,* Taylor & Francis Group, 2007, London.
[14] Rackwitz R. (2002) Optimization and risk acceptability based on the Life Quality Index, *Structural Safety,* pp. 297-331.
[15] Holický M. (2012) Optimization of the target reliability for temporary structures, *Civil Engineering and Environmental System,* Taylor & Francis, U.K.
[16] Hambly E. C., Hambly E. A. (1994) Risk evaluation and realism. In: Proceedings ICE Civil Engineering, Volume 102; pp. 64-71.
[17] ISO 2394, General principles on reliability for structures, revised draft

(2014).
[18] Vrijling J. K., van Hengel W., Houben R. J. (1998) Acceptable risk as a basis for design, *Reliability Engineering and System Safety,* 59, Elsevier Science Limited, Northern Ireland, pp. 141-150.
[19] Diamantidis D., Bazzurro P. (March 2007) Target safety criteria for existing structures, Workshop on risk acceptance and risk communication, Stanford University, CA, USA.
[20] CSA S6-1990, Design of highway bridges: Supplement No. 1-Existing bridge evaluation, Canadian Standards Association, Ottawa, Ontario (1990).
[21] Schueremans L., van Gemert D. (2004) Assessing the safety of existing structures: Reliability based assessment framework, examples and applications. *Journal of Civil Engineering and Management*, Vol. X, No. 2, pp. 131-141.
[22] Joint Committee on Structural Safety (JCSS) Probabilistic Assessment of Existing Structures, RILEM publications, editor D. Diamantidis (2001).
[23] ISO 13822: Bases for design of structures – General principles on risk assessment of systems involving structures (2009).
[24] Eldukair Z. A. and Ayyub B. M. (1991) Analysis of recent U.S. structural and construction failures. *Journal of Performance of Constructed Facilities* 5(1), pp. 57-73.

CHAPTER 8: DECISION UNDER UNCERTAINTY

M. Holický[1]

[1]Klokner Institute, Czech Technical University in Prague, Czech Republic

Summary

The target reliability levels as a basis of decision under uncertainty recommended in national and international documents vary within a broad range as the reference to relevant costs and failure consequences is mentioned only very vaguely. In some documents the target reliability index β is indicated for one or two reference periods (1 year and 50 years) without providing any link to the design working life. This chapter attempts to clarify the relationship between the target reliability levels, construction costs, failure consequences, reference period, the design working life and the discount rate to be considered when assessing aging infrastructures. The theoretical study based on probabilistic optimization is supplemented by recommendations useful for code developers and practicing engineers. It appears that the optimal reliability level depends primarily on the construction costs, failure costs, and relative cost for improving structural safety, and less significantly on the discount rate and the time to failure. In design based on the partial factor method, it is recommended that the characteristic values of the basic variables be defined independently of the design working life and specified target reliability.

1 INTRODUCTION

1.1 Background Documents

Decision under uncertainty based on a target reliability level that is an acceptable probability of failure or other unfavourable event. However, the target reliability levels recommended in various national and international documents for new structures are inconsistent in terms of the values and the criteria according to which the appropriate values are to be selected. Almost no recommendations are available for temporary structures. In general, optimum reliability levels can be obtained by considering both the cost of the structure and the expected cost of failure over the design working life.

The design working life is understood as an assumed period of time for which a structure is to be used for its intended purpose without any major repair work being necessary. Indicative values of design working life (10 to 100 years for different types of new structures) are given in EN 1990 (2002) as discussed in [1] and [2]. Recommended values of reliability indexes are given for two reference periods, 1 year and 50 years, without any explicit link to the design working life that generally differs from the reference period, while no specific indicative values are available for temporary structures.

It should be emphasized that the reference period is understood as a chosen period of time used as a basis for statistically assessing the time variant basic random variables, and the corresponding probability of failure. The concept of reference period is therefore fundamentally different from the concept of design working life. Confusion is often caused when the difference between these two concepts is not recognized.

It should be recognized that the couple of β values (for 1 year and 50 years) given in EN 1990 (2002) [1] for each reliability class correspond to the same reliability level (see also and discussion in previous Chapter 6 and in documents [3], [4] and [5]). Practical application of these values, however, depends on the time period considered in the verification, which may be linked to available probabilistic information concerning time variant basic variables (imposed load, wind, earthquake, etc.). It should be noted that the reference period of 50 years is also accepted as the design working life for common structures (see the discussion by Diamantidis (2009) [1]).

For example, considering a structure of reliability class 2 having a design working life of 50 years, the reliability index $\beta = 3.8$ should be used, provided that probabilistic models of basic variables are available for this period. The same reliability level is achieved when a reference period of 1 year, and a target of $\beta = 4.7$ are applied using the theoretical models for a reference period of one year. Thus, when designing a structural member, similar dimensions (reinforcement area) would be obtained considering $\beta = 4.7$ and basic variables related to 1 year or $\beta = 3.8$ and basic variables related to 50 years.

A more detailed recommendation concerning the target reliability is provided by ISO 2394 (1998) [6], where the target reliability indexes are indicated for the whole design working life without any restriction concerning its length, and are related not only to the consequences, but also to the relative costs of safety measures.

Similar recommendations are provided in the JCSS (2001) [7] Probabilistic Model Code based on the previous study of Rackwitz (2000) [8]. The recommended target reliability indexes are also related to both the consequences and to the relative costs of safety measures, though for a reference period of 1 year. The consequence classes in JCSS (2001) (similar to EN 1990, 2002) are linked to the ratio ρ defined as the ratio $(C_{str} + C_f)/ C_{str}$ of the total cost induced by a failure (cost of construction C_{str} plus direct failure costs C_f) to the construction cost C_{str} as follows:

- Class 1 Minor Consequences: ρ is less than approximately 2; risk to life, given a failure, is small to negligible and the economic consequences are small or negligible (e.g. agricultural structures, silos, masts);
- Class 2 Moderate Consequences: ρ is between 2 and 5; risk to life, given a failure, is medium and the economic consequences are considerable (e.g. office buildings, industrial buildings, apartment buildings);
- Class 3 Large Consequences: ρ is between 5 and 10; risk to life, given a failure, is high, and the economic consequences are significant (e.g. main bridges, theaters, hospitals, high rise buildings).

However, it is not quite clear what is meant in JCSS (2001) [7] by "the direct failure costs". This term indicates that there may be some other "indirect costs" that may affect the total expected cost. Here it is assumed that the failure costs C_f cover all additional direct and indirect costs (except the structural costs C_{str}) induced by the failure. The structural costs are considered separately and related to the costs needed for an improvement of safety (costs per unit of decision parameter C_1).

1.2 General Principles

The documents ISO 2394 (1998) [6] and JCSS (2001) [7] seem to recommend reliability indexes that are lower than those given in EN 1990 (2002) [1] even for the "small relative costs" of safety measures. It should be noted that EN 1990 (2002) [1] gives the reliability indexes for two reference periods (1 and 50 years) that may be accepted as the design working life for common structures (see also the discussion provided by Diamantidis

(2009) [1]). ISO 2394 (1998) [6] recommends indexes for "life-time, examples", thus related to the design working life, without any restrictions, while Probabilistic Model Code by JCSS (2001) [2] provides reliability indexes for the reference period of 1 year.

A clear link between the design working life and the target reliability level is not apparent from any of the above-mentioned documents. Thus, it is not clear and unambiguous which target reliability index should be used for a given design working life different from 50 years (say 10 years) that may be the case of aging infrastructure.

A new promising approach to specify the target reliability based on the concept of Life Quality Index (Fischer et al., 2012) [2] is considered in an ongoing revision of the International Standard ISO 2394 (1998) [6].

The basic aim of this chapter is to clarify the link between the design working life and the reliability index, and to provide guidance for specification of the target reliability level for a given design working life. The submitted theoretical study based on probabilistic optimization is supplemented by practical recommendations. This contribution is an extension of the previous study by Holický and Retief (2011) [5].

2 GENERAL PRINCIPLES OF PROBABILISTIC OPTIMIZATION

Probabilistic optimization is based on a fundamental form of the objective function (not covering monitoring and maintenance) expressed as the present value of the total expected cost $C_{tot}(x,q,n)$

$$C_{tot}(x,q,n) = C_{str} \sum_1^n P_f(x,i) + C_f \sum_1^n P_f(x,i) Q(q,i) + C_0 + x\, C_1 \quad (1)$$

Here x denotes the decision parameter of the optimization (a parameter of structural resistance), q is the annual discount rate (e.g. 0.03, an average long run value of the real annual discount rate in European countries), n is the number of years to the failure, which may differ from the design working life (specified usually as 50 or 100 years).

Further, $P_f(x,i)$ is the failure probability in year i, $Q(q,i)$ is the discount factor dependent on the annual discount rate q and the year number i, C_0 is the initial cost independent of the decision parameter x and failure (a quantity not affecting the optimization), and C_1 is the cost per unit of the decision parameter x (a structural parameter quantity affecting the structural resistance and optimization). The initial costs C_0 cost and cost of construction C_{str} are not discounted as they are paid at present.

Note that the design working life may generally differ from the time to failure denoted by the number of years n and considered here as an independent variable affecting the probability of failure.

It should be also noted that the maintenance and inspection costs are not explicitly included in the objective function (1) and should be considered in relevant cases, particularly when they are dependent on decision parameter x. These cost should be discounted as they will be paid during the assume design working life.

Maintenance and possible repair of the structure is not included, and these aspects are to be considered in further studies.

Assuming independent failure events in subsequent years, the annual probability of failure $P_f(x,i)$ in year i may be approximated by the geometric sequence

$$P_f(x,i) = p(x)\,(1 - p(x))^{i-1} \quad (2)$$

The initial annual probability of failure $p(x)$ is dependent on the decision parameter x. Note that annual failure probabilities can be assumed to be independent when failure

probabilities are chiefly influenced by time-variant loads (climatic actions, traffic loads, accidental loads). Then the failure probability $P_{fn}(x)$ during n years can be estimated by the sum of the sequence $P_f(x,i)$, that can be expressed as

$$P_{fn}(x,n) = 1 - (1 - p(x))^n \approx n\, p(x) \tag{3}$$

Note that the approximation indicated in Equation (3) is fully acceptable for small annual probabilities $p(x) < 10^{-3}$.

The discount factor of the present value of the expected future costs in year i is considered in the usual form as

$$Q(q,i) = 1 / (1+q)^I \tag{4}$$

Thus, the cost of malfunctioning C_f is discounted by the factor $Q(q,i)$ depending on the discount rate q and the point in time (year number defined as i) when the loss of structural utility occurs.

Considering Equations (2) and (4) the total costs $C_{tot}(x,q,n)$ described by Equation (1) may be written in a simplified form as

$$C_{tot}(x,q,n) = C_{str}\, n\, p(x) + C_f\, p(x)\, PQ(x,q,n) + C_0 + x\, C_1 \tag{5}$$

Here the total sum of expected malfunction costs during the period of n years is dependent on the product of the present value of malfunction cost C_f, the annual probability $p(x)$ and a sum of the geometric sequence having the quotient $(1- p(x))/(1+ q)$, denoted as the time factor $PQ(x,q,n)$:

$$PQ(x,q,n) = \frac{1 - \left[\dfrac{1-p(x)}{1+q}\right]^n}{1 - \left[\dfrac{1-p(x)}{1+q}\right]} \tag{6}$$

In general the total cost $C_{tot}(x,q,n)$ depends on the costs C_0, C_1, C_f, the annual probability of failure $p(x)$, the discount rate q, and the design working life n. Note that for small probabilities of failure $p(x)$ (for appropriate structural parameter x) and very small (zero) discount rate q, the time factor $PQ(x,q,n) \approx n$.

The necessary condition for the minimum of the total cost follows from (1) as

$$\frac{\partial C_{tot}(x,q,n)}{\partial x} = C_{str} \sum_{i=1}^{n} \left[\frac{\partial P_f(x,i)}{\partial x}\right]_{x=x_{opt}} + C_f \sum_{i=1}^{n} Q(q,i) \left[\frac{\partial P_f(x,i)}{\partial x}\right]_{x=x_{opt}} + C_1 = 0 \tag{7}$$

Equation (8) represents a general form of the necessary condition for the minimum of total cost $C_{tot}(x,q,n)$, the optimum value x_{opt} of the parameter x, and the optimum annual probability of failure $p_{opt} = p(x_{opt})$. The optimum probability for the total design working life $T_d = n$ years follows from Equation (7) as

$$\sum_{i=1}^{n} Q(q,i) \left[\frac{\partial P_f(x,i)}{\partial x}\right]_{x=x_{opt}} + (\rho - 1) \sum_{i=1}^{n} Q(q,i) \left[\frac{\partial P_f(x,i)}{\partial x}\right]_{x=x_{opt}} = -\frac{C_1}{C_f} \tag{8}$$

Equation (8) represents a general form of the necessary condition for the minimum of total cost $C_{tot}(x,q,n)$, the optimum value x_{opt} of the parameter x, and the optimum annual probability of failure $p_{opt} = p(x_{opt})$. The optimum probability for the total design working life $T_d = n$ years follows from Equation (3) as

$$P_{fn,\mathrm{opt}} = 1 - (1 - p_{\mathrm{opt}})^n \approx n\, p_{\mathrm{opt}} \tag{9}$$

The corresponding optimum reliability index $\beta_{\mathrm{opt}} = -\Phi^{-1}(P_{fn,\mathrm{opt}})$. These quantities are in general dependent on the cost ratio C_f/C_1, discount rate q, and the design working life n.

3 FAILURE PROBABILITY OF A GENERIC STRUCTURAL MEMBER

Consider a generic structural member described by the limit state function $Z(x)$ as

$$Z(x) = x\, f - (G+Q) \tag{10}$$

Here x denotes a deterministic structural parameter (e.g. the cross-section area), f the strength of the material, G the load effect due to permanent load and Q the load effect due to variable load. Theoretical models of the random quantities f, G and Q considered in the following example are given in Table 1 (adopted from JCSS (2001) and Holický (2009)).

Table 1 Theoretical models of the random variables *f*, *G* and *Q* (annual extremes)

Variables	Distribution	Mean	Standard deviation	Coef. of variation
F	Lognormal	100	10	0.10
G	Normal	35	3,5	0.10
Q	Gumbel	10	5	0.50

Considering the theoretical models given in Table 1, the reliability margin $Z(x)$ may be well approximated by the normal distribution $\Phi_{Z(x)}$ that provides sufficient accuracy. The annual failure probability $p(x)$ is then given as

$$p(x) = \Phi_{Z(x)}(Z(x) = 0) \tag{11}$$

In Equation (11) the normal distribution is evaluated for $Z(x) = 0$; then for x=1 and n=50 the probability $P_{fn}(1,50) \approx 6.7\ 10^{-5}$ and corresponding $\beta \approx 3.8$.

4 AN EXAMPLE

The following example illustrates the general principles, as well as a special case of probabilistic optimization. To simplify the analysis, the total costs $C_{\mathrm{tot}}(x,q,n)$ given by Equation (5) are transformed to the standardized form $\kappa_{\mathrm{tot}}(x,q,n)$ given as

$$\kappa_{\mathrm{tot}}(x,q,n) = \frac{C_{\mathrm{tot}}(x,q,n) - C_0}{C_1} = p(x) C_{\mathrm{str}} / C_1 \left[n + (\rho - 1) PQ(x,q,n) \right] + x \tag{12}$$

Here ρ denotes the cost ratio $(C_{\mathrm{str}} + C_f) / C_{\mathrm{str}}$ of the sum of structural cost and failure costs $(C_{\mathrm{str}} + C_f)$ to structural costs C_{str}.

The annual probability of failure $p(x)$ considered here for a general structural member is given by Equation (11). However, the following procedure may be applied for any relevant dependence of the failure probability $p(x)$ expressed as a function of a suitable structural parameter x.

In the example illustrated in Figure 1, it is assumed that the discount rate is q = 0.03, and the year number when the failure occurs is n = 50. Under these assumptions, Figure 1 shows the variation of the total standardized costs $\kappa_{\mathrm{tot}}(x,q,n)$ (given by Equation (12)), and the

optimum reliability index β_{opt}, with structural parameter x. The optimal values $x_{opt}(q,n)$ of the structural parameter x, given by Equation (8), are indicated by the dashed vertical lines. Indicated values of β = 2.6, 3.2, 3,5 correspond to annual rates 3,7, 4.2 and 4.4 recommended by JCSS (2001) for "normal relative costs of safety measures".

Note that the cost ratio $\rho = (C_{str} + C_f) / C_{str} = 1$ describes the extreme case when the failure costs C_f are negligible, $C_f \approx 0$, and the failure consequences are confined to the structural costs C_{str}. Assuming the conditions considered in Figure 1 (q =0.03, n = 50, C_{str}/C_1 = 100) the corresponding optimal reliability index $\beta_{opt} \approx 3.3$.

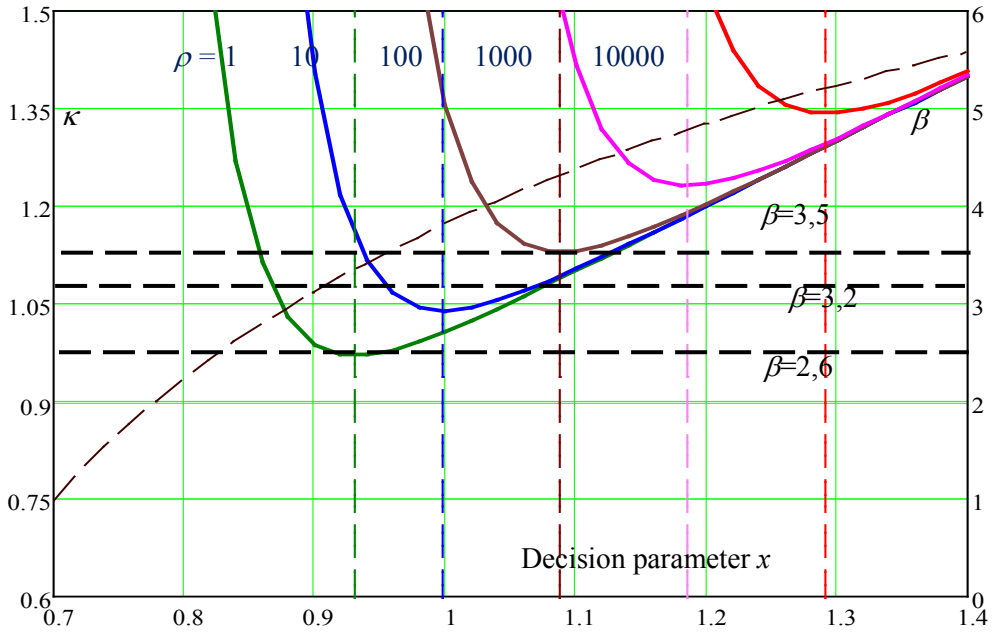


Figure 1 Variation of the total standardized cost $\kappa_{tot}(x,q,n)$ and the optimum reliability index β_{opt} with the decision parameter x for q =0.03, n = 50, C_{str}/C_1 = 100, and selected cost ratios ρ

5 THE OPTIMAL RELIABILITY INDEX

The optimal decision parameter x_{opt} given by Equation (8) is indicated in Figure 2. Figure 3 shows the variation of the optimal reliability index β_{opt} with the time to failure (number of years) n for the cost ratios $C_f/C_1 = \rho$ = 100 for selected discount rates q = 0.01, 0.03 and 0.05. The optimal reliability index β_{opt} is almost independent of q and n. Thus, as a conservative approximation, the target β can be determined considering a small discount rate q (for example 0.01), and a minimum number of years n (say 10).

The following Figure 4 shows the variation of the index β_{opt} (q,n,C_{str}/C_1,ρ) with the cost ratio ρ for q =0.03, n = 50, and selected ratios C_{str}/C_1 = 10, 100, 1000. Similarly, Figure 5 shows the variation of the index β_{opt} (q,n,C_{str}/C_1,ρ) with the cost ratio C_{str}/C_1, for q = 0.03, n = 50, and selected ratios, and for ρ = 10, 100, 1000.

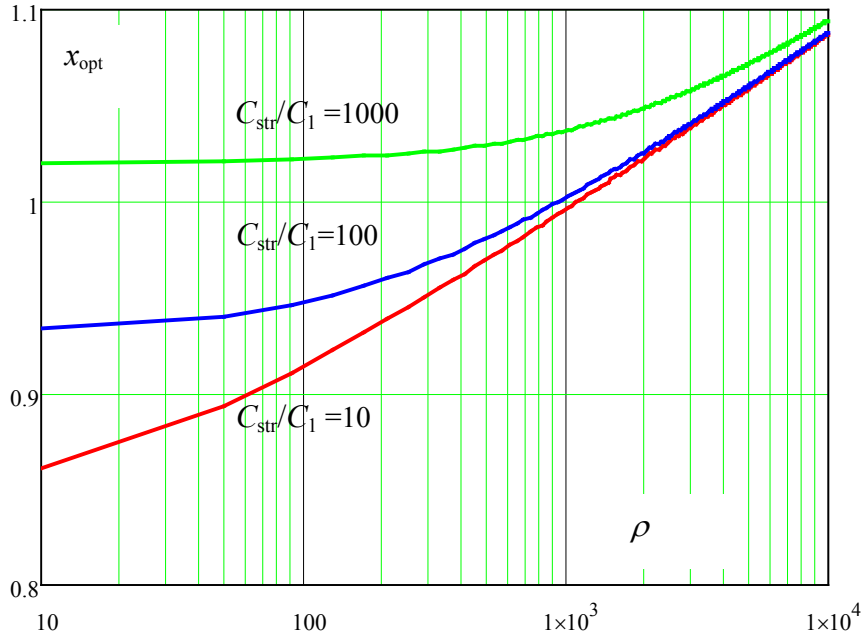


Figure 2 Variation of the index x_{opt} (q,n,C_{str}/C_1,ρ) with the number of years n for the cost ratios C_f/C_1 = ρ = 100 and discount rates q = 0.01, 0.03, 0.05

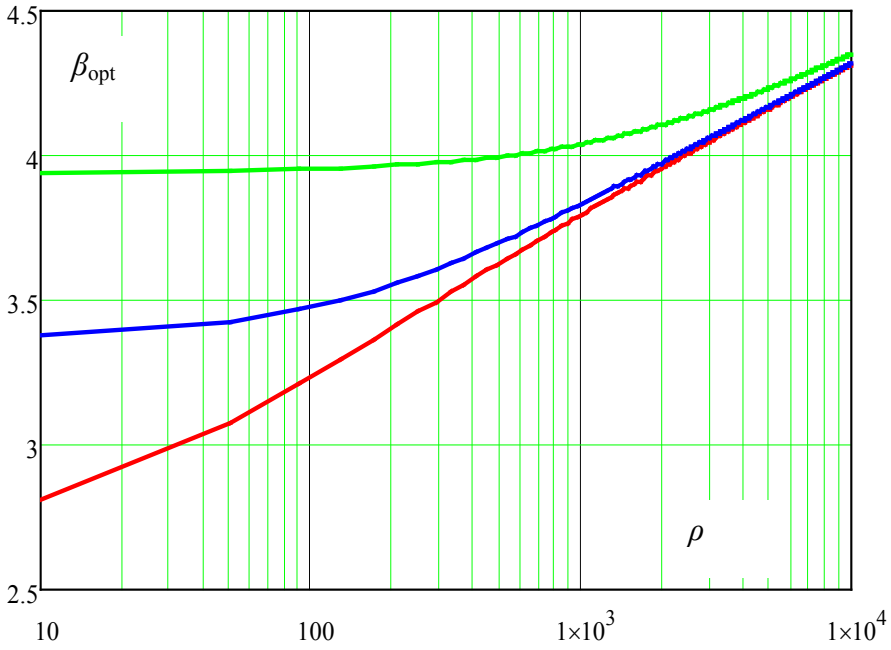


Figure 3 Variation of the index β_{opt} (q,n,C_{str}/C_1,ρ) with the number of years n for the cost ratios C_f/C_1 = ρ = 100 and selected discount rates q = 0.01, 0.03, 0.05

The variation of the index $\beta_{opt}(q,n,C_{str}/C_1,\rho)$ with the cost ratio C_{str}/C_1, shown in Figure 5, slightly differs from the variation with the ratio ρ, indicated in Figure 1. This difference is due to the effect of the discounting of the failure consequences C_f. Nevertheless, the optimal reliability level depends also on the costs C_1 (in addition to the costs C_{str} and C_f). That is why the recommendation provided by ISO and JCSS consider also "the relative costs

of safety measures". Nevertheless, the values of the ratio $\rho = (C_{str} + C_f)/C_{str}$ indicated within the interval from 2 to 10 seem to be low.

The optimization results shown in Figures 2, 3 and 4 clearly indicate that the optimal reliability index β_{opt} depends primarily on both the cost ratios C_{str}/C_1, and $\rho = (C_{str} + C_f)/C_{str}$. For the cost ratios $C_f/C_1 = \rho = 100$ (a reasonable estimate) the optimal reliability index β_{opt} is about 3.5, and the corresponding failure probability is $2.3\ 10^{-4}$.

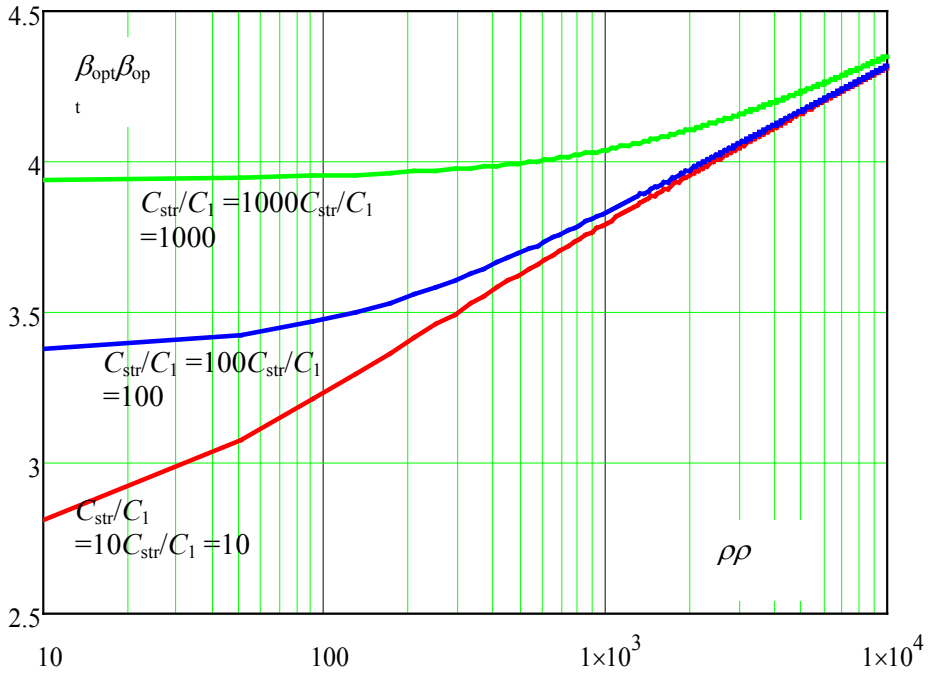


Figure 4 Variation of the index β_{opt} (q,n,C_{str}/C_1,ρ) with the cost ratio $\rho = (C_{str} + C_f)/C_{str}$ for q =0.03, n = 50, and selected ratios C_{str}/C_1 = 10, 100, 1000

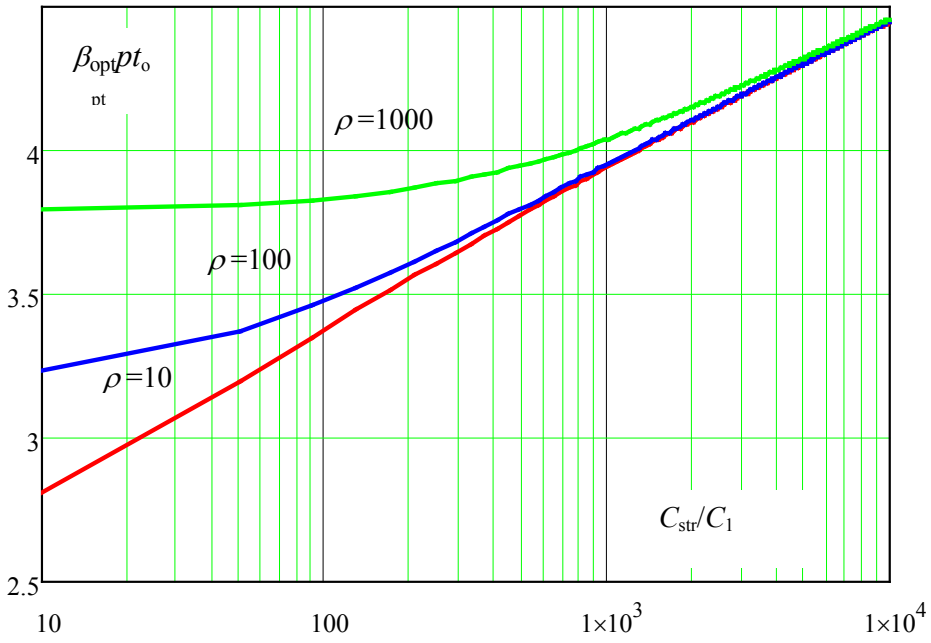


Figure 5 Variation of the index β_{opt} (q,n,C_{str}/C_1,ρ) with the cost ratio C_{str}/C_1 for q =0.03, n = 50, and selected ratios $\rho = (C_{str} + C_f)/C_{str}$ = 10, 100, 1000

It is interesting to mention that the actual reliability level observed in common design using the partial factor method is quite conservative (Holický and Schneider (2001)), and often corresponds to the index β above 5.

6 CONCLUSIONS AND RECOMMENDATIONS

The available documents and codes for structural design do not provide a clear link between the design working life and the target reliability level, and no recommendations are offered to specify the target reliability index for temporary and long life structures.

It has been shown that the target reliability of structures can be derived from theoretical principles of probabilistic optimization considering the objective function as the total costs expressed as a sum of the initial costs C_0, the marginal costs $x\ C_1$ (where x denotes the decision parameter and C_1 the incremental cost of decision parameter x), and the failure consequences consisting of the construction costs C_{str} and failure costs C_f (the loss of structural utility at the time of failure), these being taken into account by the cost ratio $\rho = (C_{str}+C_f)\ /C_{str}$. The failure cost C_f is discounted considering an annual discount rate q and the time to failure (number of years) n. In such a way the total cost is affected by the discount rate q, and the number of years n.

An example of the probabilistic optimization of a generic structural member clearly shows (see Figure 1, 4 and 5) that the optimal reliability level, i.e. the reliability index β, depends primarily on:

- The structural costs C_{str}
- Failure costs (malfunctioning costs) C_f
- Costs for improving structural safety C_1

The discount rate q and the time to failure n (Figure 3) seem to be less significant.

REFERENCES

[1] Diamantidis D. (2009) Reliability differentiation. In: Guidebook 1, Load effects on buildings. CTU in Prague, Klokner Institute, pp. 48-61.

[2] Fischer K., Barnardo-Viljoen C. and Faber M. H. (2012) Deriving target reliabilities from the LQI, LQI Symposium in Kgs. Lyngby, Denmark.

[3] Holický M. and Schneider J. (2001) Structural design and reliability benchmark study In: Safety, risk and reliability – Trends in engineering c/o IABSE, ETH Zűrich, International Conference in Malta, pp. 929-938.

[4] Holický M. (2009) Reliability analysis for structural design, SUN MeDIA Stellenbosch, ZA.

[5] Holický M. and Retief J. (2011) Theoretical basis of the target reliability. In: 9th International Probabilistic Workshop. Braunschweig: Technische Universität, pp. 91-101.

[6] ISO 2394. General principles on reliability for structures, International Organization for Standardization, Geneva, Switzerland (1998).

[7] JCSS. Joint Committee for Structural Safety. Probabilistic Model Code, http://www.jcss.ethz.ch/, 2001.

[8] Rackwitz R. (2000) Optimization — the basis of code-making and reliability verification. *Structural Safety*, **22**(1), pp. 27-60.

CHAPTER 9: SIMPLIFIED RISK ASSESSMENT OF EXISTING TUNNELS

Dimitris Diamantidis[1], Milan Holický[2]

[1]OTH Regensburg, Germany
[2]Klokner Institute, Czech Technical University in Prague, Czech Republic

Summary

Tunnels represent a significant part of the transportation infrastructure. The risk and the robustness of existing road tunnels are discussed in this chapter based on experience with various industry projects. Actual standards such as the EU directive are reviewed. The importance of risk analysis is highlighted. Risk and reliability acceptance criteria based on the theoretical aspects of Chapters 5, 6 and 7 are presented. Finally a case study is provided demonstrating the applicability and the benefits from the risk analysis approach. In the next Chapter 10 the optimization procedure for the cost optimal selection of the safety measures is described and illustrated in examples.

1 INTRODUCTION

Tunnel structures usually represent complex technical systems that may be exposed to hazard situations leading to unfavourable events with serious consequences. Moreover the risk during tunnel operation has increased in recent years. This has been manifested in significant accidents occurred in the last decade, such as the fire in the Mont-Blanc tunnel in France (see Figure 1), the accident in the Tauern tunnel in Austria and the fire in the Gotthard tunnel in Switzerland. The repair costs in case of the Mont Blanc tunnel accident were approximately 90 million Euros and the total costs over 190 million Euros. The consequences in case of such a tunnel accident can reach catastrophic dimensions for human, environment and economy.

Accidental events become therefore more and more important for the design and operation of tunnels. The combination of mixed traffic and extreme length (up to 20km) has a significant impact on the inherent risk and robustness of the tunnel system. In addition other than excavated type of tunnels such as submerged or floating tunnels are subjected to external hazards with high consequences in case of accident. Consequently risk and robustness have a significant influence in the selection of design parameters of a tunnel such as tunnel configuration, tunnel cross section and safety devices including prevention and mitigation measures.

Minimum requirements for tunnels in the Trans-European Road Network are reviewed first in this contribution such as those provided in the Directive of the European Parliament and of the Council of the European Union 2004/54/EU [1] or the recommendations given by the World Road Association PIARC [2], [3]. Many of the existing road tunnels in Europe cannot fulfil the new safety criteria and need special attention. Simplified risk analysis methods and acceptance criteria are summarized herein and applied in order to evaluate the safety of existing road tunnels. The objective is to illustrate basic aspects related to risk based approaches. More refined procedures are shown in the next Chapter 10.

2 TUNNELS AND STANDARDS

2.1 Road Tunnels Overview

The road network world-wide has grown steadily over the last century, with a greater than 50% in-crease in traffic volumes in the last decade. The development of the network has sometimes been obstructed by mountainous regions which have, on occasion, required the use of underground infrastructure. In urban areas tunnels have been relied upon to relieve congestion in heavily developed areas and remove visible traffic from sensitive environs. The longest road tunnels in existence are in mountainous countries such as Norway, Italy, Japan, and Austria. The longest road tunnel is Loerdal in Norway with a length of 24.5 kilometres (see Figure 2). Several long road tunnels are in the construction or in the design phase. Also various immersed and floating tunnels are under design and construction as under the sea links between cities or parts of a city. Therefore risk studies are undertaken as a decision support for basic configuration aspects such as tunnel system, escape ways, infrastructural and mechanical safety measures etc.

Especially with respect to existing tunnels many efforts are undertaken in order to upgrade them to reach acceptable safety levels. Risk analysis methods are thereby used and safety measures are proposed based on cost-benefit considerations.

Figure 1 Fire in the Mont Blanc Tunnel (March 1999)

Figure 2 Loerdal Tunnel in Norway with safety areas

2.2 Standards and Tunnel Classification

To assist in this decision process a scheme for tunnel classification has been developed within the European Union based on traffic volume and length. It reflects results of recent investigations and it builds the basis for the design of safety measures in road tunnels [1].

According to that guideline tunnels are classified according to their traffic volume in vehicles per lane and their tunnel length. Both parameters have a significant influence on the risk of the tunnel. On the basis of this classification scheme different safety measures such as tunnel configuration, emergency walkways, distance of emergency exits, fire resistance of structures, ventilation, water supply, monitoring systems etc. are proposed as minimum satisfactory requirements according to each tunnel class. The assessment of risk and the evaluation of possible safety measures should be accomplished based on quantitative acceptance criteria. The selection of safety measures is in addition based on cost-benefit considerations.

The EU Directive is thereby aiming to achieve the following with respect to tunnel structures:

- Prevention of all critical events that may endanger the life of tunnel users, the environment the tunnel structure as well as the provision of protection in case of accident;
- Prevention of progressive collapse, i.e. the main parts of the tunnel where local failure can lead to catastrophic consequences should provide sufficient fire resistance;
- Implementation of safety measures according to the tunnel class;
- Performing of a risk analysis by considering all hazards to determine the level of risk.

Specific risk analysis methods are not described in the directive. Also quantitative risk acceptance criteria are not provided. The member states shall ensure that, at national level, a detailed and well-defined methodology, corresponding to the best available practices is used. The risk analyses shall be carried out by a body which is functionally independent from the tunnel manager i.e. from the public or private body responsible from the management of the tunnel. The recommended safety measures represent minimum requirements.

Risk analysis procedures and associated tools have been analytically described by the World Road Association PIARC [2], [3]. Thereby the worldwide application of risk analysis methodologies for road tunnels is summarized and the suitability to meet specific requirements is discussed. The findings of the PIARC reports highlight that the possibilities for the harmonization of the methods of risk analyses for road tunnels are limited. It is also stated that the results of a quantitative risk analysis must be interpreted as an order of magnitude and not as a precise number due to the various uncertainties involved in the analyses.

In Germany specific guidelines with respect to tunnel safety RABT [5] are widely used. They provide practical recommendations regarding the implementation of safety systems in relation to the tunnel configuration and the tunnel length. For fire protection the U.S. standards of NFPA [6] are often applied. A review of fire safety guidelines is provided in [7] tunnel fire safety issues are described in [8].

2.3 Tunnel Types and Hazards

The risk of a tunnel depends on the hazards which may occur. Hazards depend also on the (construction) type of a tunnel:

- Bored tunnel
- Low buried tunnel (cut and cover)
- Immersed tunnel
- Floating tunnel

The hazards can be classified into:

– External hazards such as external impact from sinking ships, anchors, blast
– Internal hazards such as fire, explosion or impact from vehicles.

These hazards result to accidental actions which may significantly affect the risk and the robustness of a tunnel structure. Thereby it is important to classify hazards according to their frequency (probability of occurrence) and to look at the possible consequences in case of occurrence in order to evaluate the performance of the global structure. This procedure is widely used in case of accidental actions including earthquake and builds the basis of the performance-based engineering. It has been discussed in many publications and has been recently reviewed in [4].

3 RISK APPRAISAL

3.1 Societal Risk

In many practical studies the societal risk of a project is given in the form of a numerical F-N-curve as discussed in Chapter 7. An F-N-curve (N represents the number of fatalities, F the frequency of accidents with more than N fatalities) shows the relationship between the annual frequency F of accidents with N or more fatalities. It expresses both the probability and the consequence associated with given activities. Such acceptability curves have been developed for various industrial fields including the chemical and the transportation industry, [9]. The ALARP recommendations can be represented in a so-called risk-matrix. For that purpose qualitative hazard probability levels have been defined as shown in Table 1 and hazard severity levels of accidental consequences as shown in Table 2. For the severity level human consequences are considered here. The hazard probability levels and the hazard severity levels can be combined to generate a risk classification matrix. The principle of the risk classification matrix is shown in Table 3. The classification scheme of Tables 1, 2 and 3 has been proposed for long tunnel projects as possible risk acceptance criteria with emphasis to the lower bound values, [4].

Table 1 Hazard probability levels

Class	Frequency	(Events per year)
A	Frequent	>10
B	Occasional	1-10
C	Remote	0.1-1
D	Improbable	0.01-0.1
E	Incredible	0.001-0.01

Table 2 Hazard severity levels

Class	Severity category	Human losses
1	Insignificant	-
2	Marginal	Injuries only
3	Critical	1
4	Severe	10
5	Catastrophic	100

Table 3 Risk acceptability matrix

	1	2	3	4	5
A	ALARP	NAL	NAL	NAL	NAL
B	ALARP	ALARP	NAL	NAL	NAL
C	AL	ALARP	ALARP	NAL	NAL
D	AL	AL	ALARP	ALARP	NAL
E	AL	AL	AL	ALARP	ALARP

Notes:
1-5: Hazard severity levels according to Table 2
A-E: Hazard probability levels according to Table 1
AL: Acceptable Level
ALARP: As Low As Reasonably Practicable (Level)

4 SIMPLIFIED RISK ANALYSIS AND VERIFICATION

The purpose of the present study is to analyse the risk of the existing tunnel of the tunnel in South Europe during the 3-year period of construction of the second tube. Safety and reliability are major issues in the design of a road tunnel. Recent accidents have shown that significant work remains to be done in relation to the optimisation of safety measures in tunnels, the development of an understanding of reliability of tunnel structures, and improvement of building materials. The tunnel characteristics are as follows:

Tunnel length: The length is approximately 1400 meters.

Cross section: The cross section of the tunnel is characterized by a max height of 7.0 m, a width of 3.9 m per lane and a walkway with a width of 0.9m on both sides of the tunnel.

Traffic volume: (Projection) The total traffic volume (bi-directional traffic) representative for the next two years is approximately 5 000 000 vehicles per year.

Type of traffic: A mixed traffic (cars, lorries) including possible transportation of dangerous goods is taken into account. The percentage of lorries is taken as 12% based on available statistics.

Velocity: The design velocity in the tunnel is 90 km/h.

Inclination: The inclination is 0.7% and it is in the range of acceptable values.

Internal accidental events were considered in the analyses and the existing safety measures were evaluated. Table 4 illustrates representative results of the evaluation of the accident probabilities based on recorded accident statistics and on literature data [12], [13] The aforementioned risk matrix approach presented before has been used in order to evaluate the risk due to the selected hazards. Since accidental frequencies and accidental consequences are associated to considerable uncertainties the matrix method represents a logical framework to assess the risk. The consequences have been computed by using a fault-tree analysis combined with evacuation models similar to the methods described for example in [14]. Sensitivity analyses have been also performed in order to assess the variation of the possible outcomes.

Table 4 Accident probabilities

Accident type	Annual rate	Probability (T=3years)
Generic accident	~1	0.950
Fatal accident	0.150	0.362
Car fire	0.088	0.232
Major lorry fire	0.002	0.006

Table 5 Risk evaluation results

Hazard	Probability Level	Severity Level	Acceptance
Generic accident	B	2	ALARP
Fatal accident	C	3	ALARP
Major Lorry fire	E	4	ALARP

The results of the obtained risk matrix are shown together with the risk acceptance criteria in Table 5. It can be concluded that the risk is in the ALARP range and can be tolerated for the three years period until the second tube is constructed. Especially if one considers the fact that it regards the risk of an existing tunnel. Some additional safety measures such as improved safety lighting, and telecommunication systems etc. have been also proposed based on safety management considerations and cost benefit aspects as shown in the next chapter. The optimal selection of safety measures is influenced by uncertainties such as:

– Effectiveness of safety measures
– Lifetime of safety measures
– Costs of safety measures
– Discount rate
– Consequences of accident
– Etc.

It appears necessary to perform sensitivity analyses in order to evaluate the impact of such uncertainties in the results and in final decisions. An optimization approach is demonstrated in the next chapter. Figure 3 illustrates safety systems in a tunnel. In a simplified approach the safety systems can be selected based on a scoring system.

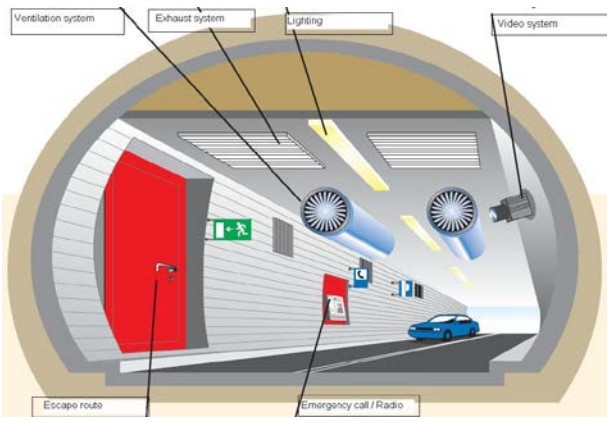


Figure 3 Safety measures in a tunnel

5 CONCLUSIONS

The treatment of risk and robustness aspects in existing long road tunnels is of significant importance. Risk analysis can be used as a decision tool for the verification of the risk level and for a preliminary selection of safety systems. Simplified methods have been demonstrated in this chapter. More refined tools are shown in the next chapter. The following conclusions can be drawn here:

1. Risk analysis is subject to uncertainties; therefore sensitivity analyses are recommended in order to examine the variation of the output.
2. A risk acceptance matrix offers a practical decision tool in order to evaluate the global performance of the tunnel system in case of accidental actions such as fire.
3. The selection of the tunnel configuration, of the escape ways geometry, of possible protective layers and of the materials plays a significant role in the risk and robustness of the tunnel.
4. Current standards provide minimum requirements regarding tunnel safety and associated safety measures.
5. The safety measures should be verified and optimized based on a risk analysis.
6. The global failure should be checked based on system considerations and on specified performance criteria.

REFERENCES

[1] European Union (EU). Directive 2004/54/EC, Official Journal of the European Union, 29th of April, 2004.

[2] PIARC Risk Analysis for Road Tunnels, PIARC Technical Committee C3.3: Working Group 2; Management of Road Tunnel Safety (2007).

[3] PIARC Integrated Approach to Road Tunnel Safety, PIARC Technical Committee. C3.3: Working Group 2; Management of Road Tunnel Safety (2007).

[4] Diamantidis D., Holický M. (2011) Risk and robustness of road tunnels, Proceedings ICASP Conference, Zürich, Switzerland.

[5] RABT, Richtlinien für die Ausstattung und den Betrieb von Straßentunneln (2006).

[6] National Fire Protection Association (NFPA), NFPA 502, Standards for Road Tunnels, Bridges, and other limited Access Highways (2004).

[7] Haack A. (2002) Current Safety Issues in Traffic Tunnels, *Tunnelling and Underground Space Technology*, 17, pp. 117-127.

[8] Hoj N. P. (2004) Guidelines for fire safety design compared fire safety features for road tunnels, First International Symposium Safe and Reliable Tunnels, Prague, Czech Republic.

[9] Diamantidis D., Zuccarelli F. & Westhäuser A. (2000) Safety of long railway tunnels, *Reliability Engineering and System Safety*, 67, pp. 135-145.

[10] ISO, ISO 2394, General principles on reliability for structures (1998).

[11] ISO, ISO 13822, Bases for design of structures -Assessment of existing structures, TC98/SC2 (2003).

[12] Bundesamt für Straßen. ASTRA Task Force, Schluss-bericht, Bern, Switzerland (2000).

[13] DATM (2003) www.gasandfire.de

[14] Vrouwenvelder A. C. W. M. & Krom A. H. M. (2004) Hazard and the Consequences for Tunnels Structures and Human Life, 1st International Symposium Safe and Reliable Tunnels in Prague, CUR, Gouda, the Netherlands.

CHAPTER 10: PROBABILISTIC RISK OPTIMIZATION OF ROAD TUNNELS

Milan Holický[1] and Dimitris Diamantidis[2]

[1]Klokner Institute, Czech Technical University in Prague, Czech Republic
[2]OTH Regensburg, Germany

Summary

Probabilistic methods of risk optimization are applied to identify the most effective safety measures considered in the design or assessment of existing of road tunnels. The total consequences of alternative tunnel arrangements are assessed using Bayesian networks supplemented by decision and utility nodes as described in Handbook 1. It is shown that the probabilistic optimization of societal and economic consequences may provide valuable information enabling a rational decision concerning effective safety measures in road tunnels. A general procedure is illustrated by the optimization of a number of escape routes using the concept of Life Quality Index. It appears that the discount rate and specified life time of a tunnel affect the total consequences and the optimum arrangements of the tunnel more significantly than the number of escape routes. Further investigation of relevant input data including societal and economic consequences of various hazard scenarios is needed.

1 INTRODUCTION

1.1 Background documents

Tunnel structures usually represent complex technical systems that may be exposed to hazard situations leading to unfavourable events with serious consequences. Minimum safety requirements for tunnels in the trans-European road network are provided in the Directive of the European Parliament and of the Council 2004/54/ES [1]. The Directive also gives recommendations concerning risk management, risk assessment and analysis.

Methods of risk assessment and analysis are more and more frequently applied in various technical systems ([2], [3]) including road tunnels ([4], [5], [6], [7], [8], [9]). This is a consequence of recent tragic events in various tunnels and of an increasing effort to take into account societal, economic and ecological consequences of unfavourable events. Available national and international documents ([10], [11], [12], [13], [14]) try to harmonise general methodical principles and terminology that can be also applied in the risk assessment of road tunnels.

1.2 General Principles

In the previous chapter a simplified risk matrix approach has been shown for the risk assessment of existing tunnels. This chapter, based, in particular, on previous studies ([6], [9] and [10]) and recent PIARC working documents, attempts to apply methods of probabilistic risk optimization using Bayesian networks supplemented by decision and utility nodes ([18]). It appears that Bayesian networks provide an extremely effective tool for investigating the safety of road tunnels. An essential extension of the presented study consists of the application of the concept of Life Quality Index (*LQI*) based on the recent studies ([19], [20], [21] and [22]).

2 PRINCIPLES OF RISK ASSESSMENT

Risk assessment is an essential part of the risk management, as indicated in Figure 1 in Chapter 1. The diagram in this figure indicates the main components of risk management and relevant basic terms. The whole process of risk management includes the risk assessment and risk control. The risk assessment further consists of the risk analysis and risk evaluation. The risk control consists of the risk decision and risk monitoring. Figure 1 in Chapter 1 also indicates another important part of the whole risk management, i.e. the risk communication concerning an unavoidable exchange of data between the risk assessment and risk control. Obviously, other types of the risk communication may be required in case of the risk management of complex tunnel systems.

The main components of the whole risk management are represented by the risk assessment and risk control. The risk control is outside the scope of this paper. Main steps of a general procedure of the risk assessment are shown in Figure 2 in Chapter 1. The flowchart is adopted from [12] and recent working materials of PIARC/C3.3/WG2. Two key steps of the risk analysis, i.e. the probability analysis and risk estimation based on probability and consequence analysis are shortly described below.

3 MODEL OF A TUNNEL

The main model of a road tunnel is indicated in Figure 3. The tunnel considered here is partly adopted from a recent study ([6]). It is assumed that the tunnel has the length of 4000 m and two traffic lanes in one direction. The traffic consists of heavy goods vehicles HGV, dangerous goods vehicles DGV and Cars.

The total traffic intensity in one direction is 20×10^6 vehicles per year (27 400 vehicles in one lane per day). The proportion of individual types of vehicles is assumed to be HGV:DGV:Cars = 0.15:0.01:0.84. The frequency of serious accidents assuming basic traffic conditions (that may be modified) is considered as 1×10^{-7} per one vehicle and one km [6], thus 8 accidents in the whole tunnel per year. The Bayesian networks used here need a number of other input data. Some of them are adopted from the study [6] (based on the event tree method), the others are estimated or specified using an expert judgement [9], [10].

Note that the main model includes three sub-models for heavy goods vehicles HGV, dangerous goods vehicles DGV and Cars. Various types of hazard scenarios are distinguished for each type of vehicles as they may cause different consequences. A detailed description of the main model, sub-models and required input data is outside the scope of this contribution.

4 PROBABILITY ANALYSIS

Probabilistic methods of risk analysis are based on the concept of conditional probabilities $P_{fi} = \mathrm{P}\{F|H_i\}$ of the event F given situation H_i occurs (see Handbook and also for example[2], [3]). In general, this probability can be found using statistical data, experience or a theoretical analysis of the situation H_i.

If the situation H_i occurs with the probability $\mathrm{P}(H_i)$ and the event F during the situation H_i occurs with the probability $\mathrm{P}(F|H_i)$, then the total probability P_F of the event F is given as

$$\mathrm{P_F} = \sum_i \mathrm{P}(F \mid H_i)\mathrm{P}(H_i) \tag{1}$$

Equation (1) makes possible to harmonize partial probabilities $\mathrm{P}(F|H_i)$ $\mathrm{P}(H_i)$ related to the situation H_i.

The purely probabilistic approach does not consider the possible consequences of the events F related to the situation H_i. Equation (1) can be, however, modified to take the consequences into account.

5 RISK ESTIMATION

A given situation H_i may lead to a set of events E_{ij} (for example, fully developed fire, explosion), which may have societal consequences R_{ij} or economic consequences C_{ij}. It is assumed that the consequences R_{ij} and C_{ij} are unambiguously assigned to events E_{ij}. If the consequences include only the societal component R_{ij}, then the total expected risk R is given as

$$R = \sum_{ij} R_{ij} \mathrm{P}(E_{ij} \mid H_i) \mathrm{P}(H_i) \tag{2}$$

If the consequences include only the economic component C_{ij}, then the total expected consequences C are given as

$$C = \sum_{ij} C_{ij} \mathrm{P}(E_{ij} \mid H_i) \mathrm{P}(H_i) \tag{3}$$

If criteria R_{d} and C_{d} are specified, then acceptable total consequences should satisfy the conditions

$$\mathrm{R} < \mathrm{R_d} \text{ and } \mathrm{C} < \mathrm{C_d} \tag{4}$$

that supplement the traditional probabilistic condition $P_{\mathrm{f}} < P_{\mathrm{fd}}$.

When the criteria are not satisfied, then it may be possible to apply a procedure of risk treatment, as indicated in Figure 2. For example, additional escape routes may be provided. Such measures might, however, require considerable costs, which should be balanced against benefits when deciding about the optimum measures.

6 RISK OPTIMIZATION

The total consequences $C_{\mathrm{tot}}(k,p,n)$ relevant to the construction and performance of a tunnel are generally expressed as a function of the decision parameter k (for example, the number k of escape routes), the discount rate p (commonly about $p \approx 0.03$) and the life time n (commonly $n = 100$ years). The decision parameter k usually represents a one-dimensional or multidimensional quantity significantly affecting tunnel safety.

The fundamental model of the total consequences may be written as a sum of partial consequences as

$$\mathrm{C_{tot}(k,p,n) = R(k,p,n) + C_0 + \Delta C(k)} \tag{5}$$

In Equation (5) $R(k,p,n)$ denotes the expected societal risk, which is dependent on the parameter k, discount rate p and life time n. C_0 denotes the initial construction cost independent of k, and $\Delta C(k)$ additional expenses dependent on k. Equation (5) represents, however, only a simplified model not reflecting all possible expenses including economic consequences of different unfavourable events and maintenance costs.

The societal risk $R(k,p,n)$ may be estimated using the following formulae

$$R(k,p,n) = N(k)\, R_1\, Q(p,n), \quad Q(p,n) = \frac{1-1/(1+p)^n}{1-1/(1+p)} \tag{6}$$

In Equation (6) $N(k)$ denotes the number of expected fatalities per one year (dependent on k), R_1 denotes the so called Societal Value of Statistical Life ($SVSL$) – an acceptable compensation cost for one fatality, and p the discount rate (commonly within the interval from 0 to 5 %). The quotient q of the geometric row is given by the fraction $q = 1/(1+p)$. The discount coefficient $Q(p,n)$ makes possible to express the actual expenses Z_1 during a considered life time n in current cost considered in (5). In other words, expenses Z_1 in a year i correspond to the current cost $R_1\, q^i$. The sum of the expenses during n years is given by the coefficient $Q(p,n)$.

The expected number of fatalities per year $N(k)$, shown in Figure 4, is derived using the model of the tunnel indicated in Figure 1. The key function $N(k)$ is dependent on a number of assumptions concerning consequences of adverse events and it may change when different hypotheses are accepted. In particular, some specific premises accepted for assessing the number of endangered persons and their ability to escape in case of fire (see studies [6], [9] and [10]) are essential. In case of route tunnels the consequence analysis appears to be the most difficult and sensitive part of the whole risk assessment.

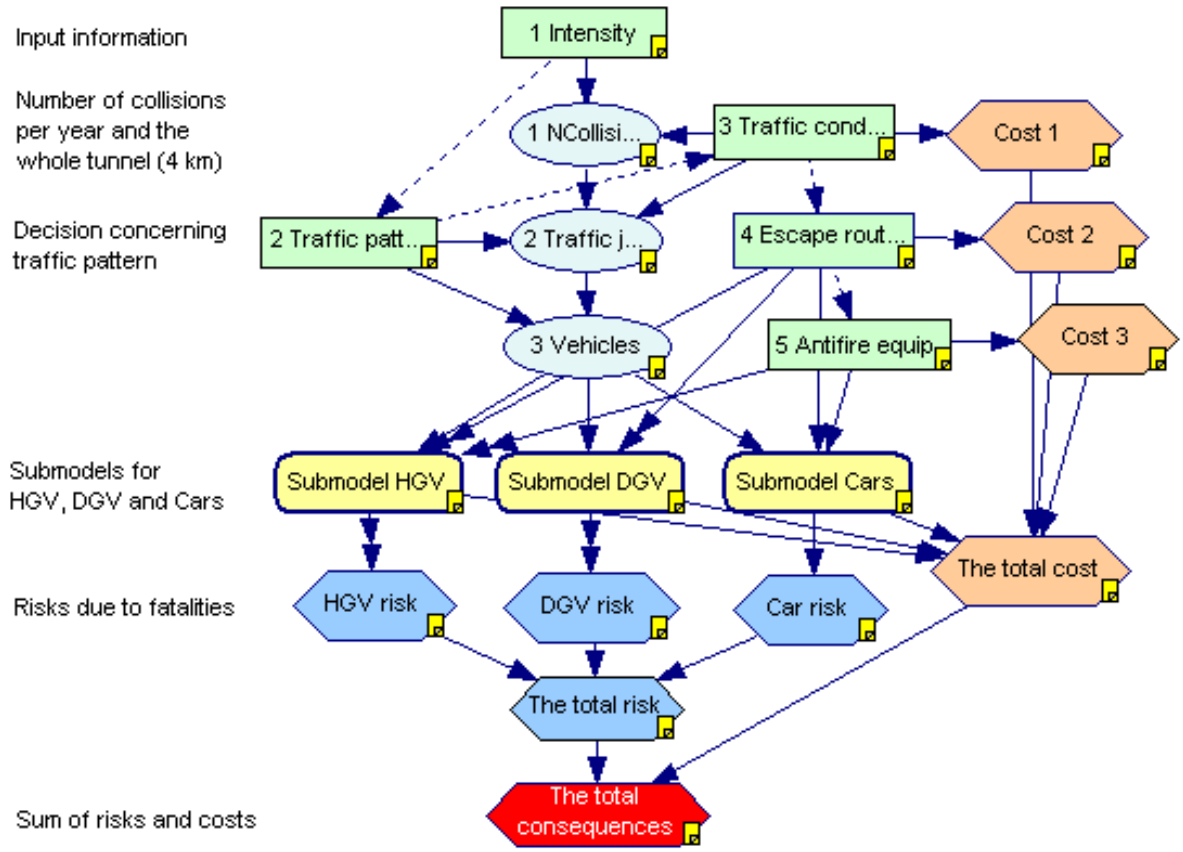


Figure 1 Main model of the tunnel

A necessary condition for the minimum of the total consequences (5) is given by the vanishing of the first derivative with respect to k written as

$$\frac{\partial N(k)}{\partial k} R_1 Q(p,n) = -\frac{\partial \Delta C(k)}{\partial k} \tag{7}$$

In some cases this condition may not lead to a practical solution, in particular when the discount rate p is small (the corresponding discount coefficient $Q(p,n)$ is large) and there is a limited number of escape routes k, which can not be arbitrarily increased.

7 STANDARDIZED CONSEQUENCES

The total consequences given by Equation (5) may be, in some cases, simplified to a dimensionless standardized form and the whole procedure of optimization may be generalized. Consider as an example the optimization of the number k of escape routes. It is assumed that the involved additional costs $\Delta C(k)$ due to k may be expressed as the product k C_1, where C_1 denotes the cost of one escape route. The cost C_1 is approximately equal to R_1 (assumed also in [6]). However in the following both variables are considered as generally independent quantities. Generally it may be considered that Equation (5) becomes

$$C_{tot}(k,p,n) = N(k)\, R_1\, Q(p,n) + C_0 + k\, C_1 \tag{8}$$

This function can be standardized as follows

$$\kappa(k,p,n) = \frac{C_{tot}(k,p,n) - C_0}{R_1} = N(k)Q(p,n) + k\zeta, \quad \zeta = \frac{C_1}{R_1}, \tag{9}$$

Both variables $C_{tot}(k,p,n)$ and $\kappa(k,p,n)$ are mutually dependent and have the minimum (if it exist) for the same number of escape routes k. A necessary condition for the extremes follows from (7) as

$$\frac{\partial N(k)}{\partial k} = -\frac{\zeta}{Q(p,n)} = -\zeta \frac{1 - 1/(1+p)}{1 - 1/(1+p)^n} \tag{10}$$

An advantage of the standardized consequences is the fact that it is independent of C_0. In addition to the number of escape routes k, discount rate p and assumed life time n, the standardized consequences $\kappa(k,p,n)$ are also dependent on the cost ratio $\zeta = C_1/R_1$, i.e. the ratio of the cost of one escape route C_1 and the Societal Value of Statistical Life $SVSL \approx R_1$, which can be based on the concept of Life Quality index, as indicated below.

8 LIFE QUALITY INDEX *LQI*

8.1 Principal Form

Based on the theory of socio-economics, the Life Quality Index (*LQI*) can be expressed in the following principal form ([19] and [20]):

$$L(g, l) = g^q l \tag{11}$$

where g denotes the part of *GDP* per capita, which is available for risk reduction purposes, l is the expected life and the parameter q is a measure of the trade-off between the resources available for consumption and the value of the time of healthy life. The parameter q may be assessed as:

$$q = \frac{1}{\beta} \frac{w}{1-w} \tag{12}$$

where β is a constant (around 0.7) taking into account the fact that only a part of the *GDP* is based on human labour (the other part is due to investments), w denotes a fraction of the life time devoted to work (around 0.15). Considering the above mentioned numerical values, the parameter q is about 0.25.

8.2 Societal Willingness To Pay *SWTP*

Assuming that any investment into life risk reduction should not lead to a decrease of *LQI*, the following risk acceptance criteria ([19]) may be obtained:

$$\mathrm{d}g + \frac{g}{q}\frac{\mathrm{d}l}{l} \geq 0 \tag{13}$$

Using this relationship, the societal willingness to invest annually into life saving activities (Societal Willingness To Pay) may be assessed as:

$$\mathrm{SWTP} = \mathrm{dg} = -\frac{g}{q}\frac{\mathrm{d}l}{l} \tag{14}$$

The annual investment dg indicated by Equation (14) denotes the annual cost the society is willing to pay to increase the life expectancy by dl.

The relative change in life expectancy dl/l may be considered as a change in mortality dμ as ([20])

$$\frac{\mathrm{d}l}{l} \approx D\,\mathrm{d}\mu = D\,\pi\,\mathrm{d}m \tag{15}$$

where dm is the rate of adverse events, D is a demographical economic constant corresponding to a given scheme for mortality reduction and π denotes the probability of dying given an adverse event. The constant D may be set approximately to 20 ([22]). It follows from Equations (14) and (15) that the annual investment into life safety denoted as dC (instead of *SWTP*) may be written as

$$dC = -\frac{g}{q} D\,\mathrm{d}\mu = -\frac{g}{q} D\,N_{PE}\;p\,\mathrm{d}m \tag{16}$$

where N_{PE} denotes the number of persons exposed to the adverse event. In case of road tunnels the change in mortality dμ was determined directly using methods of risk assessment as dN(k), where N(k) is indicated in Figure 2.

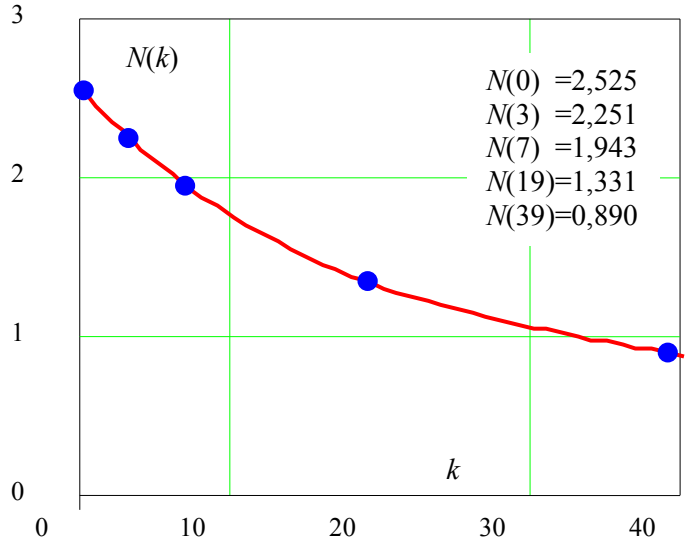


Figure 2 Variation of the number of fatalities N(k) with the number of escape routes k

If the failure rate depends on a decision parameter k (number of escape routes), then the acceptance criterion may be written as

$$dC(k) \geq -\frac{g}{q} D \, \mathrm{d}N(k) \tag{17}$$

Here d$C(k)$ denotes the amount that should be annually invested in life safety as a function of the number of escape routes.

As an example, consider g, the part of *GDP* per capita, which is available for risk reduction purposes, be the notional value g = 10 000 USD (corresponding to the demographical data valid for the Czech Republic). Assuming further the above mentioned parameter and $q = 0.25$ and $D = 20$, then the annual investment

$$dC(k) \geq -800\,000 \, \mathrm{d}N(k) \tag{18}$$

Variation of the annual investment d$C(k)$ with the number of escape routes given by Equation (18) is shown in Figure 3. Note that the Criterion (17) or (18) serves as the limitation of a possible investment decision and not as a condition for compensation in case of accident.

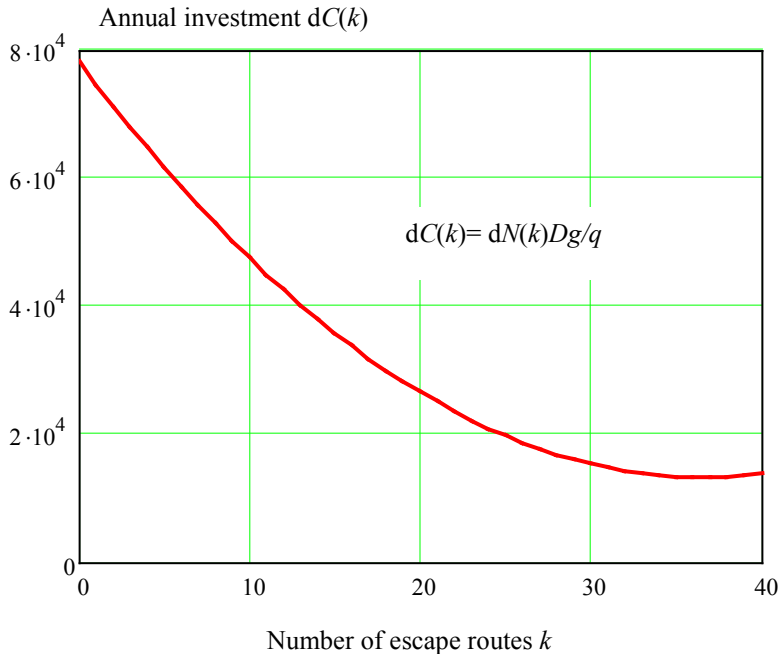


Figure 3 Variation of the annual investment d$C(k)$ with the number of escape routes k

The minimum annual investment d$C(k)$ decreases with the number of escape routes k from almost 80 000 USD to about 15 000 USD for 30 escape routes. It should be mentioned that the function $N(k)$ was obtained numerically (see Figure 2) and then approximated by a polynomial.

It should be noted that N (k) is the number of fatalities per year. Consequently, d$C(k)$ are yearly payments. But all the costs must be raised at a certain decision point. Then, privately financed and publicly financed projects should be distinguished. Financiation costs make the tunnel more expensive due to the annuity and this aspect could be included in Equation (8), but possibly with another discount rate (market rate).

It is interesting to note that Equation (17) may be used to specify an acceptable number of escape routes k_{acc}. Equation (17) may be written in the form

$$C_1 \geq -\frac{g}{q} D \, \mathrm{d}N(k) \tag{19}$$

Equation (19) can be used to specify the number of escape routes k complying with the acceptance Criterion (17). The minimum k satisfying Equation (19), called the acceptable number of escape routes k_{acc}, should be considered as a lower bound for any specified number k of escape routes. The tunnel is acceptable if $k > k_{acc}$. This condition should be verified also for an optimum number of escape routes k_{opt} (derived below). Thus, the tunnel is acceptable if $k_{opt} > k_{acc}$, when $k_{opt} < k_{acc,}$ then k_{acc} should be accepted.

8.3 Societal Value of Statistical Life *SVSL*

The *LQI* principles may be also used as a consistent basis for the assessment of compensation costs ([22]). Considering the principal form of *LQI* given by Equation (11), the Societal Value of a Statistical Life (*SVSL*) may be assessed as

$$\text{SVSL} = \frac{g}{q} E \tag{20}$$

In Equation (20) E denotes the so-called age-averaged discounted life expectancy ([22]). In this concept discounting means a societal discounting, roughly given by the natural discount rate, which equals the real economic growth rate. The age-averaging takes account of the fact that the age distribution of fatalities should mirror the age distribution of the population (equality principle). Considering an effective discounting of 3% per annum, the discount life expectancy for the Czech Republic is about 30 years and the corresponding *SVSL* is

$$\text{SVSL} = \frac{g}{q} E = 40\,000 \times 30 = 1\,200\,000 \text{ USD} \tag{21}$$

The *SVSL* is considered as an estimation of the compensation cost for one fatality R_1. However, in an optimisation concept R_1= *SVSL* may give a rather high value. The real compensation cost (carried by the social system of the state or an insurance) hardly exceeds the lost income in an event, approximately $g\,l/2$ =10000*70/2=350000 USD.

9 RISK OPTIMIZATION

The total consequences $\kappa(k,p,n)$ of the above described tunnel is indicated in Figures 4 to 8, showing the variation of the standardized total consequences $\kappa(k,p,n)$ given by Equation (9) with the number of escape routes k for selected discount rates p (up to 5 %) and a life time n (= 50 and 100 years). In Figures 4, 5 and 6 the cost ratio $\zeta = C_1/R_1 \approx 1$ is assumed; in Figures 7 and 8 the variation of the standardized consequences with the cost ratio ζ is illustrated.

Figure 4 shows the variation of the components of standardized total consequences $\kappa(k,p,n)$ with the number of escape routes k for the cost ratio $\zeta = 1$, a common value of the discount rate $p = 0.03$ and the assumed life time $n = 100$ years. Figure 5 shows the variation of the standardized total consequences $\kappa(k,p,n)$ with k for the cost ratio $\zeta = 1$, selected discount rates p and the life time $n = 50$ years only. Figure 6 shows similar curves as Figure 5, but for the expected life time $n = 100$ years (a common value). Note that the approximations $\zeta = C_1/R_1 \approx 1$, considered in Figures 4, 5 and 6, were assumed also in previous studies [6], [9] and [10].

It follows from Figures 5 and 6 that the minimum total consequences depend considerably on the discount rate p. For example, for the expected life time $n = 100$ years (Figure 6) and the discount rate $p = 0.05$ the optimum number of escape routes k is about 14, for the discount rate 0.01 the optimum number of escape routes k is more than 40 (most likely an unrealistic solution as the corresponding distance of escape routes would be less than 100 m). However, if the expected life time, considered in the optimization is 50 years only (Figure 7), then for the discount rate $p = 0.01$ the optimum number of escape routes k is about 30.

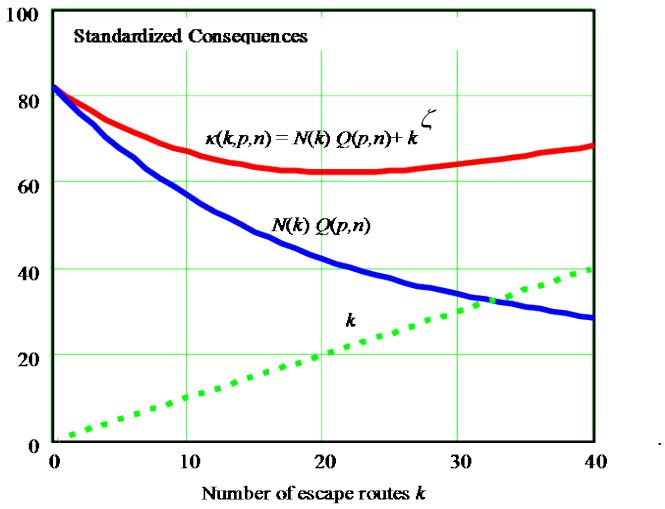


Figure 4 Variation of the components of standardized total consequences $\kappa(k,p,n)$ with the number of escape routes k for a cost ratio $\zeta = 1$, discount rate $p = 0.03$ and life time $n = 100$ years

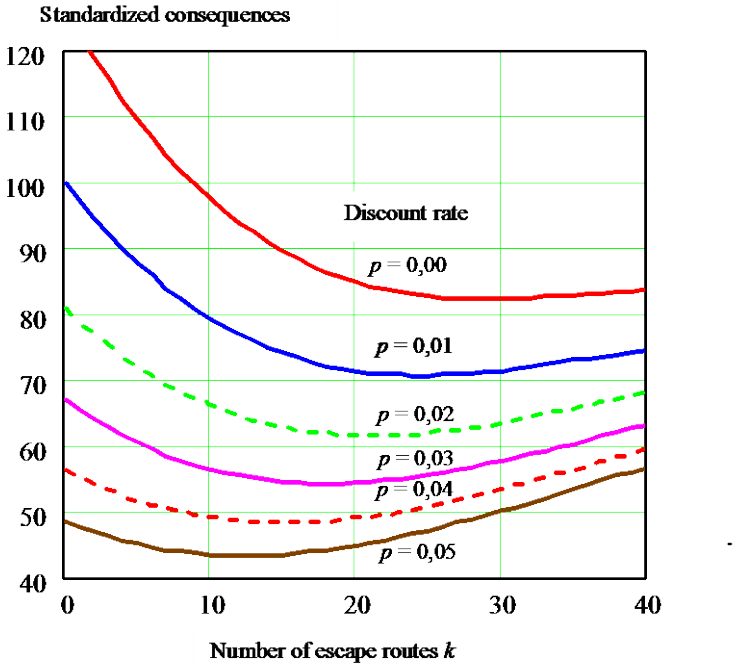


Figure 5 Variation of the standardized total consequences $\kappa(k,p,n)$ with the number of escape routes k for a cost ratio $\zeta = 1$, selected discount rates p and life time $n = 50$ years

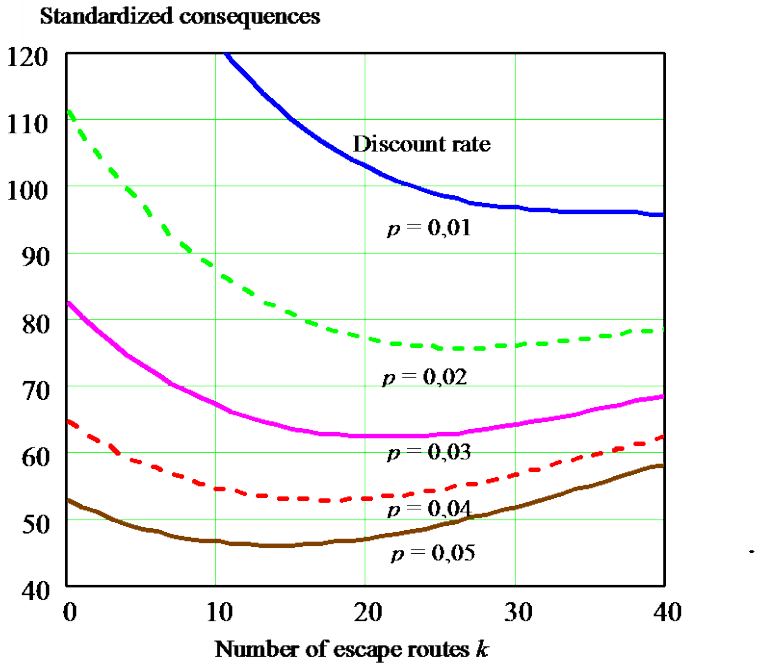


Figure 6 Variation of the standardized total consequences $\kappa(k,p,n)$ with the number of escape routes k for a cost ratio $\zeta = 1$, selected discount rates p and life time $n = 100$ years

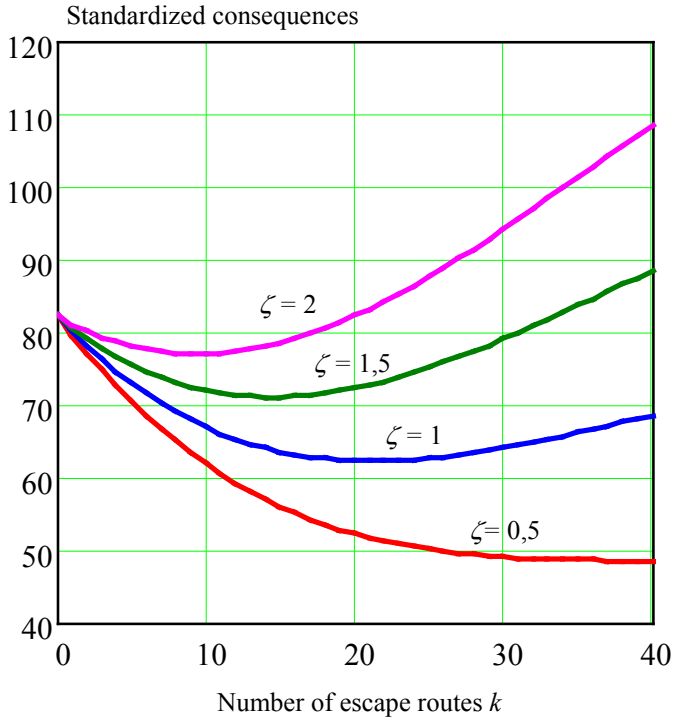


Figure 7 Variation of the standardized total consequences $\kappa(k,p,n)$ with the number of escape routes k for selected cost ratios ζ, a discount rates $p = 0.03$ and a life time $n = 100$ years

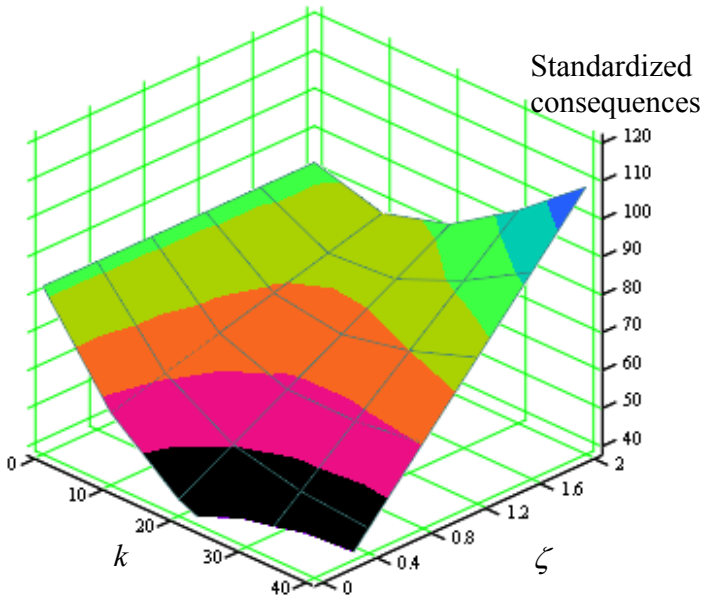


Figure 8 Variation of the standardized total consequences $\kappa(k,p,n)$ with the number of escape routes k and cost ratios ζ, for a discount rates $p = 0.03$ and life time $n = 100$ years

It appears that the cost ratio $\zeta = C_1/R_1$ of the cost of one escape route C_1 and the Societal Value of Statistical Life $SVSL \approx R_1$ may affect the total consequences and the optimum number of escape routes even more dramatically than the discount rate and the expected life time. Figure 7 shows the variation of the total standardized consequences with the number of escape routes k for selected cost ratios ζ (considered within the interval from 0.5 to 2), the discount rate $p = 0.03$ and life time $n = 100$ years. Variation of the standardized total consequences $\kappa(k,p,n)$ with both the number of escape routes k and cost ratios ζ, the discount rate $p = 0.03$ and the life time $n = 100$ years is illustrated in Figure 8.

It follows from Figures 7 and 8 that the optimum number of escape routes k increases with decreasing the cost ratio ζ (i.e. with decreasing the cost of one escape route C_1 or increasing the Societal Value of Statistical Life $SVSL \approx R_1$). For example, for $\zeta = 2$ ($C_1 = 2R_1$) the optimum k is about 9, for $\zeta = 1$ ($C_1 = R_1$) the optimum k is about 20 and for $\zeta = 0.5$ ($C_1 = 0.5\ R_1$) the optimum k is more than 40 (most likely an unrealistic solution as the corresponding distance of escape routes would be less than 100 m). Obviously, the correct specification of the cost of one escape route C_1 and the Societal Value of Statistical Life $SVSL \approx R_1$ is of the uttermost importance.

10 CONCLUSIONS

Probabilistic risk optimization based on the comparison of societal and economic consequences may provide valuable background information for a rational decision concerning effective safety measures applied to road tunnels. The concept of Life Quality Index *LQI* and derived notions of the Societal Willingness To Pay SWTP and Societal Value of Statistical Life *SVSL* seem to provide an effective and powerful tool to balance societal and economic aspects. It appears that the cost ratio of the cost of one escape route and *SVSL*, the

assumed life time and the discount rate may significantly affect the total consequences and the optimum number of escape routes.

However, as tunnels are commonly designed for a long period of time (100 years), probabilistic risk optimization may be questioned even though the assumed life time and discount rate are taken into account. The following conclusions should be, therefore, considered as qualitative (informative) statements only.

- Bayesian networks supplemented by decision and utility nodes seem to provide an effective tool for risk analysis and optimization.
- The requirement for the minimum total consequences covering both the societal and economic aspects may provide valuable information for the specification of the optimum number of escape routes.
- The total consequences are primarily affected by the cost ratio between the cost for one escape route and acceptable societal compensation cost *SVSL* and the discount rate, less significantly by the assumed life time and the number of escape routes.
- The optimum number of escape routes depends primarily on the ratio between the cost for one escape route and acceptable societal compensation cost *SVSL*, partly on the discount rate and assumed life time.

A correct specification of the discount rate and required life time is essential for making proper decisions. Further investigations of input data on conditional probabilities, describing individual hazard scenarios and models for their societal and economic consequences, are needed. In particular, consistent values for the costs of various safety measures and acceptable societal compensation costs *SVSL* based on Life Quality Index are required. The applied procedure can similarly be used for the optimization of safety measures of existing tunnels.

REFERENCES

[1] Directive 2004/54/EC of the European Parliament and of the Council of 29 April 2004 on minimum safety requirements for tunnels in the trans-European road network. Official Journal of the European Union L 201/56 of 7 June 2004.

[2] Stewart M. S. and Melchers R. E. (1997) Probabilistic risk assessment of engineering system, Chapman & Hall, London.

[3] Melchers R. E. (1999) Structural reliability analysis and prediction. John Wiley & Sons, Chichester.

[4] Worm E. W. and Hoeksma J. (1998) The Westerschelde Tunnel: Development and application of an integrated safety philosophy, Safety in Road and Rail Tunnels, 3rd International Conference organised by University of Dundee and ITC Ltd., Nice, France.

[5] Kruiskamp M. M., Weger D. and Hoeksma J. (2002) TUNprim: A Spreadsheet model for the calculating road tunnels risk, Tunnels Management International, Volume 5, Number 1, pp. 6–12.

[6] Vrouwenvelder A. C. W. M. and Krom A. H. M. (2004) Hazard and the Consequences for Tunnels Structures and Human Life, 1st International Symposium Safe and Reliable Tunnels in Prague, CUR, Gouda, The Netherlands.

[7] Knoflacher H. and Pfaffenbichler P. C. (2004) A comparative risk analysis for selected Austrian tunnels, 2nd International Conference Tunnel Safety and Ventilation, Graz.

[8] Ruffin E. &. Cassini P. and Knoflacher H. (2005) Transport of hazardous goods, Chapter 17 of Beard A and Carvel R. The Handbook of Tunnel Fire Safety. Thomas Telford Ltd., London.
[9] Holický M. and Šajtar L. (2005) Risk Assessment of Road Tunnels Based on Bayesian Network, Advances in Safety and Reliability, ESREL 2005, Taylor & Francis Group, London, pp. 873-879.
[10] Holický M. and Šajtar L. (2006) Probabilistic Risk Assessment and Optimization of Road Tunnels, Advances in Safety and Reliability, ESREL 2006, Taylor & Francis Group, London, pp. 2065-2071.
[11] NS 5814. Requirements for risk analysis (1991).
[12] CAN/CSA-Q634-91. Risk analysis requirements and guidelines (1991).
[13] ISO 2394. General principles on reliability for structures (1998).
[14] ISO/IEC Guide 73. Risk management – Vocabulary - Guidelines for use in standards (2002).
[15] ISO/IEC Guide 51. Safety aspects – Guidelines for their inclusion in standards (1999).
[16] ISO 9000. Quality management systems – Fundamentals and vocabulary (2000).
[17] Vrouwenvelder A., Holický M., Tanner C. P., Lovegrove D. R. and Canisius E. G. (2001) Risk assessment and risk communication in civil engineering, CIB Report Publication 259.
[18] Jensen Finn.V. (1996) Introduction to Bayesian networks. Aalborg University, Denmark.
[19] Rackwitz R. (2002) Optimization and Risk Acceptability Based on the Life Quality Index. Structural Safety, Vol. 24, pp. 297-332.
[20] Rackwitz R. (2005) The Philosophy Behind the Life Quality Index and Empirical Verifications. Memorandum to JCSS. Munich, Germany.
[21] Rackwitz R., Lentz A. and Faber M. H. (2005) Socio-economically sustainable civil engineering infrastructures by optimization. Structural Safety, 27(3), pp. 187-229.
[22] Rackwitz R., Faber M. H. and Schubert M. (2007) Prinzipien der Sicherheitsbewertung. Bundesamt für Strassenwesen ASTRA.

ANNEX A: EVALUATION OF DATA

Milan Holický[1]

[1]Klokner Institute, Czech Technical University in Prague, Czech Republic

A.1 General

The evaluation of statistical data representing a random sample taken from a particular population is frequently the first step in assessment of existing structures. The concept of a general population and the random samples taken from it is introduced and supplemented by the definition of commonly used sample characteristics. Emphasis is put on the moment characteristics that usually provide the initial background information for the specification of a theoretical model of population. Sample characteristics regularly used in engineering and science describe the location, dispersion, asymmetry and kurtosis of statistical data. The general rules and computational techniques used for determining sample characteristics of a single random sample, and also for the combination of two random samples, are illustrated by examples.

The concepts of population and random sample are extremely important for the appropriate interpretation of statistical data and their analysis. Population, or "the universe", is the totality of items under consideration. A population may be finite (N sampling units) or infinite. Rather than examining the entire group of N units a small part of the population, that is a sample of n units, may be examined instead. A precise definition regarding a population is often difficult to come by, but must be provided in order to interpret outcomes of statistical investigation correctly [1, 2]. An excellent description of the basic technique is given in [3, 4] and a short review is provided in [5]. The correct terminology and procedures are available in International Standards [6, 7, 8, 9].

A sample is one or more units taken from a population and is intended to provide information on that population. It may serve as a basis for decision-making about the population, or about the process which produced it. The term "random sample" refers to the samples that are taken from a population in such a way that all possible units have the same probability of being taken. The number of sampling units, called sample size n, may be considerably different. Commonly, samples are considered to be very small ($n < 10$), small ($n < 30$), large ($n > 30$) or very large ($n > 100$). Obviously, with increasing size the samples become more representative. However, the sampling procedure is equally important.

If a sample is representative of a population, important conclusions about it can often be inferred from an analysis of the sample. This phase of statistics is called inductive statistics, or statistical inference, and is covered in subsequent chapters. The phase of statistics that seeks only to describe and analyse a given sample is called descriptive, or deductive statistics to which is devoted this chapter.

Example 1

A structure consists of 70 members of the same type. A random sample of 10 members can be taken from the population of 70 units using a table, or a generator of random numbers within a range of 1 to 70. A sample can then be created by taking the units whose serial numbers are equal to ten generated random numbers.

A.2 Characteristics of Location

The basic characteristic of sample location (or its main tendency) is the sample mean m_X given as

$$m_X = \frac{1}{n}\sum_1^n x_i \tag{A.1}$$

Here x_i denotes sample units. If the sample units are ordered from the smallest to greatest unit then the subscripts *i* are generally changed to (*i*), and the units are then denoted $x_{(i)}$.

Another characteristic of location is median m_X defined the point separating ordered sequence of data into two parts such that half of the data is less than the median and half of the data greater than the median.

Example 2

A random sample of measurements of concrete strength contains ten measurements x_i = {27; 30; 33; 29; 30; 31; 26; 38; 35; 32}in MPa. The measured data, in order of scale, is $x_{(i)}$ = {26; 27; 29; 30; 30; 31; 32; 33; 35; 38}in MPa:

The sample mean m_X and the median m_X are given as

$$m_X = \frac{1}{10}(\sum x_i) = 31.1 \text{ MPa},\ m_X = \frac{1}{2}\left(x_{(5)} + x_{(6)}\right) = 30{,}5 \text{ MPa}$$

A.3 Characteristics of Dispersion

The basic characteristic of dispersion is called the variance

$$s_X^2 = \frac{1}{n}\sum_1^n (x_i - m_X)^2 \tag{A.2}$$

In practical applications the standard deviation s_X is commonly used instead of "variance".

Another measure of dispersion that is frequently applied in engineering and science is called the coefficient of variation

$$v_X = \frac{s_X}{m_X} \tag{A.3}$$

This is, in fact, a measure of relative dispersion normalised by the sample mean m_X. It is frequently used in engineering when the sample mean m_X is not very small. If the sample mean m_X is relatively small then the standard deviation should be used instead.

In the case of very small samples ($n \leq 10$) additional measure of dispersion, called sample range, is sometimes used; it is defined simply as the difference between of the greatest and smallest sample unit, $x_{(n)} - x_{(1)}$.

In same specific cases also the mean deviation MD, or average deviation, defined as the mean of differences $|x_i - m_X|$ is also used

$$MD_X = \frac{1}{n}\sum_{i=1}^n |x_i - m_X| \tag{A.4}$$

Example 3

The variance of the sample given in Example 3.1 x_i = {27; 30; 33; 29; 30; 31; 26; 38; 35; 32}in MPa is given as

$$s_X^2 = \frac{1}{n}\sum_{1}^{n}(x_i - m_X)^2 = 11.69\ (\text{MPa})^2$$

The standard deviation is thus

$$s_X = \sqrt{s_X^2} = \sqrt{11.69} = 3.42\ \text{MPa}$$

Example 4

The coefficient of variation of the data in the random sample given in Example 3.2 x_i = {27; 30; 33; 29; 30; 31; 26; 38; 35; 32} in MPa, is given as

$$v_X = \frac{3.42}{31.1} = 0.11 = 11\%$$

Example 5

Considering ordered measurements from Example 3.2 $x_{(i)}$ = {26; 27; 29; 30; 30; 31; 32; 33; 35; 38}in MPa, the variation range and the mean deviations are:

$$x_{(n)} - x_{(1)} = 38 - 26 = 12\ \text{MPa}$$

$$MD_X = \frac{1}{n}\sum_{i=1}^{n}|x_i - m_X| = 2.72\ \text{MPa}$$

A.4 Characteristics of Asymmetry and Kurtosis

The characteristics of asymmetry and peakedness (kurtosis) are used less frequently than the characteristics of location (the mean m_X) and the characteristic of dispersion (the variance s_X^2). However, the characteristics of asymmetry and peakedness provide valuable information about the nature of the sample, in particular the distribution of observation to the left and right of the mean and the concentration of observation about the mean. This information may be extremely useful for determining the appropriate theoretical model (probability distribution) of population.

The following moment characteristics are most often used. The coefficient of asymmetry is defined on the basis of the central moment of the third order as

$$a_X = \frac{1}{ns_X^3}\sum_{i=1}^{n}(x_i - m_X)^3 \qquad \text{(A.5)}$$

Similarly the coefficient of kurtosis is related to the central moment of the fourth order as

$$e_X = \frac{1}{ns_X^4}\sum_{i=1}^{n}(x_i - m_X)^4 - 3 \qquad \text{(A.6)}$$

Note that the above defined coefficients of asymmetry and kurtosis should be close to zero for samples taken from population having normal distribution.

The coefficient of asymmetry is positive when more sample data is on the left of the mean, positive when more data is on the right of the mean. The coefficient of kurtosis is positive when the sample data is located mostly in the vicinity of the mean, negative when the data is distributed more uniformly. Both these characteristics (skewness a_X and kurtosis e_X) are strongly dependent on abnormal deviations of some sample units (outliers), or errors, particularly in the case of small samples (n < 30). Then their evaluation may be highly uncertain (and may suffer from so-called statistical uncertainty due to limited data).

Example 6

Considering again data from Example 3.2 given as x_i = {27; 30; 33; 29; 30; 31; 26; 38; 35; 32} in MPa, the coefficients of asymmetry and kurtosis are:

$$a_X = \frac{1}{ns_X^3}\sum_{i=1}^{n}(x_i - m_X)^3 = 0.46$$

$$e_X = \frac{1}{ns_X^4}\sum_{i=1}^{n}(x_i - m_X)^4 - 3 = -0.44$$

The positive coefficient of asymmetry indicates that more observations are on the left of the mean (in fact 6 of 10 values are on the left of the mean). A slightly negative coefficient of kurtosis indicates low peakedness (observed values seem to be distributed slightly more uniformly than those of normal distribution). Note that the investigated sample is very small (10 values only), and the coefficients obtained, a_X and e_X may be inaccurate.

It is interesting to note that there is an empirical relationship between the skewness a_X the mean m_X, the median m_X and the standard deviation s_X (called sometimes as Pearson coefficient of skewness) in the form

$$a_X \approx 3(m_X - \tilde{m}_X)/\, s_X$$

Considering the results of previous Examples 3.2 and 3.3 $m_X = 31.1$ MPa, $m = 30{,}5$ MPa and $s_X = 3.42$ MPa it follows that

$$a_X \approx \frac{3(31.1-30.5)}{3.42} = 0.53$$

This seems to be a good approximation of the above obtained moment skewness $a_X = 0.46$. It also demonstrates the intuitively expected result that if the median m_X is less than the mean m_X, then the skewness a_X is positive. Consequently more data is located left of the mean than right of the mean.

A.5 General and Central Moments

Most of the samples characteristics described above belong to so called moment characteristics that are based on general or central moments of the data. The general moment

(about the origin) of the order l (l = 1, 2, 3, …) is defined as the arithmetic mean of the sum of l-powers

$$m_l^* = \frac{1}{n}\sum_{i=1}^{n} x_i^l \tag{A.7}$$

The central moment (about the mean) of the order l is similarly given as

$$m_l = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_X)^l \tag{A.8}$$

The moment characteristics can be then defined as follows.

$$m_X = m_1^* \tag{A.9}$$

$$s_X = \sqrt{m_2} \tag{A.10}$$

$$a_X = \frac{m_3}{m_2^{3/2}} \tag{A.11}$$

$$e_X = \frac{m_4}{m_2^2} - 3 \tag{A.12}$$

In numerical calculation it is sometime useful to apply the following relations between the general and central moments

$$m_2 = m_2^* - m_X^2 \tag{A.13}$$

$$m_3 = m_3^* - 3m_X m_2^* + 2m_X^3 \tag{A.14}$$

$$m_4 = m_4^* - 4m_X m_3^* + 4m_X^2 m_2^* - 3m_X^4 \tag{A.15}$$

When computers are used to evaluate statistical samples Equations (A.13) to (A.15) are not directly used.

A.6 Combination of Two Random Samples

Sometimes it is necessary to combine two random samples taken from one population, assuming that the characteristics of both the samples are known, but the original observations x_i are not available. It must be emphasised that only homogeneous samples of the same origin (taken from one population under the same conditions) should be combined. Violation of this important assumption could lead to incorrect results.

Assume that a first sample of the size n_1 has the characteristics m_1, s_1, a_1, while a second sample of the size n_2 has the characteristics m_2, s_2, a_2. Only three basic characteristics are considered here (the coefficients of kurtosis are rarely available for combined samples). The resulting characteristics of a combined sample of the size n can be determined from the following expressions:

$$n = n_1 + n_2 \tag{A.16}$$

$$m = \frac{n_1 m_1 + n_2 m_2}{n} \tag{A.17}$$

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n} + \frac{n_1 n_2}{n^2}(m_1 - m_2)^2 \tag{A.18}$$

$$a = \frac{1}{s^3}\left[\frac{n_1 s_1^3 a_1 + n_2 s_2^3 a_2}{n} + \frac{3 n_1 n_2 (m_1 - m_2)(s_1^2 - s_2^2)}{n^2} - \frac{n_1 n_2 (n_1 - n_2)(m_1 - m_2)^3}{n^2}\right] \tag{A.19}$$

It is interesting to note that the standard deviation s is dependent not only on the standard deviations of two initial samples s_1 and s_2, but also on the means of both the samples. Similarly, the skewness a also depends on the characteristics of the lower order (means and standard deviations). The relationship for the kurtosis is not included as it is not commonly used.

It should be noted that if the original data is available then it can be analysed as one sample; Relationships (A.16) to (A.19) can then be used for checking newly obtained results. The most important thing is the verification of the hypothesis that both samples are taken from one population.

Example 7

An example of the practical application of Equations (A.16) to (A.19) is shown underneath.

Samples	n	m	s	a	v
Sample 1	10	30.1	4.4	0.5	0.15
Sample 2	15	29.2	4.1	0.5	0.14
Combined	25	29.56	4.25	0.53	0.14

Note that a different number of sample units may affect the characteristics of the resulting combined sample. An EXCEL sheet has been developed for calculation if this is the case.

Sometimes it may occur that the size of one sample, say n_1, is not known, and only the first two characteristics m_1, s_1 are available. This is a typical situation when updating previous data with the characteristics m_1, s_1, using newly observed data of the size n_2 with the characteristics m_2, s_2. Then the Bayesian approach may be used for assessing the unknown value n_1 and a corresponding degree of freedom ν_1. The following text is presented here as a guide on how to proceed in that case, just for information and without the appropriate mathematical clarification.

In accordance with the Bayesian concept [1, 3], the unknown value n_1 and a corresponding degree of freedom ν_1 may be assessed using the relations for the coefficients of variation of the mean and standard deviation $V(\mu)$ and $V(\sigma)$, (the parameters μ and σ are considered as random variables in Bayes' concept) for which it holds

$$n_1 = [s_1 / (m_1 \, V(\mu))]^2, \; \nu_1 = 1 / (2 \, V(\sigma)^2) \tag{A.20}$$

Both unknown variables n_1 and ν_1 may be assessed independently (generally $\nu_1 \neq n_1 - 1$), depending on previous experience with a degree of uncertainty of the estimator of the mean μ and the standard deviation σ of the population. Note that for a new sample it holds that $\nu_2 = n_2 - 1$.

When the sample size n_1 and the degree of freedom ν_1 are estimated, the degree of freedom ν is given as [3, 11]

$$\nu = \nu_1 + \nu_2 - 1 \text{ if } n_1 \geq 1,\ \nu = \nu_1 + \nu_2 \text{ if } n_1 = 0 \tag{A.21}$$

Then the resulting size of the combined sample n and the mean m is given by Equations (3.59) and (3.60); the standard deviation s is determined from a modified Equation (3.61) as

$$s^2 = \left[\nu_1 s_1^2 + \nu_2 s_2^2 + \frac{n_1 n_2}{n}(m_1 - m_2)^2 \right] / \nu \tag{A.22}$$

The above relationship may be easily applied using the EXCEL sheet or other software tools.

Example 8

Suppose that from the prior production of a given type of concrete the following information is available regarding its strength

$$m_1 = 30.1 \text{ MPa},\ V(\mu) = 0.50,\ s_1 = 4.4 \text{ MPa},\ V(\sigma) = 0.28$$

For the unknown characteristics n_1 and ν_1 it follows from Equation (3.20) that

$$n_1 = \left(\frac{4,4}{30,1} \frac{1}{0,50} \right)^2 \approx 0,\ \nu_1 = \frac{1}{2 \times 0,28^2} \approx 6$$

Thus, the following characteristics are subsequently considered: $n_1 = 0$ and $\nu_1 = 6$.

To verify the quality of the concrete, new measurements have been carried out using specimens from the same type of concrete. The following strength characteristics have been obtained:

$$n_2 = 5,\ \nu_2 = n_2 - 1 = 4,\ m_2 = 29.2 \text{ MPa},\ s_2 = 4.6 \text{ MPa}.$$

Using Equations (3.16), (3.17), (3.18) and (3.19), the updated characteristics are as follows:

$$n = 0 + 5 = 5$$

$$\nu = 6 + 4 = 10$$

$$m = \frac{0 \times 30,1 + 5 \times 29,2}{5} = 29,2 \text{ MPa}$$

$$s^2 = \left[6 \times 4,4^2 + 4 \times 5,6^2 + \frac{0 \times 5}{5}(30,1 - 29,2)^2 \right] / 10 = 4,5^2 \text{ MPa}^2$$

Thus, using the previous information, the standard deviation of the new measurements could be decreased from s = 5.6 MPa to s = 4.5 MPa.

However, it should be noted that the combination of the previous information with the current measurements might not always lead to favourable results. For example, if the

coefficients of variation are $w(\mu)$=0.2 and $w(\sigma)$=0.6, then the unknown characteristics n_1 and ν_1 follow from Equation (3.20) as

$$n_1 = \left(\frac{4,4}{30,1} \frac{1}{0,2} \right)^2 \approx 1; \quad \nu_1 = \frac{1}{2 \times 0,6^2} \approx 1$$

In this case

$$n = 1 + 5 = 6$$

$$\nu = 1 + 4 - 1 = 4$$

$$m = \frac{1 \times 30,1 + 5 \times 29,2}{6} = 29,35 \text{ MPa}$$

$$s^2 = \left[1 \times 4,4^2 + 4 \times 5,6^2 + \frac{1 \times 5}{6} (30,1 - 29,2)^2 \right] / 4 = 6,03^2 \text{ MPa}^2$$

In this case, the mean increased slightly from 29.2 to 29.35, while the standard deviation increased considerably, from 5.6 to 6.03. However, this is an extreme case, caused by unfavourable estimates of n_1, ν_1 and ν following on from Equations (3.20) and (3.21). In practical applications these equations should be applied with caution, particularly in extreme cases similar to the above example. In connection with this warning, an important assumption mentioned at the beginning of this section should be stressed. Only those samples that are evidently taken from the same population can be used for combining or updating statistical data; otherwise the results of the combination of two random samples may lead to incorrect results.

A.7 Note on Terminology and Software Products

It should be mentioned that documents such as ISO 3534 [7], [8] and software products EXCEL, MATHCAD and STATISTICA provide slightly modified terminology and definitions for basic moment characteristics.

In general two modifications are commonly used for the characteristic of dispersion:

– The characteristic called here “the sample standard deviation” is also denoted as “the standard deviation of a sample”, or as “the population standard deviation” (when n is the population size), and is given as

$$s_X = \sqrt{\frac{1}{n} \sum_1^n (x_i - m_X)^2} \tag{A.23}$$

– The sample estimate of the population standard deviation called here a point estimate of the population standard deviation and denoted by the symbol $\hat{s}_X$ is sometimes called the sample standard deviation

$$\hat{s}_X = \sqrt{\frac{1}{n-1} \sum_1^n (x_i - m_X)^2} \tag{A.24}$$

Expression (A.23) corresponds to Equation (A.2) for the sample standard deviation. Expression (A.24) represents a point estimate of standard deviation that is derived from the

mean of the distribution describing the sample variance (based on the χ^2 random variable and discussed in [1], [2], [3] and [4]).

Similar modifications of sample characteristics are also available for the skewness and kurtosis. The "sample skewness" *a* defined here by Equation (A.5) can be written in simplified form as

$$a_X = \frac{m_3}{m_2^{3/2}} = \frac{1}{n s_X^3} \sum_{i=1}^{n} (x_i - m_X)^3 \tag{A.25}$$

STATISTICA, EXCEL, MATHCAD and some other software products provide a point estimate of the population skewness $\hat{a}_X$ (see Chapter 8) as

$$\hat{a}_X = \frac{n^2}{(n-1)(n-2)} \frac{1}{n \hat{s}_X{}^3} \sum_{i=1}^{n} (x_i - m_X)^3 = \frac{\sqrt{n(n-1)}}{(n-2)}\, a_X \tag{A.26}$$

Note that the population estimate $\hat{s}_X$ is used in Equation (3.26). If the sample standard deviation is used then the estimate of the population skewness would be

$$\hat{a}_X = \frac{n}{(n-1)(n-2)} \frac{1}{s_X^3} \sum_{i=1}^{n} (x_i - m_X)^3 = \frac{n^2}{(n-1)(n-2)}\, a_X \tag{A.27}$$

The factor enhancing the sample skewness a_X in Equation (A.27) (the fraction containing the sample size *n*) is slightly greater than the similar factor in Equation (A.26) (for $n > 30$ by less than 5 %); the difference diminish with increasing sample size *n*

Similar modifications of sample characteristics may be found for kurtosis based on the central moment of the fourth order (see Equation (A.6)). The relevant formulae can be found in the help component of the relevant software products. However, kurtosis is evaluated in practical applications very rarely and only for very large samples ($n > 100$).

A.8 Grouped Data, Histogram

When analysing large size of statistical data *n*, it is often useful to group them into a limited number of classes *k* (usually $7 \le k \le 20$) and to determine the number of units belonging to each class n_i ($i = 1,2,\ldots k$), called class frequency ($\Sigma n_i = n$). Each class is represented by class mark x_i^* which is the midpoint of the class interval limited by its lower and upper class limit.

Commonly, the grouped data are presented graphically in the form of a histogram, which is a column diagram showing frequency n_i or relative frequency n_i/n for each class. Histograms are very useful graphical tools providing valuable information about the overall character of the sample. Visual investigation of the histogram is always recommended. It may provide an initial understanding of the sample nature.

The mean m_X is given by the general moment of the first order (A.7), which for grouped data is written as

$$m_X = m_1^* = \frac{1}{n} \sum_{i=1}^{k} n_i\, x_i^* \tag{A.28}$$

The central moments (about the mean) of the order *l* are for grouped data given as

$$m_l = \frac{1}{n} \sum_{i=1}^{k} n_i (x_i^* - m_X)^l \tag{A.29}$$

The moment characteristics of grouped data can be determined using the general Formulae (A.10) to (A.12). Also the relationships between the general and central moments provided by Equation (A.13) to (A.15) can be used in the numerical evaluation of grouped data.

Example 9

Results of $n = 90$ tests of concrete strength are grouped into $k = 9$ classes as indicated in the table below and in the histogram in Fig. A.1. Visual investigation of the histogram indicates that the sample is well-ordered (without outliers), symmetric (the skewness is expected to be close to zero) and slightly less spiky (more flat) than commonly used normal distribution (a bit of negative kurtosis is expected).

Class i	Class interval in MPa	Class mark x_i^* in MPa	Frequency n_i	Product $n_i\, x_i^*$	Product $n_i\, (x_i^* - m_X)^2$
1	16 to 18	17	1	17	71.309
2	18 to 20	19	3	57	124.593
3	20 to 22	21	12	252	237.037
4	22 to 24	23	15	345	89.630
5	24 to 26	25	20	500	3.951
6	26 to 28	27	18	486	43.556
7	28 to 30	29	11	319	139.062
8	30 to 32	31	8	248	246.914
9	32 to 34	33	2	66	114.173
Sum	-	-	90	2290	1070.222

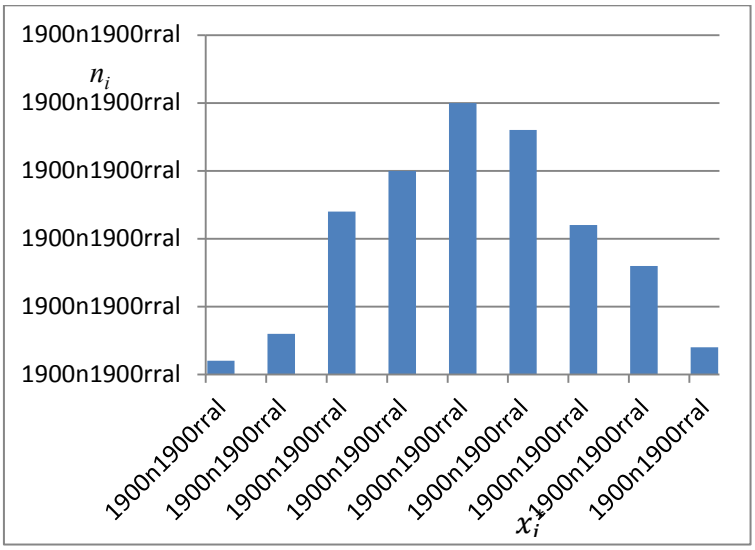


Figure A.1 Histogram of the grouped data from Example 3.9 (90 observations of concrete strength)

The table shows the class intervals, class marks x_i^* (in MPa), frequency n_i and products $n_i\, x_i^*$ and $n_i\, (x_i^* - m_X)^2$ used to calculate the general moments of the first order, and the central moment of the second order. The moments of the order three and four would be necessary for calculation of the skewness a_X and kurtosis e_X.

It follows from Equations (A.7), (A.10) and the numerical results shown in the last row of the above table that the sample mean and standard deviation are

$m_X = m_1^* = 2290/90 = 25.44$ MPa and $s_X = \sqrt{m_2} = (1070.222/90)^{0.5} = 3.45$ MPa

The coefficient of variation $v_X = 3.45/25.44 \approx 0.14$ is relatively high and indicates a somewhat low quality of material. The other moment characteristics can be similarly found using the central moments of higher order and general Equations (A.11) and (A.12). This way it can be found that the sample skewness is almost zero, $a = 0.03$, and the kurtosis $e = -0.53$. So the sample is really symmetrical and slightly more uniform than the normal distribution.

REFERENCES

[1] Ang A. H.-S. and Tang W. H. (2007) Probabilistic Concepts in Engineering. Emphasis on Applications to Civil Environmental Engineering. John Wiley and Sons, New York.

[2] Devore J. and Farnum N. (2005) Applied Statistics for Engineers and Scientists. Thomson, London.

[3] Dunin-Barkovskij I. V. and Smirnov N. V. (1955) The theory of probability and mathematical statistics in engineering (in Russian). Moscow.

[4] Gurskij E. I. (1971) The theory probability with elements of mathematical statistics (in Russian), Moscow.

[5] Holický M. (2009) Reliability analysis for structural design. SUNN MeDIA, Stellenbosch.

[6] ISO 12491, Statistical methods for quality control of building materials and components (1997).

[7] ISO 3534-1, Statistics – Vocabulary and symbols – Part 1: Probability and general statistical terms (1993).

[8] ISO 3534-2, Statistics – Vocabulary and symbols – Part 2: Statistical quality control (1993).

[9] Holický M. (2013) Introduction to Probability and Statistics for Engineers, Springer-Verlag, ISBN 978-3-642-38299-4.

This project has been funded with support from the European Commission.
This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.