DISS. ETH NO. 18238

Issues of sustainability in engineering decision analysis

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

Kazuyoshi Nishijima

Master of Environmental Studies, University of Tokyo

born 13.08.1978

citizen of
Japan

accepted on the recommendation of

Prof. Dr. Michael Havbro Faber, examiner
Prof. Lucas Bretschger, co-examiner
Prof. Niels Lind, co-examiner

2009

Acknowledgements

"You are one of the freest PhD students in the world," my supervisor Professor Faber once said to me - and yes, I am. He gave me this topic and motivated me to conduct this research work. However, during the course of the work he never gave me any "fixed" assignments. Instead, he gave me the opportunity to explore different research disciplines. At the same time, whenever necessary, he gave me exceptional support, both professionally and personally. Here, I express my deep gratitude to him.

My colleagues at the group of Risk and Safety have created a friendly, pleasant atmosphere, conductive to research work. It is thanks to them that I could always concentrate on the research work. Among others, I express my wholehearted thanks to Matthias Schubert, who was my office mate over the last few years. Whenever I encountered technical or personal problems, he generously gave of his time to discuss them with me: he understood the problems and provided me with useful advice. I also express my gratitude to Patricia Meile, who is a secretary in our research group. She constantly supported me in setting up and improving the working environment, including administrative issues such as acquiring working permission for a foreign student. Dr. Daniel Straub, who is a senior ex-colleague in the group, greatly influenced my motivation towards research and my research style, directly and indirectly, which I appreciate very much.

I also want to thank all my friends and especially my family for their support. Their support has encouraged me to continue with the PhD work. This work is dedicated to them.

Finally, my gratitude goes to Professors Bretschger and Lind for acting as external examiners and providing valuable comments.

Zurich, 28.02.2009
Kazuyoshi Nishijima

Abstract

Sustainable societal development has become a subject of increased and widespread societal attention especially during the last two decades. The tremendous economic development of former developing nations such as China and India and the general impact of globalization have put even larger pressures on our limited natural resources and fragile environment. Faced with an ever increasing amount of evidence that the activities of our own generation might actually impair the possibilities for future generations to meet their needs, it has become a major political concern that societal development must be sustainable. The issuing of the famous Brundtland report “Our Common Future” (1987) formed a political milestone. This important event has enhanced the public awareness that substantial changes of consumption patterns are called for and has further significantly influenced research agendas worldwide.

The realization of a sustainable development of society necessitates that a holistic perspective is taken in operational and strategic societal decision-making. In principle, a joint consideration of the preferences, needs and capabilities of the present and future generations across all nations, industrial and public sectors is required if we are to fully succeed in achieving sustainable societal development. It may be realized that decisions made to enhance sustainability of societal development not only concern reduced emissions of pollutants but also directly and indirectly involve a redistribution of globally available resources and not least a reassessment of the societal affordability of lifestyle and quality of life. So far, the available research literature in this field has mainly reported on results relating to individual aspects of sustainable development; as of yet a general framework that facilitates the joint consideration of the many dimensions of sustainability in supporting decision-making for sustainable societal development is still missing.

Whereas the development of a general framework for sustainable decision-making is one of the most relevant tasks in the research agenda, it is unlikely that this task could be accomplished in the foreseeable future. However, at the same time, there is an urgent need for methods that enable societal decision-makers to identify "sustainable" policies in different sectors of society. Here, the "sustainable" policies imply policies that conform to current preventive measures, regulations, principles, ethics and whatever else is regarded as best practice for the realization of the sustainable development of society.

Motivated by this and focusing on the civil engineering sector, the present thesis has two aims. The first aim is to reformulate the classical life-cycle cost optimization concept, which has been advocated in civil engineering as the decision principle, in such a way that relevant aspects of sustainability can be incorporated into engineering decision-making. The aspects of sustainability considered in depth in this

reformulation are intergenerational equity and allocation of limited resources. Furthermore, for the purpose of facilitating the applications in practical decision situations, a platform is proposed for the modelling and optimization of decision problems based on Bayesian probabilistic networks. Thereby, it is possible with the proposed platform to consider the constraints relating to societal sustainability posed by present society in the decision problems. The second aim is to present a fundamental approach for incorporating the reliability of civil infrastructure in general economic models so that the sustainable policies on design and maintenance of civil infrastructure can be identified from a macroeconomic perspective.

In the present thesis, two types of engineering decision analyses are differentiated in order to clarify the extent of the consequence of decisions; marginal engineering decision analysis and non-marginal engineering decision analysis. In marginal engineering decision analysis, it is assumed that the economic growth path is exogenously given and the consequence of decisions does not affect the economic growth; the life-cycle cost optimization concept corresponds to the marginal engineering decision analysis; the first aim of the present thesis can be regarded as the formulation of engineering decision problems from a sustainability perspective in the context of the marginal decision analysis. In contrast, non-marginal decision analysis considers the change of economic growth as a consequence of decisions; the second aim of the present thesis can be regarded as a proposal for a decision framework for the non-marginal engineering decision analysis.

The present thesis consists of eight chapters. Chapter 1 introduces the background, aim, scope and outline of the thesis. A literature survey is also provided in the fields of economics and civil engineering, where the formulation and optimization of sustainable decision making in civil engineering is dealt with. The core of the present thesis consists of six chapters (Chapters 2 to 7). Each of the chapters, except Chapter 7, represents a part of my research work published during the PhD study. Chapter 2 considers the general treatment of uncertainties in engineering decision analysis, which is the philosophical basis for decision-making subject to uncertainties. Chapters 3 to 5, respectively, investigate the modelling and optimization of sustainable decision problems, the issue of intergenerational equity and the issue of allocation of limited resources in the context of marginal engineering decision analysis. In Chapter 6 the approach for incorporating the reliability of civil infrastructure in general economic models is proposed based on economic growth theory. This approach corresponds to non-marginal engineering decision analysis. The proposed approach is then applied to a simplistic economic model in Chapter 7 in order to show how the optimal reliability of civil infrastructure can be identified and the sustainable policy on the design and maintenance of civil infrastructure can be examined. Thereby, an objective function is derived in the context of non-marginal decision analysis that is different from the objective function employed in the classical life-cycle cost optimization concept. The reason for this is provided by looking at the differences in the formulation of the decision problems in marginal and non-marginal decision analysis. In this chapter the assumptions of the derivation of the classical life-cycle cost optimization and its limitations are also introduced in order to emphasize the difference between non-marginal decision analysis and marginal decision analysis. Chapter 8 concludes the present work.

In the reformulation of the classical life-cycle cost optimization, its practical applicability is emphasized. Hence, the proposed methods in the corresponding chapters (Chapters 3 to 5) can be readily applied to practical decision situations. Practical examples are provided in these chapters. On the other hand, the approach presented in Chapters 6 and 7 serves as a relevant building block for further development of the general framework for sustainable decision-making, whereby scientific insights are provided on how sustainable design and maintenance policies on infrastructure can be investigated in a macroeconomic context.

Zusammenfassung

Die Frage nach einer nachhaltigen gesellschaftlichen Entwicklung hat insbesondere in den letzten zwei Jahrzehnten zunehmend an Bedeutung gewonnen. Im Fokus stehen dabei die begrenzten natürlichen Ressourcen und die fragile Umwelt, die durch die enorme wirtschaftliche Entwicklung von Schwellenländern wie China und Indien noch stärker unter Druck geraten. Da es immer offensichtlicher wird, dass die Aktivitäten unserer eigenen Generation die Entwicklungsmöglichkeiten der folgenden Generationen beeinträchtigen könnten, wurde die Forderung nach einer nachhaltigen gesellschaftlichen Entwicklung ein wesentliches politisches Ziel. Ein politischer Meilenstein wurde 1987 durch den Brundtland Report "Unsere gemeinsame Zukunft" gesetzt. Dieses entscheidende Ereignis verstärkte das öffentliche Bewusstsein, dass substantielle Änderungen im Konsumverhalten zukünftig notwendig sind. Seit der Veröffentlichung des Brundlandt Reports beeinflusst das Thema der Nachhaltigkeit weltweit viele Agenden von Forschergruppen.

Die Umsetzung einer nachhaltigen gesellschaftlichen Entwicklung erfordert eine Einnahme einer holistischen Perspektive sowohl für die operationelle als auch für die strategische Entscheidungsfindung in der Gesellschaft. Prinzipiell ist eine integrale Berücksichtigung der Präferenzen, Bedürfnisse und Fähigkeiten der heutigen und der zukünftigen Generationen über alle Nationen und alle Sektoren hinweg notwendig, wenn eine Steuerung hin zu einer nachhaltigen gesellschaftlichen Entwicklung erfolgreich sein will. Es muss erreicht werden, dass Entscheidungen zur Förderung der nachhaltigen Entwicklung einer Gesellschaft nicht nur unter Berücksichtigung monokausaler Zusammenhängegetroffen werden, z.B. die Verringerung von schädlichen Emissionen, sondern auch unter Berücksichtigung der direkten und indirekten Umverteilung globaler Ressourcen, der Neubewertung von Lebensstilen und nicht zuletzt der Qualität des Lebens in der globalen Welt. Der Grossteil der verfügbaren wissenschaftlichen Literatur zum Thema Nachhaltigkeit fokussiert auf einzelne Aspekte, die für eine nachhaltige Entwicklung notwendig sind. Ein genereller Rahmen, der die gemeinsame Betrachtung des mehrdimensionalen Problems der Nachhaltigkeit erlaubt und gesellschaftliche Entscheidungsträger unterstützen kann, fehlt bisher noch.

Die Entwicklung eines solchen Rahmens ist die relevanteste Aufgabe, die die Forscher im Bereich der nachhaltigen Entscheidungsfindung zu bewältigen haben. Es ist nicht abzusehen, dass in naher Zukunft in diesem Bereich eine Lösung gefunden wird. Dennoch ist derzeit der Druck gross, Methoden zur Verfügung zu haben, die es Entscheidungsträgern aus allen Bereichen ermöglicht, die "nachhaltigste" Handlungsalternative zu identifizieren. Der Ausdruck " nachhaltigste" impliziert, dass die Handlungsalternativen konform sind zu den Massnahmen, Regulierungen, Prinzipien, Ethiken und allen anderen Gegebenheiten in einer Gesellschaft, die als "beste Praxis" für die Umsetzung der nachhaltigen Entwicklung in einer Gesellschaft gelten.

Diese vielschichtigen Aspekte waren die Motivation für diese Arbeit, die sich auf den Bereich des Bauingenieurwesens bezieht. Zwei wesentliche Ziele werden in dieser Arbeit verfolgt. Das Erste ist, den klassischen Ansatz des Konzeptes zur Optimierung der Lebenszykluskosten, der im Bereich des Bauingenieurwesens als das Entscheidungsprinzip betrachtet wird, so umzuformulieren, dass Aspekte der Nachhaltigkeit im Entscheidungsprozess Berücksichtigung finden können. Die Aspekte der Nachhaltigkeit, die insbesondere Berücksichtigung in der Neuformulierung finden sind das Prinzip der intergenerationellen Gleichheit und der Allozierung von beschränkten Ressourcen. Für die Anwendbarkeit in realen Entscheidungssituationen wird eine Plattform für die Modellierung und Optimierung von Entscheidungsproblemen vorgeschlagen, die auf Bayes'schen Probabilistischen Netzen basiert. Dies ermöglicht es, die Einschränkungen, die durch die Aspekte der Nachhaltigkeit gegeben sind, im Entscheidungsprozess zu berücksichtigen. Das zweite Ziel ist, einen fundamentalen Ansatz vorzustellen, der es ermöglicht, strukturelle Zuverlässigkeit von baulichen Infrastrukturen in allgemeinen ökonomischen Modellen zu berücksichtigen, so dass nachhaltige Entscheidungen in Bezug auf den Entwurf und den Unterhalt solcher Anlagen von einer makroökonomischen Perspektive aus identifiziert werden können.

Zwei Typen von Entscheidungsanalysen im Ingenieurwesen wurden in dieser Arbeit unterschieden, um das Ausmass der Konsequenzen aus Entscheidungen klar herauszustellen; es werden sowohl marginale Entscheidungsanalysen als auch nicht-marginale Entscheidungsanalysen beleuchtet. In der marginalen Entscheidungsanalyse im Ingenieurwesen wird angenommen, dass das wirtschaftliche Wachstum exogen gegeben ist und die Konsequenzen, die aus Entscheidungen resultieren, keinen Einfluss auf das wirtschaftliche Wachstum haben. Das Konzept der Optimierung der Lebenszykluskosten von baulichen Infrastrukturen ist ein Beispiel für eine marginale Entscheidungsanalyse. Damit kann das zuvor genannte erste Ziel dieser Arbeit als Formulierung von Entscheidungsproblemen im Hinblick auf Nachhaltigkeit im Kontext der marginalen Entscheidungsanalyse gesehen werden. Im Gegensatz dazu kann das zweite formulierte Ziel als ein Rahmen für Entscheidungen gesehen werden, die einen nicht-marginalen Einfluss auf das Wirtschaftswachstum haben.

Die vorliegende Arbeit gliedert sich in acht Kapitel. Kapitel 1 stellt die Ziele der Arbeit vor, grenzt die Arbeit ab und erläutert die Hintergründe zu dieser Arbeit. Im ersten Teil wird ein Überblick über die Literatur in den relevanten Gebieten der Wirtschaftswissenschaften und des Bauingenieurwesens, insbesondere in den Bereichen Formulierung und Optimierung von nachhaltigen Entscheidungsproblemen, gegeben. Der Kern dieser Arbeit besteht aus sechs Kapiteln (Kapitel 2 bis 7). Jedes dieser Kapitel (mit Ausnahme von Kapitel 7) repräsentiert einen Teil meiner Forschungsarbeiten während des Doktorats, die bereits veröffentlicht sind oder zur Veröffentlichung akzeptiert sind. Kapitel 2 behandelt den allgemeinen Umgang mit

Unsicherheiten in der Entscheidungsanalyse im Ingenieurwesen und stellt die philosophische Basis für die Entscheidungsfindung im Ingenieurwesen unter Unsicherheit dar. Kapitel 3 bis 5 untersucht die Modellierung und Optimierung von Entscheidungsproblemen unter Berücksichtigung der zuvor genannten Aspekte der Nachhaltigkeit. Kapitel 6 stellt einen Ansatz vor, mit dem die strukturelle Zuverlässigkeit baulicher Infrastrukturen in allgemeinen wirtschaftswissenschaftlichen Modellen und Modellen zur Beschreibung des Wirtschaftswachstums berücksichtigt werden kann. Dieser Ansatz korrespondiert zu nicht-marginalen Entscheidungsanalysen. In Kapitel 7 wird dieser Ansatz an einem einfachen wirtschaftswissenschaftlichen Modell angewendet, um zu zeigen, wie die optimale Zuverlässigkeit baulicher Infrastrukturen identifiziert werden kann, und eine nachhaltige Strategie in Bezug auf den Entwurf und den Unterhalt verfolgt werden kann. Dazu wird eine Zielfunktion in einem nicht-marginalen Kontext hergeleitet, die grosse Unterschiede zur Zielfunktion aufweist, die im klassischen Ansatz zur Optimierung der Lebenszykluskosten verwendet wird. Der Grund für diese Unterschiede liegt in der Formulierung des Problems im marginalen und im nicht-marginalen Entscheidungsraum. In diesem Kapitel wird auch auf die klassischen Annahmen und Einschränkungen eingegangen, um die Unterschiede in diesen beiden Ansätzen beleuchten zu können. Kapitel 8 schliesst die Arbeit.

In der Neuformulierung des klassischen Lebenszyklusansatzes wird die praktische Anwendbarkeit unterstrichen. Daher können die Methoden, die in den Kapiteln 3 bis 5 vorgestellt werden, direkt in praktischen Problemen angewendet werden. Hierzu werden in diesen Kapiteln praktische Beispiele gegeben. Auf der anderen Seite ist der Ansatz, der in Kapitel 6 und 7 vorgestellt wird, ein relevanter Baustein für die weitere Entwicklung eines allgemeinen Rahmenwerks für die nachhaltige Entscheidungsfindung, wobei wissenschaftliche Einblicke gegeben werden, wie nachhaltige Entwurfs- und Unterhaltsstrategien an baulichen Anlagen in einem makroökonomischen Kontext untersucht werden können.

TABLE OF CONTENTS

1. Introduction

1.1. Relevance

Sustainable design and maintenance policies on civil infrastructure have become a relevant subject in both developed and developing countries. Many developed countries are presently experiencing severe deterioration of older infrastructure. Developing countries are repeatedly faced with the losses of infrastructure due to natural hazards. In addition, these countries continuously suffer from losses of infrastructure due to deterioration that arises from the lack of appropriate maintenance work.

In some developed countries, a considerable amount of economic resources is allocated for maintenance work for civil infrastructure. For instance, in 2006 Switzerland allocated 2.3% of its GDP to the investment into civil infrastructure and 54% of this investment was used for maintenance work[1]. This ratio is high in comparison to the average ratio for the European countries, which was found to be 31.4%[2]. However, in other developed countries, the resources allocated for maintenance work for civil infrastructure are not sufficient, and additional resources are urgently called for in order to restore deteriorated infrastructure to a good condition. The Report Card for America's Infrastructure (ASCE (2005)) estimates that US$1.6 trillion is needed over the next five-year period in the United States, which amounts to approximately 10% of the country's annual GDP. JSCE (2008) reports that Japan is expected to experience severe deterioration of infrastructure by 2025 like what the United States is presently experiencing, since infrastructure in Japan was mainly constructed between 1970s and 1980s, and the infrastructure constructed in this period will exhibit severe deterioration in the near future. Developing countries have the same problem, i.e. lack of resources for maintenance work. However, they are faced with an even more difficult situation, since they also suffer from the lack of resources for the construction of new infrastructure. In these countries the optimal balance of resource allocation between construction and maintenance work is not yet obvious, whereas the World Bank (1994) assesses that an additional US$12 billion spent for maintenance work for road networks in African countries could save US$45 billion which otherwise have to be spent on the reconstruction of the severely deteriorated road networks.

1 These numbers are calculated based on the statistics provided by EUROCONSTRUCT (2007).

2 The average over Austria, Belgium, Czech republic, Denmark, Finland, France, Germany, Hungry, Ireland, Italy, the Netherlands, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland and the United Kingdom, which are included in EUROCONSTRUCT (2007).

The statistics on economic losses due to natural hazards are summarized by Guha-Sapir et al. (2004). These statistics show that during the period 1974 to 2003 the highest economic losses due to natural hazards were brought about by: an earthquake in Japan in 1995, US$159 billion[3]; flooding in China in 1998, US$22.6 billion; a hurricane in the United States in 1992, US$39.4 billion. However, the same statistics show the opposite story if the economic impact is measured in terms of a proportion of GDP[4]. For instance, the greatest economic impact was caused by: earthquake in Guatemala in 1976, 27% of the GDP; flood in Yemen in 1996, 28% of the GDP; wind storm in St Lucia in 1988, 413% of the GDP. These countries are small in economic terms and/or geographical size. Other developing countries have suffered from major natural hazards, e.g. the flood event of 1987 in Bangladesh, the earthquake event of 1990 in Iran, the earthquake and associated tsunami event of 2004 in Southeast Asian countries[5].

Deterioration of infrastructure and losses of infrastructure due to natural hazards are inevitable. However, these are manageable to a large extent by means of design and maintenance policies on civil infrastructure. Thus, the statistics shown above raise the question: were the past policies on design and maintenance of infrastructure optimal? And if this is not the case, which are the optimal policies for the long-term development of societies, i.e. what policies are sustainable?

Today, due consideration of sustainability is required in almost all civil engineering decision situations. These decision situations include appraisal of new civil infrastructure projects, ranking of rehabilitation measures for deteriorating infrastructure, preparing design codes, donations and investments by international organizations to civil infrastructure projects in poor countries. Since these activities are supported by the public and undertaken on behalf of society, it is of the utmost importance that the process of decision-making in such activities is clear, transparent and consistent.

1.2. Aim of the thesis

The issue of sustainability is a complex issue that concerns many different aspects of society and the environment, involving different stakeholders. Thus, it is unlikely that a commonly agreed, general framework for sustainable decision-making can be established in the near future. On the other hand, there is an urgent need for methods that enable societal decision-makers to identify "sustainable" policies for civil infrastructure projects. Herein, the "sustainable" policies imply the policies that conform to current preventive measures, regulations, principles, ethics and whatever

3 Adjusted to US dollar in 2003. The same applies in the following unless otherwise stated.
4 GDP in the previous year of the hazard event occurrence.
5 Note that economic loss induced by Hurricane Katharina in 2005 is estimated at US$125 billion, Munich Re (2005). However, this amounts to only slightly more than 1% of the GDP of the United States in 2004, i.e. US$11 trillion (World Development Indicators Database, World Bank).

are regarded as best practices for the realization of a sustainable development of society; due to the absence of a general framework for sustainable decision making, these best practices may be less efficient, but these are often undertaken in preventive manners to avoid irreversible consequences.

Motivated by this and focusing on the civil engineering sector, the present thesis has two aims. The first aim is to reformulate the classical life cycle cost optimization concept advocated in civil engineering as the decision principle, in such a way that relevant aspects of sustainability can be incorporated in engineering decision-making. The relevant aspects of sustainability considered in this reformulation are intergenerational equity and allocation of limited resources. Furthermore, for use in practical decision situations, a platform is proposed for the modelling and optimization of decision problems based on Bayesian probabilistic networks. The proposed platform enables one to consider the constraints dictated by society in terms of, e.g., regulations for the realization of the sustainable development of society. The second aim is to provide a fundamental approach for incorporating the reliability of civil infrastructure in general economic models so that the appropriate policies for design and maintenance on civil infrastructure can be identified in the context of macroeconomics.

To achieve these aims systematically and also to facilitate a clear focus on individual problems, the following four issues are identified. In the present thesis, each of these issues is investigated individually.

Issue 1: Uncertainties

Decisions involving design and maintenance policies on civil infrastructure must be made subject to significant uncertainties. These uncertainties are associated with the randomness of natural phenomena such as the physical process of material deterioration, a change of the environment surrounding the infrastructure and the occurrence of natural hazards, in two ways. Firstly, the randomness of nature itself is one of the uncertainties (aleatory uncertainty). By definition, this type of uncertainty cannot be reduced. Secondly, modelling the characteristics of the randomness of nature constitutes the other type of uncertainty (epistemic uncertainty). In principle, this type of uncertainty can be reduced by a better understanding of the phenomena; however, although some of the epistemic uncertainties may be reduced by merely collecting more information, for others a reduction may not be possible in the foreseeable future. Both types of uncertainty are relevant to decision problems when looking at the choice of optimal policies, and they must be consistently taken into account in the decision problems.

Issue 2: Adaptation of optimization problems to sustainable decision-making

Economic growth is a societal goal. At the same time, besides economic growth there are a number of societal preferences. These preferences concern, for instance, the preservation of natural resources including landscape, biodiversity and non-renewable resources, degree of homogeneity of welfare between members in society, and human safety. These preferences must be fully taken into account in societal decision-making. Thus, as part of such societal decisions, decisions regarding design and maintenance policies on civil infrastructure often take the form of multi-objective optimization problems, or otherwise constrained optimization problems where societal preferences and other boundary conditions such as constraints on the amount of resources available act as the constraints in optimization problems.

Issue 3: Inter-generational equity
Civil infrastructure provides benefits to society in terms of direct increase of economic output and direct as well as indirect increase of social welfare over the long period of its operation, possibly over a number of generations. At the same time, construction and maintenance work of the civil infrastructure incur costs over the entire operation period. Since the temporal distribution of such costs depends on the chosen design and maintenance policies, the optimal choice of the policies is considered as a decision problem in regard to fair distribution of the benefits and costs over different generations.

Issue 4: Balance between quality and quantity
Civil infrastructure is important for economic growth. An increase in the quantity of civil infrastructure capital increases economic output. Thus, given an amount of investment in civil infrastructure, it is possible to achieve a higher economic output at least in the short term by reducing the quality of infrastructure. This is because a unit of infrastructure capital can be constructed and maintained less expensively, and as a result the amount of constructed infrastructure can be increased. One of the consequences of this strategy is a higher deterioration rate of the infrastructure in the long term; this strategy may partly correspond to the strategies taken in the past by some developed countries that are presently suffering from severe deterioration of civil infrastructure. In contrast, high-quality infrastructure can be much more durable, though it can be realized only at higher costs – not only higher costs of construction and maintenance work but also a lower economic output in the short term due to a smaller accumulation rate of the capital.

1.3. Scope of the thesis

In the course of investigating these issues, the present work makes several assumptions. The most critical assumptions are: definition and formulation of sustainability are assumed to be given; acceptable levels concerning several aspects, e.g. human safety, environment and use of resources etc., are assumed to be given. These assumptions effectively mean that the forms of the objective function (utility function or social welfare function) and constraints, i.e. the general rule set for sustainability, are assumed to be given. In fact, the general rule set could be established given general agreement on the implications of sustainability within/between groups in society, e.g. individuals, communities, scientists and politicians. Therefore, the present work, which focuses on engineering decision analysis, does not directly discuss these topics, but instead relies on relevant related research works presently available. The state of the art in these topics is briefly summarized in the next section, in addition to research work on the structural performance of civil infrastructure and the socio-economic role of civil infrastructure.

The present thesis defines two types of engineering decision analysis; marginal engineering decision analysis and non-marginal engineering decision analysis. An engineering decision is marginal if the consequence of the decision does not influence the economic growth of society. As will be discussed in Section 7.2 this condition is the assumption required for the application of the life-cycle cost optimization concept. The marginal decision analysis is thus most suitable e.g. for decision situations in which: private firms optimize individual engineering projects under constraints such as budget constraints and regulations imposed by authorities; societal decision-makers optimize the allocation of given resources in a portfolio of public engineering projects in which the benefits from the projects are not reinvested into capitals but are consumed. In contrast, an engineering decision is non-marginal if the consequence of the decision affects the economic growth. An important example of a non-marginal engineering decision is code-making for civil infrastructure; a higher acceptance criterion for human safety imposes higher construction and maintenance costs on civil infrastructure, which results in a smaller rate of capital accumulation.

In principle, any engineering decision-making may affect economic growth. Hence, marginal decision analysis should be regarded as an approximation of non-marginal decision analysis, although often formal non-marginal decision analysis may not be feasible in practical decision situations due to the complexity of the analysis.

The scope of the present thesis is thus to investigate the issues mentioned in the previous section in these two contexts; Issues 1 to 3 in the context of marginal engineering decision analysis and Issue 4 in the context of non-marginal engineering decision analysis.

1.4. State of the art in relevant research topics

Sustainable policy-making on civil infrastructure is interdisciplinary. It necessitates not only an understanding of the structural performance of civil infrastructure, but also of the socio-economic role of civil infrastructure. Furthermore, philosophical discussions and practical agreements on what sustainability implies are required. In the following sub-sections the state of the art in these areas is examined.

1.4.1. Structural performance of civil infrastructure

Modelling the performance of structures has a long history. Until now, significant effort has been directed towards the development of theories that describe the performance of structures. Here, one of the most important paradigm shifts is the introduction of the concept of probability: the concept that the performance of structures can/should be evaluated in a probabilistic manner. This concept is especially suited to the evaluation of the structural performance of civil infrastructure, since civil infrastructure is typically exposed to random natural phenomena, e.g. earthquakes, storms and floods, and the structural capacity of infrastructure and its modelling involves large uncertainties.

Whereas some attempts were made to base structural performance on probability (see Mayer (1926), Wierzbicky (1936) and Freudenthal (1947)), this important concept was clearly formulated by Freudenthal (1954), wherein the failure and unserviceability of structures are defined with due consideration given to uncertainties associated with both loading on and the resistance of structures. Subsequently, the theory was extended in many directions, which presently constitute the structural reliability theory. The so-called second-moment concept gained its reputation at an earlier stage in the development of structural reliability theory. This concept does not assume the form of a probability distribution function to measure the reliability of structures (reliability index), but only requires the first two orders of moments of the random variables that characterize the reliability of structures. Due to this relatively simple way of measuring the reliability, and also enhanced by the work by Cornell (1969), the concept was widely accepted.

However, for the same reason, the concept has several disadvantages. One of the most significant disadvantages is that the reliability index measured in accordance with this concept is not invariant; the measured reliability index can differ in the algebraic reformulations of the equations that mathematically represent the failure of structures, i.e. limit state functions. This "invariance" problem was solved by Hasofer and Lind (1974) with the introduction of the geometrical definition of reliability index. Thereafter, a number of its extended variants have been proposed to incorporate more information on the distributions of the random variables that characterize the reliability of structures, e.g. the first order reliability methods (FORM) and the second order reliability methods (SORM), see Ditlevsen and Madsen (2005) for an overview.

Other extensions are directed at application to the analyses for cases where the reliability of structures may change over time, see e.g. Lin (1967), Ferry-Borges and Castanheta (1971) and Vanmarcke (1983). The techniques developed for time-variant reliability analysis have been widely applied to examine e.g. the reliability of deteriorating structures and the dynamic response of structures in a probabilistic manner. However, the techniques practically applicable for these analyses are highly dependent on the nature of the stochastic processes that characterize the resistances of structures and the loads on the structures.

The structural reliability theory has also been extended to investigate the reliability of structural systems. Earlier contributions to this extension primarily focus on the development of algorithms for evaluating the probability of system failure defined by a set of limit state functions, see e.g. Hohenbichler and Rackwitz (1982), Der Kiureghian and Moghtaderi-Zadeh (1982), Ditlevsen and Bjerager (1986). Later, based on these earlier contributions, more systematic and realistic approaches have been developed for evaluating the reliability of structural systems. These approaches include the consideration of the statistical dependence of the performance of structural system components, e.g. Straub and Der Kiureghian (2008), Song and Kang (2008) and Der Kiureghian and Ditlevsen (2008).

Today, some generic software tools for the reliability analysis of structures and structural systems are available, e.g. STRUREL/COMREL (RCP GmbH) and CalREL/FERUM (Der Kiureghian et al. (2006)).

The probability-based concept for the evaluation of structural performance has been applied to the design optimization of structures within the framework of life-cycle cost analysis. Therein, the optimal design is obtained by minimizing the sum of the initial cost and the expected future costs due to possible failures. This life-cycle cost optimization concept was first introduced by Rosenblueth and Mendoza (1971) in civil engineering. At the same time, Bayesian decision theory was developed, see e.g. Raiffa and Schlaifer (1961), Lindley (1965) in general and Benjamin and Cornell (1970) for the application to civil engineering in particular. Later, the life-cycle cost optimization concept was formally integrated into the framework of Bayesian decision theory. Presently, the life-cycle cost optimization concept and Bayesian decision theory are widely accepted and employed as the guiding philosophical principles in a variety of engineering decision problems. The most important and successful applications of the concept and the theory in civil engineering include: risk-based inspection planning e.g. Tang (1973), Thoft-Christensen and Sørensen (1987), Faber et al. (2000) and Straub (2004); reassessment of existing structures, e.g. JCSS (2001a); code making, e.g. JCSS (2001b) and Rackwitz (2000).

Recently, the life-cycle cost optimization concept has been applied in the context of sustainable societal development. However, most of these applications do not explicitly consider intergenerational aspects; the utility function assumed in these applications corresponds to the utility of one representative individual who is assumed to live for an infinite time. The exception is Rackwitz et al. (2005), who consistently consider the intergenerational aspect and apply discounting accordingly for the marginal cost-benefit analysis of individual civil infrastructure projects. However, no general framework for sustainable decision-making on civil infrastructure in a macroeconomic context, i.e. a non-marginal manner, has been developed.

1.4.2. Socio-economic role of civil infrastructure

In the last two decades, the role which civil infrastructure plays in the economy has been intensively discussed. One of the relevant research questions in the discussion is the social return rate of investment in civil infrastructure. The social return rates have been estimated using a variety of historical datasets from different time periods and different countries/regions. These estimates have then been utilized to discuss the effectiveness of investment in civil infrastructure. Meanwhile, significant research efforts have been made to develop economic models, within the framework of the growth theory, that incorporate the role of civil infrastructure capital in the economy. The primary goal of these efforts is to describe the effect of investment in civil infrastructure on the long-term development of the economy, and to facilitate societal policy-making on civil infrastructure.

A pioneering work on the effectiveness of investment in civil infrastructure is that of Aschauer (1989). Based on statistics from the USA, he reveals that investment in civil infrastructure has strong explanatory power for economic productivity. Subsequently, a number of studies confirmed and generalized this observation, see review papers by Munnell (1992) and Gramlich (1994). However, this observation is critically analyzed by, among others, Holtz-Eakin and Schwartz (1995), arguing that there is little support for claims of drastic productivity boost from increased infrastructure capital. Further investigation was made by Canning and Bennathan (2000), focusing on the complementarities of civil infrastructure capital to other types of capital, e.g. physical and human capital. The results suggest that the investments in civil infrastructure are not sufficient by themselves, and the investments should be undertaken in coordination with investments in other types of capitals. Presently, whether or not current policies on investment in civil infrastructure are effective is still a controversial question, and a considerable amount of literature is available, see the review paper by Nijkamp and Poot (2004).

The assessment of the social return rate is mostly made by relying on statistical analysis techniques, especially regression analysis, see e.g. Chapters 11 and 12 in Barro and Sala-i-Martin (2004). One of the problems of standard regression analysis is that it is difficult to identify the causality between economic growth and infrastructure investment; whether economic growth demands more infrastructure capital, or whether increased infrastructure capital leads to an increase of economic output, see e.g. Duffy-Deno and Eberts (1991) and Canning and Bennathan (2000). In order to avoid the causality problem, several techniques have been developed, e.g. Engle and Granger (1987) and Canning (1999), and applied to the estimation of the social return rate.

Using these techniques, Canning and Bennathan (2000) show that investment in civil infrastructure can result in an increase of economic output.

The results of these assessments on the productivity of civil infrastructure are useful not only in discussing the effectiveness of investment in civil infrastructure, but also serve as building blocks of economic models that represent the productivity of the civil infrastructure.

The development of economic models for the economic role of civil infrastructure capital is often based on the growth theory. The growth theory aims, in general, at describing the long-term development of the economy in which different stakeholders, e.g. households, firms and governments, maximize their own objective functions. The original work on the growth theory is by Ramsey (1928). It investigates the optimal saving rate of households to achieve their maximum utility in an infinite time horizon. Today the theory presented therein forms the fundamental basis for a variety of economic theories, ranging from consumption theory, asset pricing and business-cycle theory (Barro and Sala-i-Martin (2004)). This work was later refined by Cass (1965) and Koopmans (1965). Meanwhile, Solow (1956) and Swan (1956) propose a model known today as the Solow-Swan model, which employs the neoclassical form of production function and the assumption that saving rate is constant and exogenously given. These conditions result in a very simple representation of the general equilibrium of the economy. For this reason, the Solow-Swan model is widely used, in spite of claims that the assumptions are not realistic and consistent with actual observations.

Whereas these classical models involve labor and (aggregated) capital as factors of production, modern models have been proposed that explicitly incorporate specific factors of production, e.g. technology (e.g. Arrow (1962)) and natural resources (e.g. Stiglitz (1974), Dasgupta and Heal (1974) and Solow (1974)). More recently, so-called endogenous models have been developed, which enable the long-term growth of the economy to be described without relying on exogenous growth factors (e.g. Romer (1986) and Lucas (1988)). Today, both these modern and classical models are widely applied as tools to investigate the sustainability of the economy, see e.g. Pezzey and Withagen (1998), Krautkraemer (1999) and Valente (2005).

Within the framework of the growth theory, several directions have been proposed to incorporate civil infrastructure capital in economic models as one of the production factors. For instance, Glomm and Ravikumar (1994) implement civil infrastructure capital into the production function of private firms as an external input. Duggal et al. (1999) incorporate civil infrastructure capital in the production function as part of the technological constraints. These production functions can then be employed to discuss sustainable policies for investment in civil infrastructure and sustainability of the economy.

However, most of the economic models that incorporate civil infrastructure capitals assume that the deterioration rate of the infrastructure capital is exogenously given and constant; the deterioration rate is not considered as a variable. This means that the average reliability of the infrastructure remains constant over the entire time period, being independent of the growing economic states – the reliability remains the same when the economy is in a poor state and in a richer state. However, this is not realistic since the deterioration rate of infrastructure can be dynamically controlled by means of the design and maintenance policies on civil infrastructure. There are only a few research studies available that consider the deterioration rate as a variable. Rioja (2003) proposes a dynamic general equilibrium model that explicitly considers investment into maintenance work of civil infrastructure, thereby incorporating the effect of the maintenance works on the deterioration rate of infrastructure. This model is extended by Kalaitzidakis and Kalyvitis (2004), which endogenizes the decision of budget allocation into both investment in the construction of new infrastructure and investment in maintenance work on existing infrastructure.

The use of these models is a promising way to investigate the optimal reliability level of infrastructure as a function of economic growth, thereby to identify the optimal policies for the design and maintenance work on civil infrastructure in a macroeconomic context. However, the assumptions made in these models are too simplistic in regard to the relations between the amount of investment in maintenance work and the deterioration rate; for instance, the investment in maintenance work at one particular time influence the deterioration rate at the same time but not for the deterioration rate in the future. Realistic models and a methodology that can incorporate engineering knowledge into the models are still missing, and thus need to be developed.

1.4.3. Implication and formulation of sustainability

Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs (Brundtland (1987)). The intuitive implication of this statement seems clear: increasing energy consumption efficiency, less dependence on non-renewable resources, preserving biodiversity, etc. However, when it comes to the formulation of sustainability, there are a huge variety of opinions, approaches, methodologies and philosophies between researchers in different disciplines, and even among researchers within the same disciplines. In this section, instead of identifying the best formulation among them, relevant discussions of three aspects of sustainability in the field of economics are briefly summarized.

The first aspect is the substitutability of different types of capital in production functions. Especially, the substitutability between man-made capital (e.g. physical capital, human capital) and natural capital (e.g. non-renewable resources) is the focus of discussion, see Chapter 4 in Perman et al. (2003). Therein a distinction of the concept of sustainability is made; weak sustainability and strong sustainability. The perspective of weak sustainability is that man-made capital can substitute natural capital, thus a certain production level can be kept by maintaining the level of the sum of both types of capital. On the other hand, the perspective of strong sustainability is that the level of production can be sustained only if natural capital is provided at a certain level. If the strong sustainability perspective is taken, the level of production can be maintained in an infinite time horizon only by exploiting natural capitals indefinitely, which seems unfeasible at least for non-renewable resources. In contrast, based on the weak sustainability perspective, the (feasible) conditions under which a certain level of production and thus consumption can be maintained are derived by Hartwick (1977) and Hartwick (1978).

The second aspect concerns the economic concepts of sustainability; opportunity-based concept or consumption-based concept. The opportunity-based concept considers that the sustainability should be based on the opportunities, i.e. opportunities to use capitals should be maintained. On the other hand, the consumption-based concept assumes that the sustainability is realized as long as the same level of (aggregated) consumption is maintained. Seen in this light, the famous sentence in the Brundtland report: *Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs*, stands for the opportunity-based concept. The opportunity-based concept is also supported by ecologists, since resources which ecologists focus on are primarily renewable, thus the preservation of the opportunities is feasible. Furthermore, the concept fits well with the preservation of biodiversity. One of the proponents of the consumption-based concept is Solow (1986), who argues that *we have no obligation to our successors to bequeath a share of this or that resources. Our obligation refers to generalized productive capacity or, even wider, to certain standards of consumption/living possibilities over time*. Although this distinction poses an important philosophical question, practically it makes little difference in economic models. This is because the economic models presently employed in the discussions on sustainability are so simple that each type of capital is nothing other than an input to the production functions. Consequently, within these economic models the opportunities to use capitals are limited to production; then, to maintain the opportunities is much the same as to maintain the production level, and thus, the consumption level.

Today, no general agreement on the definition and criteria for sustainability is made; a steady increase of consumption or utility over time is often considered as the criterion for sustainability, see e.g. Withagen (1996), in which relevant works that employ this definition are listed. However, this is not a unique criterion, see e.g. Pezzey (1992) and

Pezzey (1997). Some concepts which are widely used and discussed in economics are stated as[6]:

- A sustainable state is one in which utility (or consumption) is non-declining through time
- A sustainable state is one in which resources are managed so as to maintain production opportunities for the future
- A sustainable state is one in which the natural capital stock is non-declining through time.

Other concepts, which originate in ecology, are stated as:

- A sustainable state is one in which resources are managed so as to maintain a sustainable yield of resource services
- A sustainable state is one which satisfies minimum conditions for ecosystem resilience through time.

1.5. Outline of the thesis

The core of the present thesis consists of six chapters. The next five chapters (Chapters 2 – 6) represent research work published or accepted for publication in four peer-reviewed journal papers and a conference paper during the PhD study[7]. Chapter 7 is devoted to illustrating the approach proposed in Chapter 6 with a simplistic example. Each of the chapters focuses on one of the four issues mentioned in the previous section.

Chapter 2 (Paper I) focuses on the treatment of aleatory and epistemic uncertainties in probabilistic assessments of extreme events. This chapter first reviews the general principle for the treatment of these uncertainties. Then, focusing on the probabilistic assessment of extreme events, it is pointed out that the general principle is often violated in practice, and it is shown that such violations can lead to biased assessments of probabilistic characteristics of extreme events. Since a consistent treatment of aleatory and epistemic uncertainties is essential for risk-based decision analysis in general, and the probabilistic assessment of extreme events is especially relevant to the risk assessment of long-term structural performance of infrastructure, the principle presented in this chapter constitutes a basis for the treatment of uncertainties in sustainable policy making for civil infrastructure.

Chapter 3 (Paper II) proposes a method for optimizing decisions for complex engineering systems under constraints. Constrained optimization problems are often encountered in engineering decision analysis, especially where societal preferences

6 From Table 4.2 in Perman et al. (2003).

7 Therein, minor modifications such as grammatical corrections are made. Also, errata in the original papers are, if any, corrected.

must be taken into account. For instance, transport networks have to be designed and maintained by satisfying requirements on human safety over their entire operation periods. An engineering facility may have to satisfy the regulations imposed for environmental protection, e.g. in terms of maximum leakage of harmful biochemical agents. The proposed method employs Bayesian probabilistic networks for the probabilistic representation of the structural performance of complex systems, and generic algorithms for solving constrained optimization problems. Since these techniques are commonly available in terms of software tools, the proposed method is directly facilitated in practical decision situations.

Chapter 4 (Paper III) considers the issue of discounting in the context of intergenerational equity. A large amount of research literature is available on the issue of discounting, focusing on different types of discount factors. Among others, the most relevant discount factors in civil infrastructure projects are the factors of pure-time preference and long-term economic growth. The former concerns the preference of individuals regarding the timing of consumption. The latter is related to the relative wealth of the members of the society at different point in time. Incorporating these two types of discount factors with due consideration of the finite lifespan of individuals, a logically consistent concept for discounting (generation-adjusted discounting) is proposed by Bayer and Cansier (1999). However, the application of the concept requires tedious calculation. Thus, based on the consideration similar but independent from Bayer and Cansier (1999), this chapter proposes a formula for deriving an equivalent discount rate which, if applied to a decision problem with the classical perspective where one decision-maker who is assumed to have an infinite lifespan, yields the same total expected utility as when the decision problem is analyzed in accordance with the consistent consideration of discounting over generations.

Chapter 5 (Paper IV) reformulates optimization problems of civil infrastructure projects from a different perspective. The classical perspective is that the projects should be optimized by minimizing the (discounted) life cycle costs. In this chapter instead the optimization of projects is seen from the perspective of optimal budget allocation. The shift of the perspective naturally introduces costs incurred by the delay of actions which in turn is caused by the lack of a budget. In the reformulated optimization problems, ultimate decision variables to be optimized are the amount of budget that needs to be allocated to individual projects. This perspective is especially useful for societal decision-makers who have to decide on the allocation of limited resources.

In Chapters 3 to 5, the primary objective is to optimize individual civil infrastructure projects. One of the underlying assumptions therein is that decisions made regarding individual projects do not influence long-term economic growth in society, i.e., the economic consequence of the projects is marginal – this assumption is required in order to justify the assumption that the discount factor for economic growth is exogenous, independent of the decisions regarding individual projects. However, whenever this assumption is violated, the marginal perspective mentioned above may be invalid and the non-marginal (macroeconomic) perspective should be chosen. In Chapters 6 and 7, a new conceptual approach for this is proposed and illustrated.

Chapter 6 (Paper V) proposes an approach for how the reliability of infrastructure can be treated in the context of macroeconomics. The proposed approach consists of two steps: (1) defining infrastructure failure by limit state representations; (2) implementing the reliability concept into economic models. The first step takes basis in the structural reliability theory and the second step employs the economic growth theory. Thus, the proposed approach can incorporate knowledge of civil engineering concerning structural performance into economic models. In order to show how the proposed approach can be applied an illustrative example is provided. Therein, a simplistic economy is assumed, which solely depends on civil infrastructure as the production factor and is subject to natural hazards, and the economic growth path is examined as a function of the policy on the design and maintenance of civil infrastructure.

In Chapter 7, the proposed approach is applied to another simplistic economy, and the steady and transition states of the economy are examined as a function of the policy on the design and maintenance of civil infrastructure. By analyzing the steady state a decision principle is derived, which differs from the decision principle adopted in the life cycle cost optimization concept. Furthermore, it is shown that by analyzing the transition state the optimal policy at each point in time depends on the current economic output level.

2. Probabilistic assessment of extreme events subject to epistemic uncertainties (Paper I)

Kazuyoshi Nishijima

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 22.3, Zurich 8093, Switzerland.

Michael Havbro Faber

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 23.3, Zurich 8093, Switzerland.

Marc A. Maes

Schulich School of Engineering, University of Calgary, 2500 University Ave. N.W., Calgary, AB T2N1N4 Canada.

Proceedings of the ASME 27th International Conference on Offshore Mechanics and Arctic Engineering, OMAE2008, Estoril, Portugal, June 15-20, 2008.

Abstract

Over the years the modeling and treatment of aleatory and epistemic uncertainties in probabilistic assessments has repeatedly been an issue of discussion and also some controversy. The philosophical and mathematical aspects may be said to be well appreciated; however, there are cases in practice where principles seem to be violated and frequently the effects of the epistemic uncertainty are treated inconsistently in the probabilistic modeling. The present paper first reviews the general principles for the modeling and treatment of uncertain characteristics subject to both aleatory and epistemic uncertainties. Thereafter, the general principles are applied considering three examples concerning the probabilistic modeling of extreme events; 1) the n-year maximum distribution, 2) the corresponding return period and 3) the exceedance probability in hazard analysis. Through these examples typical inconsistencies made in practical probabilistic assessments are pointed out. The results from the examples are interpreted and discussed from a structural design perspective and from a rational risk-based decision perspective. Finally, a practical solution to avoid the inconsistencies is suggested emphasizing the analogy of the analysis of extreme events with the analysis of portfolios.

2.1. Introduction

The probabilistic modeling of events and not least extreme events forms a crucial corner stone in risk based decision making concerning the design, assessment, inspection and maintenance planning for engineering structures and facilities. The assessment of probabilities can be performed based on probabilistic models that describe the events of interest; extreme wave heights, current and wind velocities, etc. In general, such probabilistic models are established through the joint consideration of knowledge, experience and observations; combining statistical assessments with subjective judgments. Consequently, very often the resulting probabilistic models are associated with not only aleatory uncertainties, i.e. the inherent natural variability associated with the phenomenon of interest but moreover with significant epistemic uncertainties. It is of utmost importance that both of these two contributions to uncertainty are treated correctly in the probabilistic assessments.

In the literature a number of discussions have been made on how uncertainties arising from different sources may be categorized and how these different categories should and/or can be considered in probabilistic risk assessment and risk-based decision making, e.g. Raiffa and Schlaifer (1961), Pate-Cornell (1996), Faber (2003), Wen et al. (2003), Faber and Maes (2005) and Der Kiureghian and Ditlevsen (2007). It can be said that the relevance of epistemic uncertainties in risk assessments is well recognized and also the general principles for modeling and assessing the relevant probabilistic characteristics seem well understood. However, there are still several situations where the general principles are violated in practice. The present paper considers the treatment of aleatory and epistemic uncertainties especially in the probabilistic modeling and assessment of extreme events. The probabilistic modeling of extreme events often requires that several probabilistic models are applied jointly and that some logical framework is assumed for extrapolation of knowledge concerning e.g. the probabilistic characteristics of annual events to the corresponding characteristics of events with much longer return periods, e.g. 100 years. If in this process the aleatory and epistemic uncertainties are inconsistently mixed up the probabilistic characteristics of the extreme events of interest are assessed incorrectly.

The present paper first reviews the general principles for the probabilistic modeling of uncertain characteristics subject to both aleatory and epistemic uncertainties. Thereafter, three examples are considered pointing out in parallel the typical inconsistent assessments often made in practice and the results of a correct assessment following the general principles. Finally, a practical procedure to avoid inconsistent probabilistic assessments of extreme events is presented based on an analogy to the probabilistic modeling and treatment of portfolio loss assessments.

2.1.1. Aleatory and epistemic uncertainties

Without going into detailed and philosophical discussions, it is taken for granted in the present paper that the probability measure is sufficient to represent any type of uncertainty e.g. O'Hagan and Oakley (2004) and that the Bayesian statistics provides a consistent basis for representing both aleatory and epistemic uncertainties, see e.g. De Groot (1970) and Lindley (1980).

Generally, it is understood that aleatory uncertainty reflects the variability of events subject to inherent natural variability and epistemic uncertainty represents imprecise models, lack of data and insufficient knowledge, e.g. Pate-Cornell (1996), Wen et al. (2003) and Der Kiureghian and Ditlevsen (2007). Pate-Cornell (1996) provides a general overview on the treatment of the uncertainties in risk assessment over different engineering applications identifying different levels of analytical sophistication. Therein, the explicit consideration of epistemic uncertainty in risk assessment is qualified as the highest level of risk assessment.

In engineering decision making the treatment and categorization of the two components of uncertainties has received attention for mainly two reasons. The first reason is that the categorization of uncertainties allows for the optimization of resource allocations aiming to reduce uncertainty and thereby to enhance ranking of options for the purpose of risk management; epistemic uncertainty can be reduced by accumulating data and knowledge. In this context the pre-posterior decision analysis provides the theoretical basis, see Raiffa and Schlaifer (1961). The pre-posterior analysis has been extensively applied in the field of engineering in general, e.g. Faber (2003) and Faber and Maes (2005) and in risk-based inspection planning in particular, e.g. Straub and Faber (2005). The second reason is that the epistemic uncertainty often may have a profound effect on the probabilistic characteristics of systems. In Nishijima and Faber (2007a) systems with quasi-identical components subjected to epistemic uncertainties are considered. There it is shown that the epistemic uncertainty can be utilized for the reduction of the uncertainty of a whole system performance by inspecting the states of some of the components in the system. Faber et al. (2007a) considers the effect of epistemic uncertainties on the portfolio loss analyses subject to seismic hazards; epistemic uncertainties concerning the resistance of types of buildings commonly affect all buildings that belong to the same type. Thus, the quantile values of the distribution of failure costs are highly dependent on the extent of the epistemic uncertainties. The present paper is strongly related to the latter considerations as discussed in more detail in the subsequent sections.

2.1.2. Probabilistic modeling approach in practice

Within the framework of probabilistic hazard analysis, the probabilistic modeling of hazards, such as the seismic ground motion, wind speed and wave height, can be established by either pure statistical modeling relying only on available relevant data or by means of engineering probabilistic models which also facilitate for the utilization of subjective information such as experience and physical understanding.

The pure statistical approach has been preferred by classical statisticians since the results of such models are coherent with the frequentistic interpretation of probabilities; there is a one to one correspondence between observations and model predictions. Typically the statistical models are formulated as annual extreme value distributions, and the extreme value theory thus provides the justification for assuming either one of the three extreme value distributions or the generalized extreme value distribution, e.g. Leadbetter et al. (1983) and Coles (2001). This approach may be a reasonable solution for cases where the detailed physical mechanisms that govern the hazard events are not well understood or too complex to represent in a practically manageable effort. However, this approach also has drawbacks; 1) direct observations of extreme events are by definition rare why the parameter estimation of the distributions generally involves large statistical uncertainties (epistemic uncertainty), and 2) the potentially available scientific knowledge and/or engineering experiences cannot be included in the modeling. To overcome these drawbacks, engineering probabilistic approaches have been developed for different types of hazards, which enables one to integrate into the hazard analysis the available knowledge and engineering experience. For instance, in Nishijima and Faber (2007b) hurricane simulation techniques have been developed for wind hazard analysis integrating several probabilistic model components each of which represents individual parts of the involved physical mechanisms, e.g. the transition of hurricanes and development of the pressure fields.

In the pure statistical modeling approach the distinction between epistemic uncertainty and aleatory uncertainty is relatively clear, since the epistemic uncertainty is primarily statistical uncertainty that is involved in the parameter estimation of the distributions (including uncertainty on the choice of distribution family). The epistemic uncertainty can be integrated into the probabilistic assessments within the Bayesian statistical framework, e.g. Coles et al. (2003), although in practice it is often neglected. On the other hand, in the engineering approach taking basis in the Bayesian framework the epistemic uncertainties are associated with each individual probabilistic model components that jointly comprise the probabilistic assessment model in terms of model uncertainty and statistical uncertainty.

As is discussed in more detail later, the integration of aleatory and epistemic uncertainties at the level of the individual probabilistic models may lead to inconsistent assessments of the probabilistic characteristics of extreme events, see Maes and Jordaan (1985) and Maes (1990). This can be seen through a simple example: consider throwing two different dice. One die is a fair die which has six numbers (one to six) and the probability of the outcome of each number is assumed equal to 1/6 (pure aleatory uncertainty). Therefore, the probability that a six comes out in one trial is 1/6. The other die is an unfair die which has an identical number, between one and six, on all six faces, yet the number is unknown. Thus, it is assumed that the probability that the number is i ($i = 1, 2, ..., 6$) is equal to 1/6 (pure epistemic uncertainty). Therefore, the probability that a six comes out in a trial is 1/6, which is the same as with the fair die. Now consider throwing each of two dice 100 times. The probability that the six comes out at least once with the fair die is equal to $1-(1-1/6)^{100} \approx 1$, while the probability that the six comes out at least once with the unfair die remains 1/6. When the different origins and/or types of uncertainty is not identified and differentiated in the probabilistic assessments, it may not be possible to assess the probability of extreme events correctly. Thus, it is of utmost importance to distinguish between aleatory and epistemic uncertainties in the probabilistic modeling of extreme events for both statistical and engineering-based approach.

2.2. General principles for the probabilistic modeling of events subject to aleatory and epistemic uncertainty

This section reviews the general principles for assessing the probabilistic characteristics of events in general and provides remarks which are relevant for the probabilistic assessment of extreme events in particular. The probabilistic models for assessing probabilistic characteristics of extreme events are assumed to have been developed aiming at describing the random nature of phenomena of interest in e.g. offshore engineering. Hence, the probabilistic models specifically focus on the aleatory uncertainties associated with e.g. extreme wave heights. However, due to the lack of data and/or knowledge the developed probabilistic models do not precisely represent the random phenomena of the real world, why epistemic uncertainty is introduced to account for such model uncertainties.

In the context of engineering decision making or reliability assessments the probabilistic modeling problem can in general be represented as a problem involving the expectation operation (in some cases a conditional expectation) over a function $g(\mathbf{X})$ of aleatory random variables $\mathbf{X} = (X_1, X_2, ..., X_n)$ as:

$$E[g(\mathbf{X})] = E_{\mathbf{\Theta}}\left[E_{\mathbf{X}}\left[g(\mathbf{X}) \mid \mathbf{\Theta}\right]\right] \tag{2.1}$$

The random variables $\mathbf{X}$ are characterized by the joint probability distribution function $F_{\mathbf{X}}(\mathbf{x} \mid \mathbf{\theta})$ conditional on the epistemic random variables $\mathbf{\Theta} = (\Theta_1, \Theta_2, ..., \Theta_m)$, which in turn are characterized by the probability distribution function $F_{\mathbf{\Theta}}(\mathbf{\theta})$. Thus, $F_{\mathbf{X}}(\mathbf{x} \mid \mathbf{\theta})$ corresponds to the developed probabilistic model and together with $F_{\mathbf{\Theta}}(\mathbf{\theta})$ constitutes the probabilistic assessment model, see Figure 2.1. From a probability theoretical viewpoint the expectation operation of $g(\mathbf{X})$ may be performed in any manners as long as it is integrated over the domain of the joint probability density function of $(\mathbf{X}, \mathbf{\Theta})$. However, the hierarchical expression on the expectation given by Equation (2.1) is useful especially for probabilistic modeling of extreme events since some of the aleatory random variables often can be assumed to be conditionally independent given the epistemic uncertainties $\mathbf{\Theta}$; this can reduce the computational effort required to evaluate the expectation significantly.

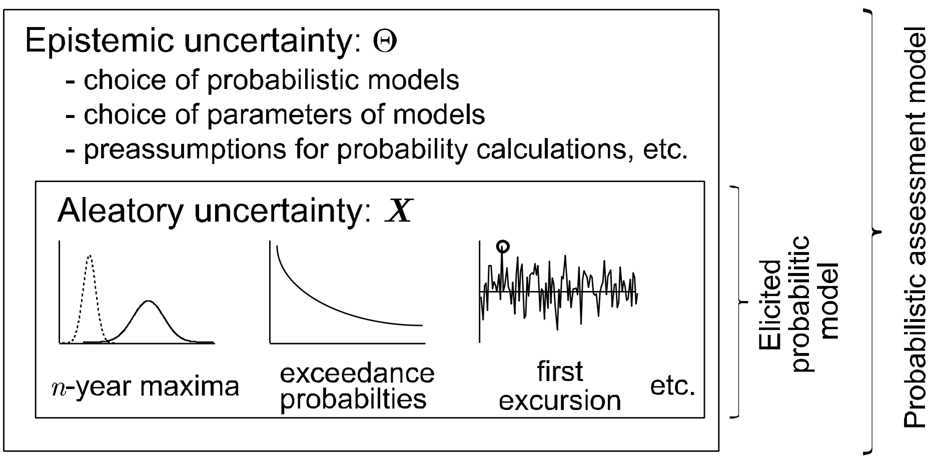


Figure 2.1. Probabilistic assessment subject to aleatory and epistemic uncertainties.

Figure 2.1 illustrates the roles of aleatory and epistemic uncertainties in probabilistic modeling. Probabilistic characteristics of extreme events are first assessed conditional on the epistemic uncertainty $\mathbf{\theta}$ then integrated over possible realizations of epistemic random variables $\mathbf{\Theta}$. The epistemic random variables $\mathbf{\Theta}$ should be interpreted heuristically; the epistemic random variables represents not only the uncertainties of the parameters of distributions but also the likelihood or degree of belief associated with different distribution families and even pre-assumptions for probabilistic calculations etc. The pre-assumptions reflect the modeler's perception of the phenomena of interest, for instance what concern causal relations and boundaries for the considered phenomena. Although these pre-assumptions are often precluded in the probabilistic modeling and simply assumed certain, it should be kept in mind that these may be significant for the probabilistic modeling. It should be also mentioned that the categorization of epistemic uncertainty and aleatory uncertainty is dependent on these pre-assumptions, a process which in itself is subject to the modeler's choice and taste why in a certain sense any assignment of aleatory uncertainties is conditional on factors or variables which are associated with epistemic uncertainty.

2.3. Examples

Three examples are now considered in order to illustrate how the general principle introduced in the previous section might be utilized in practice. Through the examples, pointing out the typical inconsistent probabilistic assessments of characteristics of extreme events which are commonly utilized in engineering design and assessment, the probabilistic models which follow from the application of the general principle are also provided. The discussion on the implications of the results is provided subsequently in Section 2.4.

2.3.1. N-year maxima

The first example considers the derivation of the cumulative distribution of the n-year maxima from the annual maximum distribution. It is pre-assumed that the annual maxima are statistically independent and identically distributed. The cumulative distribution function of n-year maxima can be calculated in accordance with Equation (2.1) by defining:

$$g(\mathbf{X}) = I\left[\max_{i=1,2,..,n}\{X_i\} \le x\right] \tag{2.2}$$

where $I[\cdot]$ is an indicator function that returns the value one if the condition in the bracket is satisfied and zero otherwise and X_i is the i^{th} year maximum. By substituting Equation (2.2) into Equation (2.1) the cumulative distribution function is obtained as:

$$F_{X,n}(x) = \int \{F(x\,|\,\theta)\}^n \, p(\theta)d\theta \tag{2.3}$$

where $F(x\,|\,\theta)$ is the conditional cumulative distribution function of the annual maxima and $p(\theta)$ is the probability density function of the epistemic random variable Θ. The epistemic random variables may be represented by a scalar or a vector. The possible sources of the epistemic uncertainty are the statistical uncertainties when the cumulative distribution function is established by a pure statistical approach and the model and statistical uncertainties when the cumulative distribution function is established based on engineering probabilistic models.

In practice deviations from the general principle are observed. One example for this concerns the utilization of probabilistic hazard maps or load recommendations for risk management purposes. Hazard maps usually provide characteristic values, e.g. quantile values including the effect of the epistemic uncertainty, e.g. in the form of conservatively assessed fractile values or median values of the fractile values relative to the epistemic uncertainties. Based on these characteristic values a distribution function of annual maxima $\tilde{F}(x)$ is established and based on this finally the distribution of the n-year maximum distribution is calculated as:

$$F_{X,n}{}^{*}(x) = \left\{\tilde{F}(x)\right\}^{n} \tag{2.4}$$

Since the annual maximum distribution $\tilde{F}(x)$ that is established utilizing the probabilistic hazard map or load recommendations already contains the effect of epistemic uncertainty, $\tilde{F}(x)$ can be written as:

$$\tilde{F}(x) = \int F(x \mid \theta) p(\theta) d\theta \tag{2.5}$$

Obviously, $F_{X,n}(x)$ and $F_{X,n}{}^{*}(x)$ are in general not identical. Furthermore, for $n > 1$ it can be shown by applying Jensen's inequality that

$$\begin{aligned} F_{X,n}(x) &= E_{\Theta}\left[\left\{F(x \mid \Theta)\right\}^{n}\right] \\ &\geq \left\{E_{\Theta}\left[F(x \mid \Theta)\right]\right\}^{n} = \left\{\tilde{F}(x)\right\}^{n} = F_{X,n}{}^{*}(x) \end{aligned} \tag{2.6}$$

The equality holds if there is no epistemic uncertainty. Thus, for any given quantile the corresponding value is larger when $F_{X,n}{}^{*}(x)$ is employed instead of $F_{X,n}(x)$; n-year maximum events are overestimated when $F_{X,n}{}^{*}(x)$ is employed.

A numerical example is shown to illustrate the degree of the difference between $F_{X,n}(x)$ and $F_{X,n}{}^{*}(x)$, considering the case of wind hazard analysis. For this purpose it is assumed that the conditional annual maximum wind speed X follows the Gumbel distribution as:

$$F(x \mid \theta) = \exp\left(-\exp\left(-\alpha(x-\theta)\right)\right) \tag{2.7}$$

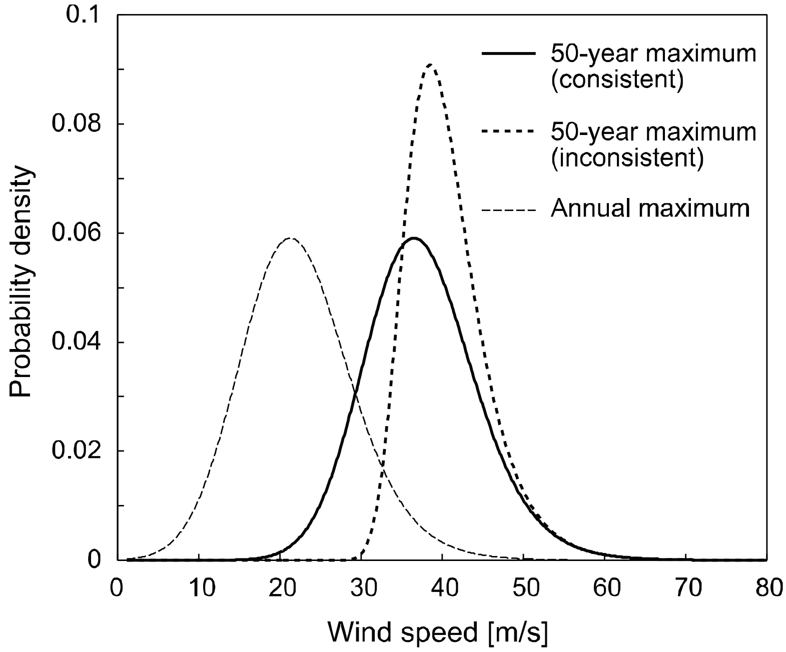


Figure 2.2. Probability density functions of maximum wind speed.

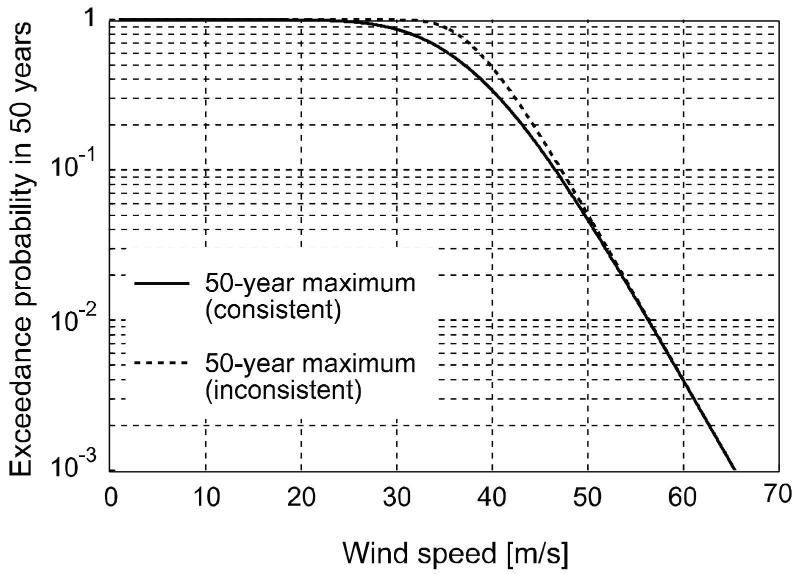


Figure 2.3. Exceedance probabilities of 50-year maximum wind speed.

where θ represents the epistemic uncertainty and $\alpha = 0.257$ (this corresponds to the standard deviation of 5 [m/s] given θ). The epistemic uncertainty represented by the random variable Θ is assumed to follow the Normal distribution with mean and standard deviation being equal to 20 [m/s] and 5 [m/s] respectively. Figure 2.2 shows the assessed probability density functions of the 50-year maximum in accordance with Equations (2.3) (denoted as "consistent") and (2.4) (denote as "inconsistent") respectively. It is seen that the probability density function looks significantly different and that the mean value of the 50-year maximum wind speed is overestimated when it is evaluated using Equation (2.4).

Figure 2.3 shows the corresponding exceedance probabilities of the 50-year maximum wind speed. Whereas the (inconsistent) Equation (2.4) overestimates the exceedance probability at the range between 10^{-1} and 1, the tendency diminishes for the range of lower probabilities. These results should be appreciated depending on the context as will be discussed further in the subsequent section.

2.3.2. Return period

In this example first the definition of the return period of events is briefly revisited and thereafter the effect of epistemic uncertainties on the return period is assessed.

The return period may be defined as the expected value of the arrival time of the event of interest, see e.g. Benjamin and Cornell (1970). Assuming that the probability of occurrence of an event in a Bernoulli sequence of trials is p, then the arrival time follows the geometric distribution. The expected value of the arrival time $E[T]$ is then calculated as $1/p$. When the event is characterized by its intensity X, e.g. a given wind speed or a given precipitation, the probability p is represented by the cumulative distribution function $F(x)$ of the maximum within a given period (e.g. one year). Thus, the return period is a function of the intensity x and may be written as:

$$E[T(x)]^* = \frac{1}{1-F(x)} \tag{2.8}$$

However, when the epistemic uncertainty represented through Θ is involved, the assumption of independence between the intensities at different times does not hold, even if this might be a reasonable assumption considering observations from the real world; the intensities are independent only conditional on the realization of epistemic uncertainty θ. Thus, the return period defined by Equation (2.8) should be reformulated as:

$$E[T(x)] = E_\Theta\left[\frac{1}{1-F(x\,|\,\Theta)}\right] = \int \frac{p(\theta)}{1-F(x\,|\,\theta)}d\theta \tag{2.9}$$

where $F(x\,|\,\theta)$ is the conditional cumulative distribution function on the epistemic uncertainty θ and $p(\theta)$ is the probability density function of θ. This formulation is coherent with the general principle given in Equation (2.1).

Probabilistic engineering models are often employed where the cumulative distribution function of the maximum intensity within a given reference period is established by combination of probabilistic models that represent the natural random nature (aleatory uncertainty) yet subject to model/statistical uncertainties (epistemic uncertainty), as e.g. in hurricane simulation for wind hazards analyses. The cumulative distribution function obtained in this manner already considers the epistemic uncertainty and can thus be written as:

$$\tilde{F}(x) = \int F(x \mid \theta) p(\theta) d\theta \tag{2.10}$$

The return period is often assessed by combining Equations (2.8) and (2.10) as:

$$E[T(x)]^{**} = \frac{1}{1-\tilde{F}(x)} \tag{2.11}$$

This is obviously not the same as Equation (2.9) and it can be shown by applying Jensen's inequality that:

$$\begin{aligned} E[T(x)] = E_{\Theta}\left[\frac{1}{1-F(x \mid \Theta)}\right] &\geq \frac{1}{1-E_{\Theta}\left[F(x \mid \Theta)\right]} \\ &= \frac{1}{1-\tilde{F}(x)} = E[T(x)]^{**} \end{aligned} \tag{2.12}$$

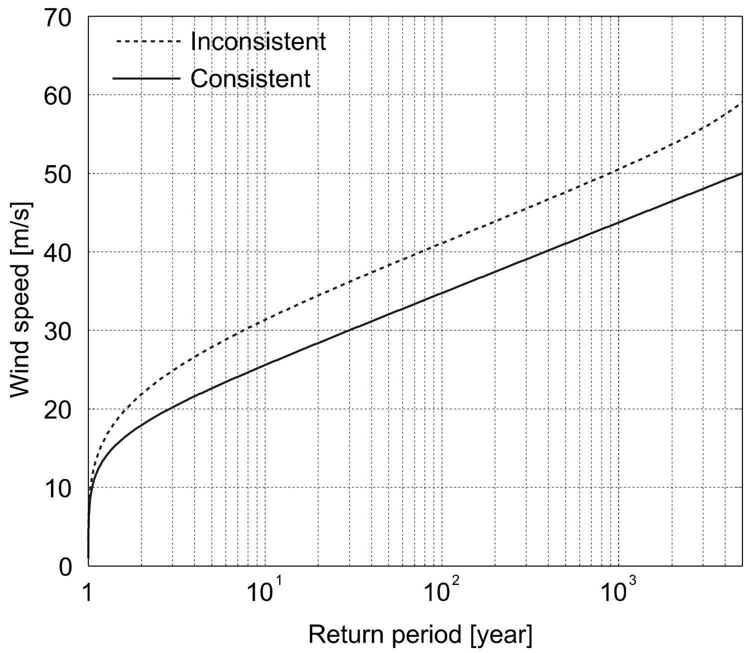


Figure 2.4. Comparison of return periods.

The equality in Equation (2.12) holds if there is no epistemic uncertainty; in that case $E[T(x)]$ and $E[T(x)]^{**}$ coincide. From this inequality, it can be said that the return period assessed by Equation (2.11) underestimates the expected arrival time.

In Figure 2.4 the results of a probabilistic assessment of the relation between extreme wind speeds and corresponding return periods are shown. Based on the same assumption as in the first example it is seen that the application of Equation (2.9) and (2.11) respectively result in different return periods. For instance, based on the application of Equation (2.11) a wind speed of 40 m/s corresponds to a return period of 80 years, whereas the correct return period using Equation (2.9) is in fact 400 years.

2.3.3. Hazard curve

In the following example it is investigated how hazard curves, i.e. the relationships between the exceedance probabilities for a given uncertain phenomenon represented by the random variable X should be calculated according to the general principle given by Equation (2.1). For illustrative purposes an example considering an earthquake hazard analysis is selected and for simplicity, only one seismic zone is considered in this example.

Seismic hazard analysis aims at assessing the probability of exceedance of any given seismic hazard intensity x for a specified reference period, e.g. one year, (seismic hazard curve). In the assessment of this probability several assumptions and probabilistic models are required; e.g. the occurrence of earthquake in the seismic zone, the magnitude of the earthquake, the distance between the epicenter and the site for which the hazard analysis is performed and the so-called attenuation law that relates the relevant parameters and the seismic hazard intensity. Essentially such assumptions and probabilistic models involve epistemic uncertainty due to the imperfection of the postulated models and scarce data available for estimating parameters in the models. Whereas the presence of epistemic uncertainty in general is appreciated and some epistemic uncertainties are considered correctly, other epistemic uncertainties are often inconsistently considered. Examples of cases where epistemic uncertainties are consistently accounted for include the epistemic uncertainty associated with the choice of attenuation law and the choice of the range of the possible magnitudes. For instance, a typical attenuation law is represented in the form of $X = \varepsilon \cdot g(a,b,c,...)$, where X denotes the hazard index, e.g. peak ground motion, and $a,b,c,...$ represent the relevant parameters in the attenuation law, e.g. magnitude and distance from the epicenter, and ε represents the residual term. Different attenuation laws are proposed by different experts. These differences are often ascribed to expert judgments, for each of which a probability is assigned in order to incorporate the different expert judgments into one unified seismic hazard curve. Such incorporations are consistent with Equation (2.1), since the inner expectation in Equation (2.1) corresponds to each hazard curve conditional on each expert judgment and the outer expectations correspond to the uncertainties associated with the expert judgments. An example of the inconsistent consideration of the epistemic uncertainties corresponds to the residual term of the attenuation law. The random variable ε can be considered to involve epistemic uncertainty, since obviously this uncertainty can be reduced by updating using data on the seismic hazard intensity from the site for which the seismic hazard analysis is performed.

Denote by $q(x|\theta)$ the probability that the seismic hazard intensity X exceeds x given the occurrence of an earthquake. The probability $q(x|\theta)$ is conditioned by the epistemic uncertainty θ, e.g. the uncertainty associated with the attenuation law. Hence, the probability that the seismic hazard intensity X exceeds x may be written in accordance with Equation (2.1) as:

$$P[X > x] = \int \left(1 - \exp\left[-\nu q(x|\theta)\right]\right) p(\theta) d\theta \tag{2.13}$$

Here it is assumed that the occurrence of an earthquake follows a Poisson process with intensity ν. However, in some practices the probability is calculated as:

$$P[X > x]^{*} = 1 - \exp\left[-\nu \int q(x|\theta) p(\theta) d\theta\right] \tag{2.14}$$

where the conditional probability of the seismic hazard intensity given the occurrence of an earthquake is first marginalized by integrating over the epistemic uncertainty θ, thereafter the assumption of the Poisson process is applied to calculate the probability of exceedance x; Equation (2.14) is inconsistent with the general principle given by Equation (2.1). Generally, Equation (2.14) does not provide the same value as Equation (2.13), although if ν is small enough both equations can be approximated as $\nu \int q(x|\theta) p(\theta) d\theta$. In this sense, the evaluation of the probability with Equation (2.14) can be seen as a numerical approximation and this may justify the use of Equation (2.14) in practice. Furthermore, by applying Jensen's inequality, it can be shown:

$$\begin{aligned} P[X > x] &= \int \left(1 - \exp\left[-\nu q(x|\theta)\right]\right) p(\theta) d\theta \\ &= E_{\Theta}\left[1 - \exp\left[-\nu q(x|\Theta)\right]\right] \\ &\leq 1 - \exp\left[-\nu E_{\Theta}\left[q(x|\Theta)\right]\right] \\ &= P[X > x]^{*} \end{aligned} \tag{2.15}$$

A similar discussion may apply to cases where non-Poisson processes are assumed for the occurrence of earthquake and for cases where two or more seismic zones are considered.

2.4. Discussion

Three examples considering the n-year maximum distribution, the return period and the exceedance probability respectively have been considered. For each of these examples typical inconsistent treatments of epistemic uncertainties found to occur in practical applications have been considered and analyzed. The results from these examples should be interpreted corresponding to the contexts: structural design in practice and optimal decision making. In the context of structural design in practice the results of the examples may be understood such that the inconsistent probabilistic assessments often made in practice are conservative and hence can be justified. Furthermore, the inconsistent probabilistic assessments are in general less complicated compared with the consistent assessments, since they allow for incorporation of the epistemic uncertainties at earlier stages of the assessments. However, in the context of optimal risk-based decision making the inconsistent probabilistic assessment should be circumvented as it leads to sub-optimal decisions.

The first example reveals that the information provided in typical hazard maps and load recommendations are not sufficient to use directly in the context of optimal decision making, since they do not differentiate the sources of uncertainties; hence the distributions of maximum values for a given reference period cannot be correctly established. The second example shows that the return period that provides the basis for structural design as well as for validation of the established probabilistic models based on observations does not correspond to the expected value of the arrival time. Therefore, the return period assessed by Equation (2.11) should not be used for these purposes. The third example justifies the seismic hazard analyses presently made in practice in a numerical sense, although it is important to realize that the analyses are not conceptually consistent with the general principle for the probabilistic assessments.

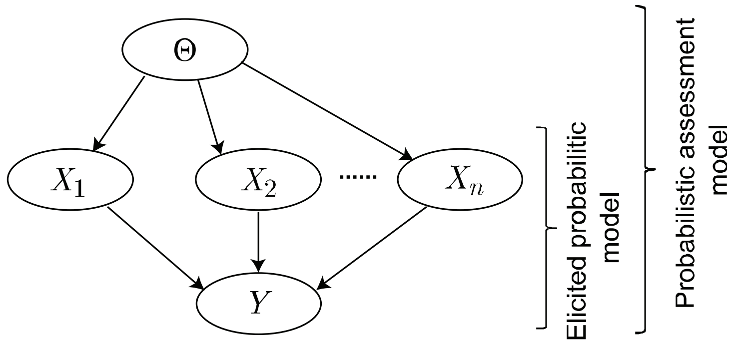


Figure 2.5. Graphical representation for interrelation between random variables.

In order to circumvent inconsistent probabilistic assessments, a causal representation, e.g. through Bayesian probabilistic networks Jensen (2001) may be useful just for the purpose to explicitly understand the interrelations between all random variables in the probabilistic assessment models. Figure 2.5 shows the causal representation corresponding to the first example, where Θ represents the epistemic uncertainty, X_i represents the annual maximum wind speed at the i^{th} year and Y represents the 50-year maximum wind speed ($Y = \max_{i=1}^{n} X_i$). When the interrelation between all the variables are explicit, it is clear that X_i ($i = 1, 2, ..., n$) are not independent but are instead exchangeable, see Maes and Jordaan (1985). Thereby it is also clear how to calculate the marginal distribution of Y according to the general graphical representation theory e.g. Jensen (2001). It is worthwhile mentioning that the random variables X_i can be seen as the components of a temporarily distributed portfolio with the analogy of a spatially distributed portfolio – the graphical representation in Figure 2.5 can be also understood to represent a spatially distributed portfolio, the component of which are subject to epistemic uncertainty, see Faber et al. (2007a). Then, it is obvious that the probabilistic characteristics of identical components X_i are subject to epistemic uncertainty Θ that simultaneous affects all the components. In this regard the distinction between aleatory and epistemic uncertainty might be useful simply to make clear which variables affects other variables. For completeness the incorporation of epistemic uncertainty in seismic hazard analysis as discussed in the third example is shown in detail in the Appendix.

2.5. Conclusion

The present paper first provides general principles on how aleatory and epistemic uncertainties should be considered in the probabilistic modeling and assessments for risk based decision making. Focusing on the probabilistic modeling of extreme events, several inconsistencies often made in practical probabilistic assessments for extreme events are pointed out; i.e. the n-year maximum distribution, the return period and the exceedance probability in hazard analysis. For the considered examples it is shown that such inconsistent probabilistic assessments overestimate the probabilistic characteristics of the extreme events. From the perspective of structural design it can be seen as a conservative assessment and thus may be justified. However, from the perspective of optimal decision making the inconsistent assessments lead to sub-optimal decisions and should thus be avoided.

2.6. Appendix

The exceedance probability is calculated assuming that the occurrence of earthquakes over time follows a Poisson process as:

$$\begin{aligned} P[X>x] &= \sum_{k=1}^{\infty} P\left[N=k \cap \max_{i=1,2,..,k} X_i > x\right] \\ &= \sum_{k=1}^{\infty} P\left[\max_{i=1,2,..,k} X_i > x \mid N=k\right] P\left[N=k\right] \end{aligned} \tag{2.16}$$

where N is the number of occurrence of earthquake and X_i is the peak ground intensity due to the i^{th} earthquake. When the intensities can be assumed independent, the calculation proceeds as:

$$\begin{aligned} P[X > x] &= \sum_{k=1}^{\infty} \left[1-(1-\tilde{q}(x))^k\right] \frac{\nu^k}{k!} e^{-\nu} \\ &= 1-\exp\left[-\nu \tilde{q}(x)\right] \end{aligned} \tag{2.17}$$

where $\tilde{q}(x)$ is the probability that the intensity exceeds x given the occurrence of an earthquake and ν is the occurrence rate. This is the same form as Equation (2.14) using that $\tilde{q}(x) = \int q(x \mid \theta) p(\theta) d\theta$. However, when epistemic uncertainties which affect all X_i are present, the calculation should proceed as:

$$\begin{aligned} P[X > x] &= \int \sum_{k=1}^{\infty} P\left[\max_{i=1,2,..,k} X_i > x \mid N = k, \theta\right] P\left[N = k\right] p(\theta) d\theta \\ &= \int \sum_{k=1}^{\infty} \left(1-\left(1-q(x \mid \theta)\right)^k\right) \frac{\nu^k e^{-\nu}}{k!} p(\theta) d\theta \\ &= \int \left(1-\exp\left[-\nu q(x \mid \theta)\right]\right) p(\theta) d\theta \end{aligned} \tag{2.18}$$

which is equivalent to Equation (2.13). In this way, the fact that the epistemic uncertainty affects the ground motion intensities for all earthquakes over time plays a crucial role.

3. Constrained optimization of component reliability in complex systems (Paper II)

Kazuyoshi Nishijima

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 22.3, Zurich 8093, Switzerland.

Marc A. Maes

Civil Engineering Department, Schulich School of Engineering, University of Calgary, 2500 University Avenue N.W., Calgary, Canada T2N1N4.

Jean Goyet

Bureau Veritas, Marine Division, Research Department, 17 bis Place des Reflets, La Defense 2, 92400 Courbevoie, France.

Michael Havbro Faber

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 23.2, Zurich 8093, Switzerland.

Structural Safety, Vol. 31, pp. 168-178, 2009,
doi: 10.1016/j.strusafe.2008.06.016.

Abstract

The present paper proposes an approach for identifying target reliabilities for components of complex engineered systems with given acceptance criteria for system performance. The target reliabilities for components must be consistent in the sense that the system performance resulting from the choice of the components' reliabilities satisfy the given acceptance criteria, and should be optimal in the sense that the expected utility associated with the system is maximized. To this end, the present paper first describes how complex engineered systems may be modelled hierarchically by use of Bayesian probabilistic networks and influence diagrams. They serve as functions relating the reliabilities of the individual components of the system to the overall system performance. Thereafter, a constrained optimization problem is formulated for the optimization of the component reliabilities. In this optimization problem the acceptance criteria for the system performance define the constraints, and the expected utility from the system is considered as the objective function. Two examples are shown: (1) optimization of design of bridges in a transportation network subjected to an earthquake, and (2) optimization of target reliabilities of welded joints in a ship hull structure subjected to fatigue deterioration in the context of maintenance planning.

Keywords

Constrained optimization, complex system, acceptance criteria, Bayesian probabilistic network, influence diagram.

3.1. Introduction

Typically engineered systems are complex systems comprised of geographically distributed and/or functionally interrelated components, which through their connections with other components provide the desired functionality of the system expressed in terms of one or more attributes. This perspective may indeed be useful for interpreting and modelling a broad range of engineered systems ranging from construction processes over water and electricity distribution systems to structural systems. One of the characteristics of engineered systems is that, while the individual components may be standardized in regard to quality and reliability, the systems themselves often cannot be standardized due to their uniqueness. The performance of the systems will depend on the way their components are interconnected to provide the functionalities of the systems as well as on the choice of reliabilities of their components. Thus, the design and maintenance of such systems effectively concern the requirements to the reliability of their components, which can be translated from given requirements to the attributes of the performance of systems in accordance with the way the components are connected.

Due to the complex nature of the problem, modelling and optimization of such systems generally require that different levels of analyses provided by different experts and supported by data are integrated interdisciplinary. Taking basis in engineered structures, at component level physical failure mechanisms may be analyzed, such as yielding, fracture and corrosion. The component failure modes now constitute the building stones for the development of systems failure modes including the formation of failure modes for sequences of sub-systems, for which the corresponding consequences may be assessed. An optimization of the target reliability for components of a given system, i.e., a system with a given interrelation between its components, must take basis in such analyses. Seen in this light, it is useful to hierarchically establish models for complex engineered systems which accommodate for the integration of the different levels of analyses. Such a hierarchical approach may also prove to be beneficial as a mean of communication between professionals representing the expertise required for the modelling of the performance of the different types of components, sub-systems and systems.

The present paper addresses the problem outlined in the foregoing in the context of a hierarchical system modelling developed for risk assessment of engineered systems by the Joint Committee on Structural Safety (Faber et al. (2007b)), where taking basis in structural systems a framework is formalized in regard to how the hierarchical system model can be established and then applied to optimize the reliability for components of structures based on specified requirements to the acceptable risks for the considered structural system.

The present paper first provides a short summary of available techniques on the modelling of complex systems. Following this, a general approach for the optimization of the reliability of system components with given criteria to the acceptable system risk is proposed. The proposed approach is composed of three steps; (1) adaptation of Bayesian probabilistic network and influence diagram representation for hierarchical system modelling, (2) linking of acceptance criteria for system level to component level through the Bayesian probabilistic networks and the influence diagrams, and (3) optimizing the target reliabilities of individual components. The original contribution of the presented approach is the effective use of the commonly available techniques, i.e. Bayesian probabilistic networks, influence diagrams and generic algorithms for constrained optimization problems. The approach suggested allows for the assessment of optimal target reliabilities for the individual components of systems for which the risk acceptance criteria are specified in regard to the system performance. The proposed approach is most useful in cases where (1) the components that constitute the system or the sub-system can be categorized into groups with identical probabilistic characteristics and/or (2) the components are hierarchically related. Finally, two illustrative examples are provided. The first example addresses the design of bridges in a transportation network subject to earthquake hazards. Through this example the individual steps of the proposed approach are explained. The second example considers a floating production storage and offloading unit (FPSO), which constitutes a typical complex engineered system. In this example, the target reliabilities of welded joints subject to fatigue deterioration in the framework of inspection and maintenance planning are optimized with given acceptance criteria for the performance of the ship hull structure as a whole.

3.2. Problem setting

3.2.1. Modelling of complex systems

The requirements to the probabilistic modelling of complex engineered systems in the context of risk based decision making concern the consistent and tractable representation of the physical characteristics of the considered system and the appropriate detailing to facilitate the assessment of the benefit associated with different decision alternatives. In addition, of course the modelling should facilitate an efficient analysis of the probabilities and consequences required for the ranking of decision alternatives. Fault tree analyses comprise classical techniques for the representation and analysis of systems failure modes, see e.g. Vesely et al. (1981). Assuming that components in a system have only two states (failure and success) and that the component failures are statistically independent, the probability that a predefined state of the system (top-event) occurs may be quantitatively assessed (Bobbio et al. (2003)). Fault tree analyses have been applied to a variety of fields, e.g., among others, risk assessments of nuclear power plants (USNRC (1975) and USNRC (1990)) and the reliability analysis of control systems for gas turbine plants (Bobbio et al. (2003)). Fault tree analysis is from a technical perspective relatively simple, and hence in many ways attractive, however, for the same reason subject to important limitations. Among these limitations, the difficulty in representing dependencies between basic events as well as the problems associated with updating based on new information should be mentioned. Bayesian probabilistic networks (BPNs) and influence diagrams (IDs) seem to provide an interesting and promising alternative to the classical techniques for system analysis. Any fault trees can be mapped into BPNs as is shown in Bobbio et al. (2001). The BPN approach for systems modelling has been utilized for the analysis of structural systems, see e.g. Baker et al. (2007). The applications of BPNs in the context of hierarchical modelling are briefly reviewed in the subsequent section.

When modelling the performance of systems it is important to consider temporal aspects. Petri Nets provide a powerful platform based on which temporal dependencies associated with e.g. repair or replacement actions which may provoke cyclic references to states of the components in the model can be accounted for, see Volovoi (2004). However, the evaluation of the reliability of a given system through a Petri Net often takes basis in Monte Carlo simulation, which in general requires a considerable amount of computational effort, and the generic algorithms applicable to a broader range of problems are not yet available. BPNs are not immediately appropriate for the representation of cyclic effects; however, by introducing time slices in a BPN (so-called dynamic BPN), BPNs may also be applied for such analysis. Several efficient time slice BPN algorithms have been developed for calculating probabilistic characteristics of state variables of BPNs, e.g. expected values and conditional probabilities, see e.g. Kjaerulff (1995). It should be noted that a dynamic BPN representation is equivalent to a Markov chain representation (Smyth (1997)).

Another approach for the probabilistic modelling and analysis of complex systems is proposed by Der Kiureghian and Song (2008). In this approach, the probability of an event of interest (related to the system performance) is formulated as a sum of the probabilities of the mutually exclusive combinations of the component states that govern this event. Upper and lower probability bounds on the system performance are calculated based on an out-crossing formulation and using linear programming techniques. Moreover, it is shown in Der Kiureghian and Song (2008) that by aggregating several components as "super-components" and applying the linear programming method in a hierarchical way, the approach provides reasonable probability bounds on the system performance with a manageable computational effort. However, the applied scheme for component aggregation affects the efficiency of the computation and the width of the obtained probability bounds. An optimization of the aggregation scheme in principle requires trial and error, although general guidelines are provided in Der Kiureghian and Song (2008).

3.2.2. Bayesian hierarchical modelling

The applications of the Bayesian hierarchical models range from, for instance, sociology, biology, environmental studies to engineering. In experiments in sociology, e.g., experiments for studying school effect in educational research, it is difficult to control all the experimental conditions. Ignoring dependences between the uncontrolled experimental conditions at different levels - for the example of school effect, student level, classroom level and school level - and applying simple statistical analysis are proven to produce misleading results as is summarized in Raudenbush and Bryk (1986). Raudenbush and Bryk (1986) propose a hierarchical approach for studying school effect taking basis in the Bayesian multi-level linear model proposed by Lindley and Smith (1972). It provides a flexible statistical tool for estimating how variations in school policies and practices influence educational processes, whereby the different levels of interrelations are taken into account. Environmental sciences face similar situations where due to the complex nature of processes and interactions between systems, observing all the relevant variables that may influence the process of interest is not realistic. Furthermore, it is difficult to realize the identical conditions in different experiments. Thus, the comprehensive use of data obtained for different conditions is necessary for efficiently estimating the parameters of the models, see Clark and Gelfand (2006). In these contexts the Bayesian hierarchical models are employed in such ways that the causal relation or interrelation of variables at different levels in whole systems are first established based on scientific knowledge without specifying the probabilistic characteristics of the variables or assuming weak prior distributions. The parameters of the variables are then estimated or updated using observed data. Other applications of Bayesian hierarchical models can be found in the area of pattern categorization/recognition, see e.g. Li and Pietro (2005) and George and Hawkins (2005). Due to the characteristics of the applications of the models for the pattern categorizations or recognitions, it is important that these models allow for promptly updating the parameters in the models for a broader range of objects. To this end, flexible representations and systematic learning algorithms which the BPN approach provides are extensively utilized. The Bayesian hierarchical approach has been applied also for engineered complex systems. Among others, Johnson et al. (2002) apply the hierarchical model for estimating the reliability of missile systems, where the fault tree analysis is extended using the Bayesian approach to accommodate the integration of available expert knowledge and data.

Emphasizing the difference of the use of the Bayesian hierarchical models, the present paper appreciates the fact that input-output relations of phenomena in engineering at different levels are often quantitatively available in probabilistic terms. For instance, given the geometry and material properties of an engineered component, it is possible to calculate the probability of failure of the component using data and by physical modeling and analysis techniques, e.g. finite element methods. Fatigue deterioration can be probabilistically modelled for given environments, using physical models and data, see Straub (2004). As the events of interest such as component failure and fatigue degradation are subject to given circumstances, which themselves might be associated with uncertainty, the probabilities of the events are appropriately represented in terms of conditional probabilities. Therefore, in the context of modelling of complex engineered systems, the main focus is how the system can be hierarchically modelled using these conditional probabilities of components at different levels.

As observed in the above the applications of Bayesian hierarchical models are rather diverse. However, all Bayesian hierarchical models utilize generic algorithms developed for estimating parameters and/or obtaining conditional or posterior distributions. The algorithms themselves are indifferent to the contexts where the Bayesian hierarchical models are employed.

3.2.3. Optimization of engineering decisions under constraints

It is often the case that the optimization of decisions for engineering systems must be performed under constraints. These constraints are typically given a priori to the decision problems in terms of acceptance criteria regarding risks and/or practical operational limitations. Acceptance criteria are generally defined for the attributes of the performance of systems considering the consequences due to possible failures. Recent design codes e.g. ASCE7-98 (2000) provide acceptance criteria in terms of minimum requirements to structural performance. The Joint Committee on Structural Safety (JCSS (2001b)) recommends different target reliabilities for engineered structures in accordance with the different magnitude of the consequence of failure as well as the relative cost of safety measures. Also, safety to personnel must be considered. Recently, a general principle for evaluating the acceptability of a life saving measure has been proposed using the concept of life quality index (LQI), e.g. Nathwani et al. (1997) and Rackwitz (2002). Based on the LQI principle it is possible to optimize and specify requirements for the reliability of engineered systems based on the costs of improving their reliability. Additionally, several practical constraints, e.g., available budget, cost-benefit ratios and allowable environmental impacts, may be given for projects involving design and maintenance of engineered systems. Together with acceptance criteria given from normative perspectives, these exogenously given constraints constitute important boundary conditions for the optimization of the performance of engineered systems.

A number of approaches have been proposed for optimizing decisions under constraints in engineering (e.g. Royset et al. (2003), Guikema and Pate-Cornell (2002) and Salazar et al. (2006)). Thereby, one of the central issues is how the optimization process can be transformed in such ways that it allows for the utilization of commonly available techniques for the probability calculations as well as for numerical optimization. Royset et al. (2003) propose algorithms for reliability-based optimal design problems with which the required calculations of reliability and optimizations are completely decoupled, hence, allowing for a flexible choice of the optimization algorithm and the reliability calculation method. Guikema and Pate-Cornell (2002) propose a method for the optimization whereby the performances of engineered systems are related discontinuously to decision variables. These approaches are in fact highly sophisticated and also efficient in the treatment of some optimization problems. However, for the same reason they may be cumbersome to apply in practical situations where complex engineered systems are of interest, since different levels of models established by different experts must be reformulated to fit the format which these approaches require. To overcome this difficulty Bayesian probabilistic network and influence diagram representations are employed in the present paper as is described in the following sections.

3.2.4. Objective of proposed approach

The acceptance criteria mentioned in the foregoing may be seen to constitute the boundary conditions, which any engineered system must satisfy during its service life. The present paper takes the standpoint that the acceptance criteria for systems are a priori given. This situation is often the situation encountered in practice. The goal of the present paper is to establish an approach for the optimization of the target reliability for components of systems for given system performance requirements in terms of acceptance criteria, by minimizing life cycle costs for the design and operation of the system, or more generally by maximizing the service life expected utility.

3.3. Proposed approach

3.3.1. Hierarchical system modelling with Bayesian probabilistic networks

A hierarchical system modelling for complex systems facilitates the representation of complex systems at an early stage of risk analysis, e.g. at the concept evaluation, but may also serve to optimize the final design as well as the management of the risk during operation. Hierarchical BPN models appear suitable as a platform for modelling complex systems, since they provide a causal and mind mapping representation of the system characteristics and functionalities. In Figure 3.1 it is illustrated how the system functions are represented in terms of a hierarchical aggregation of components and their interrelations. At the same time the requirements to the system performance may be disaggregated into reliability performance requirements for the components. In what follows, the proposed approach is explained in accordance with Figure 3.1.

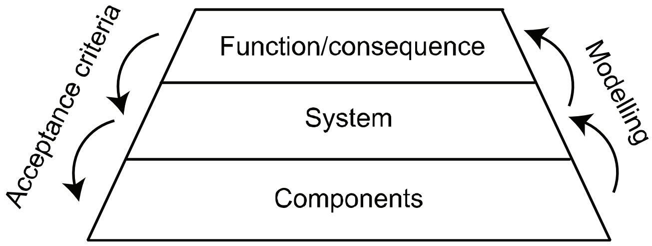


Figure 3.1. Hierarchical modelling and translation of acceptance criteria.

Let A and $E=(E_1,E_2,...,E_n)$ denote the sets of possible actions and possible states of a system respectively. The combination of $a\in A$ and $\mathbf{e}\in E$ specifies the joint probability conditional on the action $P[\mathbf{e}\,|\,a]$ and the consequences $\mathbf{C}(a,\mathbf{e})=(C_1(a,\mathbf{e}),C_2(a,\mathbf{e}),...,C_m(a,\mathbf{e}))$. In general these quantities are the functions describing how the components and the sub-systems in the system are interconnected. However, in the following it is assumed that the interconnectivity is fixed. Note that the consequences $\mathbf{C}(a,\mathbf{e})$ may be a vector when two or more attributes of the system performance are considered, e.g. financial cost, fatalities and damages to the qualities of the environment. It is assumed that the consequences $\mathbf{C}(a,\mathbf{e})$ can be represented as an attribute-wise sum of the consequences $\mathbf{C}_A(a)$ associated with action a and the consequences $\mathbf{C}_E(\mathbf{e})$ associated with event $\mathbf{e}$, namely

$$\mathbf{C}(a,\mathbf{e})=\mathbf{C}_A(a)+\mathbf{C}_E(\mathbf{e}) \tag{3.1}$$

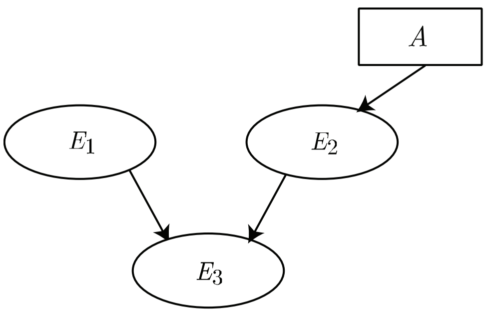
A Bayesian probabilistic network is a probabilistic model representation in terms of a directed acyclic graph that consists of nodes representing uncertain state variables, so-called chance nodes and edges that logically link the nodes, and conditional probability assignments, see Figure 3.2 for example, and see e.g. Jensen (2001) for general introduction. An influence diagram (ID) is an extension of a Bayesian probabilistic network that includes so-called decision nodes and utility nodes in a graph in addition to chance nodes. Using the chain rule for Bayesian probabilistic networks (Jensen (2001)), the joint probability $P(E \mid a)$ can be decomposed as

$$P(E \mid a) = \prod_i P(E_i \mid pa(E_i), a) \tag{3.2}$$

where $pa(E_i)$ is the parent set of E_i. From Equation (3.2) it can be seen that the joint probability $P(E \mid a)$ can be built up by conditional probabilities. Any marginal probabilities of the states of the subset of E can be efficiently calculated from the joint probability $P(E \mid a)$ with generic algorithms and software tools commonly available, see the appendix of Korb and Nicholson (2004). For the BPN shown in Figure 3.2, the parents of E_3 are the nodes E_1 and E_2, and the node E_2 is a function of action A. The joint probability is then written as

$$P(E \mid a) = P(E_3 \mid E_1, E_2) P(E_1) P(E_2 \mid a) \tag{3.3}$$

Each term in Equation (3.3) thus the joint probability is fully characterized by the conditional probability tables shown in Figure 3.2.



$P(E_1)$

e_1^1	p_1
e_1^2	p_2
⋮	⋮

$P(E_2 \mid A)$

	a_1	a_2	····
e_2^1	q_{11}	q_{12}	····
e_2^2	q_{21}	q_{22}	····
⋮	⋮	⋮	

$P(E_3 \mid E_1, E_2)$

	e_1^1		····
	e_2^1	e_2^2	····
e_3^1	r_{11}	r_{12}	···
e_3^2	r_{21}	r_{22}	···
⋮	⋮	⋮	

Figure 3.2. Example of a BPN and conditional probability tables.

Let $\mathbf{F}(\mathbf{C}, P) = (F_1(\mathbf{C}, P), F_2(\mathbf{C}, P), ..., F_l(\mathbf{C}, P))$ denote a vector function of $\mathbf{C}(a, \mathbf{e})$ and $P(E \mid a)$. For instance, the expected total cost, may be one of the attribute of a system performance to be considered, and is written as one element of $\mathbf{F}(\mathbf{C}, P)$ as

$$F_i(\mathbf{C}, P) = \sum_{\mathbf{e} \in E} C_i(\mathbf{e}, a) P(\mathbf{e} \mid a) \tag{3.4}$$

where $C_i(\cdot,\cdot)$ represent the cost. The probability that the damage to environmental quality exceeds a given threshold c_{acc} may be another element of $\mathbf{F}(\mathbf{C},P)$ and is written as

$$F_j(\mathbf{C},P) = \sum_{\mathbf{e}\in E} I\left[C_j(\mathbf{e},a) > c_{acc}\right] P(\mathbf{e}\,|\,a) \tag{3.5}$$

where $C_j(\cdot,\cdot)$ represents the environmental damage and $I[\cdot]$ is the indicator function, which returns unity if the condition in the bracket is satisfied and zero otherwise. Such environmental damages may be represented e.g. in terms of release volumes, the geographical release extent and/or temporal release extent of agents. The conditional expected value of the number of fatalities given the state $E_m = e_m$ may be other element of $\mathbf{F}(\mathbf{C},P)$ and is written as

$$F_k(\mathbf{C},P) = \frac{\sum_{\mathbf{e}'\in E\setminus E_m} C_k(a,(e_m,\mathbf{e}'))P((e_m,\mathbf{e}')\,|\,a)}{\sum_{\mathbf{e}'\in E\setminus E_m} P((e_m,\mathbf{e}')\,|\,a)} \tag{3.6}$$

where $C_k(\cdot,\cdot)$ represents the number of fatalities and $\mathbf{e}'\in E\setminus E_m = \{E_1, E_2, ..., E_{m-1}, E_{m+1}, ..., E_n\}$. Note that any functions represented in terms the elements of $\mathbf{F}(\mathbf{C},P)$ can be systematically calculated by the algorithms developed for the analyses of BPNs and IDs when the state variables $E = (E_1, E_2, ..., E_n)$ and their interrelations and the (conditional) probabilities corresponding to the interrelations of the variables are defined in an ID, see e.g. Jensen (2001). Thus, the remaining task for developing models for engineered complex systems is to represent the physical understanding, the relevant experience and the data available at different hierarchical levels in terms of (conditional) probabilities of states of variables or in terms of decision nodes or utility nodes, and then link them together. Thereby, the general characteristic that engineered systems are comprised and built up by components, which are standardized by codes and industrial standards in regard to quality and reliability may add value to the use of object-oriented BPN representations. This special type of BPN models allows for creating classes of BPNs, which are representative for sub-systems that have identical characteristics, see e.g. Bangso et al. (2003) and Bangso and Olesen (2003).

3.3.2. Objective function and constraints

Having established the hierarchical system model in terms of IDs, the objective function such as service life utility or expected total cost may be assessed from the ID as a function of the chosen action utilizing the functional representation of $\mathbf{F}(\mathbf{C},P)$ as shown in the previous section, i.e.:

$$u(a) = F_1(\mathbf{C}(a,\cdot), P(\cdot\,|\,a)) \tag{3.7}$$

Acceptance criteria are typically defined in regard to the functionality or performance of the considered system measured in terms of risks and/or probability of failure. Since the design and maintenance of a system usually specifically addresses the components of the system, it is of interest how the acceptance criteria for the components may be derived from the acceptance criteria specified for the system performance. Thus, the optimization of reliabilities for components in a system constitutes an inverse problem, see Figure 3.1. The acceptance criteria for the system performance can be related to the target reliabilities for the components using the function type of $\mathbf{F}(\mathbf{C}, P)$ as is shown in the previous section as

$$F_i(\mathbf{C}(a,\cdot), P(\cdot\,|\,a)) \le c_i, \quad (i = 2,3,...,m) \tag{3.8}$$

where F_i $(i = 2,3,...,m)$ represent the functions on the ID calculating the quantities for which the acceptance criteria for the system are defined, and c_i are acceptance levels for the corresponding quantities.

3.3.3. Optimization of actions for components of complex system

Since several combinations of target reliabilities for different components in a system may satisfy the prescribed acceptance criteria for the system, the optimal combination of target reliabilities for components may be identified as the combination of the target component reliabilities associated with action a which maximizes the expected utility u using Equations (3.7) and (3.8) formulated in accordance with the previous sections as

$$\begin{aligned}&\text{Maximize } u(a) = F_1(\mathbf{C}(a,\cdot), P(\cdot\,|\,a)) \text{ s.t.}\\ &F_i(\mathbf{C}(a,\cdot), P(\cdot\,|\,a)) \le c_i,\ (i = 2,3,...,m)\end{aligned} \tag{3.9}$$

Since the functions F_i $(i = 1,2,...,m)$ are readily calculated, the problem is reduced to a standard non-linear constrained optimization problem for which efficient algorithms are available, see e.g. Press et al. (1988).

3.4. Example 1

This example considers the simple optimization of the design of bridges subject to earthquake hazards. The aim of this example is to explain in detail how the proposed approach may be applied in practical situations. The bridges b_1, b_2 and b_3 geographically connect the location a with c, and thus constitute the system components in a transportation network system, see Figure 3.3. It is assumed that the state of the system is fully described through the combinations of the states of the three bridges, and hence, the failures of e.g. the road sections besides the bridges in the network are not considered. The system failure is assumed to be defined as the joint failures of all three bridges. The objective function to be minimized is the expected discounted total cost, which consists of the initial cost and the expected cost associated with the failures of bridges. The acceptance criteria are assumed to be given for (1) the expected number of fatalities in the system given that an earthquake occurs as 10, and (2) the conditional probability that the system fails given that an earthquake occurs as 1%. The life time considered in the design of the bridges is 100 years, and it is assumed that an earthquake occurs at most once in the system's life time. The discounting rate applied for evaluating the future costs is assumed equal to 3% per annum.

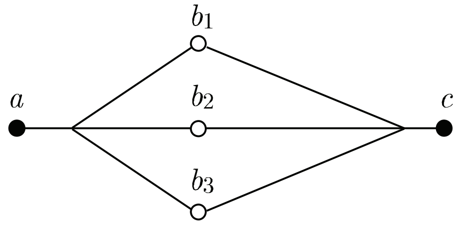


Figure 3.3. Transportation network system.

3.4.1. Model description

The earthquake hazard is modelled in the earthquake class BPN as is shown in Figure 3.4 (left). It consists of five nodes, namely, "Scenario", "Time", "V1", "V2" and "V3". The node "Scenario" contains different possible earthquake scenarios with corresponding probabilities. The term scenario may refer to an earthquake occurring at different seismic zones and different faults, or more specifically, different combinations of the values of ground motion intensities at different locations. The latter corresponds to the cases where the joint probability density of ground motion intensities at different sites is identified by seismic hazard analyses and thereafter the joint probability density is discretized into a finite number of probabilities corresponding to the intervals of the ground motion intensities at different sites. When the different combinations of the values of ground motion intensities are taken as the identifiers of the scenarios, the spatial correlations between the intensities at different locations can be suitably considered in the earthquake hazard model. In this example, however, for illustrative purposes only one scenario "eq1" is considered.

The node "Time" specifies the probability of the yearly discretized time T when the scenario eq1 occurs. T is assumed to follow a geometric distribution with an occurrence probability for each year given as $\nu\Delta t = 0.01$. The nodes "V1", "V2" and "V3" represent the logarithms of the peak ground accelerations (cm/s^2) at the locations where the bridges b_1, b_2 and b_3 are to be built, and are assumed to follow normal distributions given the scenario eq1 with the parameters shown in Table 3.1.

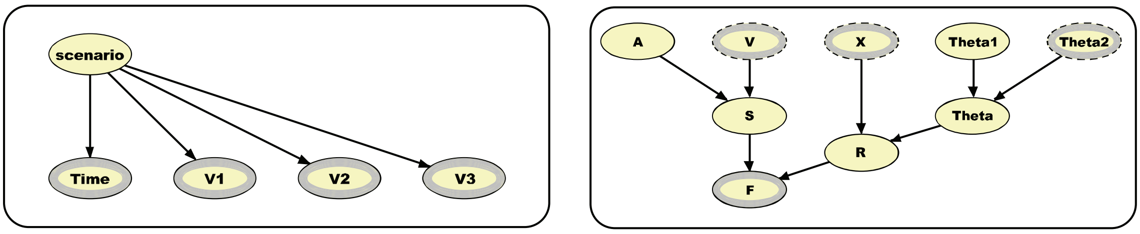


Figure 3.4. Classes of BPNs for Earthquake hazard (left) and for Bridge (right).

Table 3.1. Assumed distributions of nodes in BPNs and ID.

Variables	Distributions	Bounds
Earthquake class BPN		
Scenario	P[Scenario=eq1]=1	
V1|eq1	Normal (ln200, 0.5)	[0, 9]
V2|eq1	Normal (ln300, 0.5)	[0, 9]
V3|eq1	Normal (ln400, 0.5)	[0, 9]
Time|eq1	Geometric (0.01)	[1, 100]
Bridge class BPN		
A	Normal (ln2, 0.1)	[0, 2]
Theta1	Normal (ln1, 0.1)	[-0.5, 0.5]
ID for transportation network system		
Theta2	Normal (ln1, 0.1)	[-0.5, 0.5]
X1,X2 and X3 given design alternative a_1	Normal (ln600, 0.1)	[0, 9]
X1,X2 and X3 given design alternative a_2	Normal (ln800, 0.1)	[0, 9]
X1,X2 and X3 given design alternative a_3	Normal (ln1000, 0.1)	[0, 9]

(Normal (μ,σ) abbreviates the normal distribution with the mean μ and the standard deviation σ, and Geometric (p) abbreviates the geometric distribution with occurrence probability p. The geometric distribution is discretized by the interval of 1 and the Normal distributions are discretized by the interval of 0.1 when implemented into the conditional probability tables in the BPNs. The last column shows the upper and lower bounds in the corresponding conditional probability tables.)

When the probabilistic characteristics are implemented into the conditional probability table in BPNs they have to be discretized. The intervals and the upper and lower bounds must be chosen carefully assuring the efficiency and accuracy of the discretization. They are chosen in this example as shown in Table 3.1. Note that the BPN in Figure 3.4 (left) assumes that "V1", "V2" and "V3" are conditionally independent given the scenario. The nodes "Time", "V1", "V2" and "V3" (surrounded by the bold line) are output nodes, and are connected to other nodes in the BPN for the transportation network system, Figure 3.5.

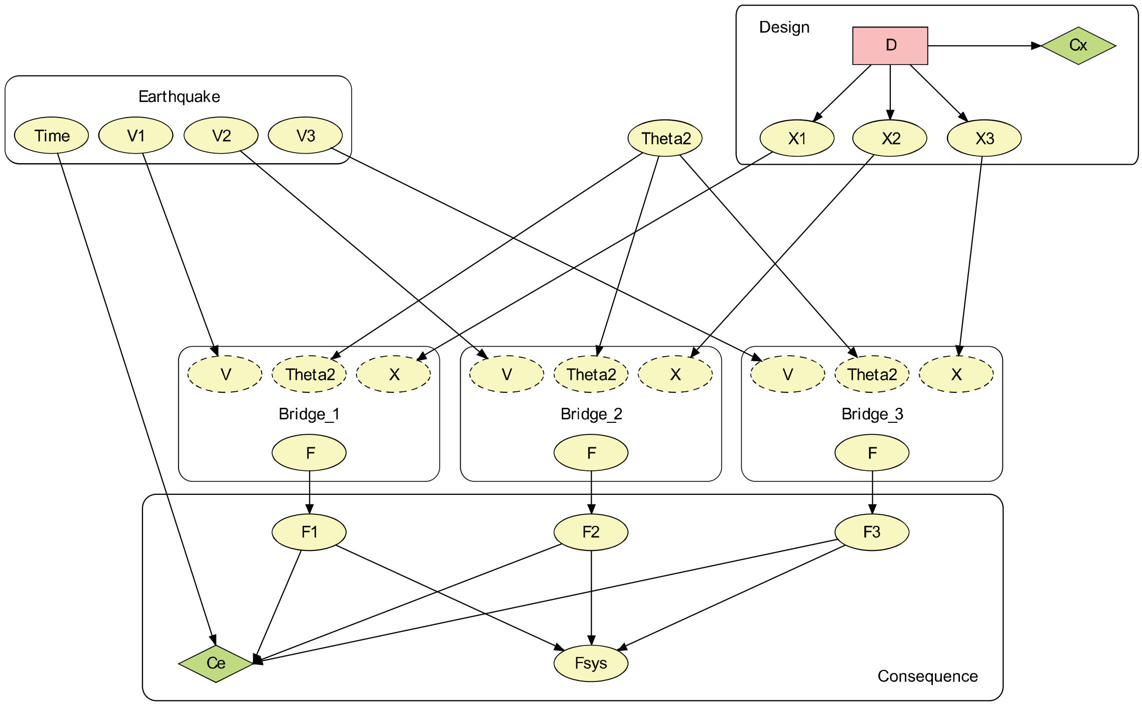


Figure 3.5. ID for transportation network system (cost).

The bridges are modelled in the Bridge class BPN as shown in Figure 3.4 (right). The bridges b_1, b_2 and b_3 are assumed to be identically modelled through the Bridge class BPN. However, the different probabilities in the input nodes "V", "X" and "Theta2" (highlighted with bold dashed line) facilitate the differentiation between the resistances of the bridges and the corresponding probabilities of failure. In the Bridge class BPN, S denotes the load effect, which is represented by

$$S = V + A \tag{3.10}$$

where A represents the logarithm of the soil amplification factor. A is assumed to follow a normal distribution with the parameters given in Table 3.1. The resistance R of the bridge is modelled as

$$R = X + \Theta = X + (\Theta_1 + \Theta_2) \tag{3.11}$$

where X specifies the design of the bridges and Θ represents the uncertainties associated with the resistance of the bridge. Θ can be decomposed into two types of uncertainties, Θ_1 and Θ_2. Θ_1 is the uncertainty associated with individual realizations of bridges, and can be assumed independent between the different bridges, whereas Θ_2 denotes the common uncertainty that affects all realizations of bridges thus introduces the statistical dependence. For example, uncertainty on material geometry or uncertainties associated with construction work may belong to the former type of uncertainty. Modelling and statistical uncertainties belong to the latter type of uncertainty. The assumed probabilistic characteristics of Θ_1 and Θ_2 are shown in Table 3.1. The failure of a bridge, which is defined as the event $R < S$, is denoted by the Boolean node "F", and the probability of failure is expressed as

$$P[F = 'true'] = P[R < S] \quad (3.12)$$

The node "F" is the output node from the Bridge class BPN and is utilized for the assessment of consequences in the ID, see Figure 3.5. Figure 3.5 shows the ID for the whole transportation network system. "Earthquake" is an instance of the Earthquake class BPN, and "Bridge_1", "Bridge_2" and "Bridge_3" corresponding to b_1, b_2 and b_3, respectively, are instances of the Bridge class BPN, for which only input and output nodes are shown. The node "Fsys" represents system failure, which is connected with the nodes "F1", "F2" and "F3" representing the individual failures of the bridges b_1, b_2 and b_3, respectively. These are required for checking if the acceptance criterion is satisfied for the conditional probability of system failure given that an earthquake occurs. The node "Theta2" specifies the probability distribution of the common uncertainty Θ_2, see Table 3.1. Finally, the decision node "D" represents the set of design alternatives for the three bridges. Three design alternatives a_1, a_2 and a_3 are considered for each bridge, hence, there are $3^3 = 27$ possible actions in the decision node. The nodes "X1", "X2" and "X3" represent the probability distribution of state of the bridges b_1, b_2 and b_3 respectively, corresponding to the choice of the design alternatives, see in Table 3.1. For each action, the corresponding initial cost is defined in the utility node "Cx" whose values are shown in Table 3.2. The utility node "Ce" defines the discounted failure costs for all combinations of the states of the three bridges for each year up to 100 years. The failure costs assumed in the example are shown in Table 3.3. From the utility nodes "Ce" and "Cx" the expected discounted total cost is calculated. Similarly, the expected number of fatalities in the system given that an earthquake occurs can be calculated with a similar ID as the one shown in Figure 3.6. In the figure input and output nodes of the instances of the class BPNs (earthquake class, bridge class, design class and consequence class) are abbreviated. The summary of the magnitudes of the consequences are given in Table 3. Failure costs and fatalities shown in the tables should be considered as the expected values over possible consequences given the states of the bridges when an earthquake occurs. In practice the development of the table requires that the consequences must be analyzed for all possible combinations of the states of all bridges in the network. While it requires considerable efforts, it allows for flexibility considering the significance of each bridge in the network, e.g. consideration of the topology of network.

Table 3.2. Initial costs.

Design alternative	Initial cost (Monetary unit)
Design alternative a_1	10
Design alternative a_2	11
Design alternative a_3	12

Table 3.3. Failure costs and fatalities.

	State of Bridge							
Bridge 1	NF				F			
Bridge 2	NF		F		NF		F	
Bridge 3	NF	F	NF	F	NF	F	NF	F
Failure cost (Monetary unit)	0	10	10	50	10	50	50	200
Fatality	0	10	10	20	10	20	20	30

(Failure costs are not discounted. F and NF are abbreviations for failure and no failure, respectively.)

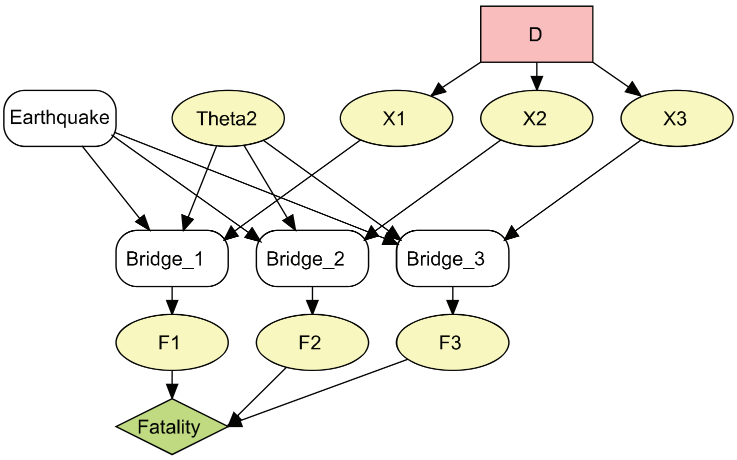


Figure 3.6. ID for transportation network system (fatality).

3.4.2. Results

The expected discounted total costs, the expected number of fatalities and the probabilities of system failure given that an earthquake occurs for the 27 possible actions are calculated using the established IDs. The result is shown in Figure 3.7. At the bottom of the figure the correspondence between the actions and the combinations of the design alternatives for the three bridges is also shown. The optimal action consistent with the two acceptance criteria regarding the expected number of fatalities and conditional probability of system failure given the occurrence of an earthquake is identified as action 25 (design alternative a_3 for the bridges b_1 and b_2, and design alternative a_1 for the bridge b_3); action 17 results in the minimum expected discounted total cost, but it does not satisfy the acceptance criteria. The strategy behind action 25 may be interpreted as follows; considering the non-linear relation between the number of failed bridges and the failure costs, a sound strategy may be to avoid, by all means, the simultaneous failures of the three bridges in an economically efficient way, which may be realized with higher reliabilities for one or two of the three bridges and comparatively low reliability for the other bridge(s). Since the earthquake hazard is smallest for bridge b_1, the highest reliability of the system can be realized most efficiently through bridge b_1 and be realized relatively efficiently for the bridge b_2, by adopting the design alternative a_3 for the bridges b_1 and b_2; corresponding to the highest design resistance in the three design alternatives. At the same time, by accepting a relatively higher failure probability for bridge b_3, the expected discounted total cost can be reduced. This becomes clearer by comparing action 25 with action 9, which is composed of the same set of design alternatives but applied for different bridges, i.e., a_1 for the bridge b_1 and a_3 for the bridges b_2 and b_3. Action 9 requires the same initial cost as action 25, and results in almost the same amount of the expected discounted total cost, but significantly high conditional probability of system failure given an earthquake. This strategy seems tricky, and may not be considered in practical situations where typically the resistances of structures may be designed in a proportional way to the magnitudes of hazards. However, from a system optimization point of view, this is the best strategy that satisfies the acceptance criteria given for the system. It should be noted that in practical situations decision makers might accept slightly higher costs to further reduce the risk of fatalities (e.g. Action 27 instead of Action 25 in this example). However, if the objective function and the constraints are established to fully represent the decision maker's preference, such a subjective decision may lead to sub-optimal decisions.

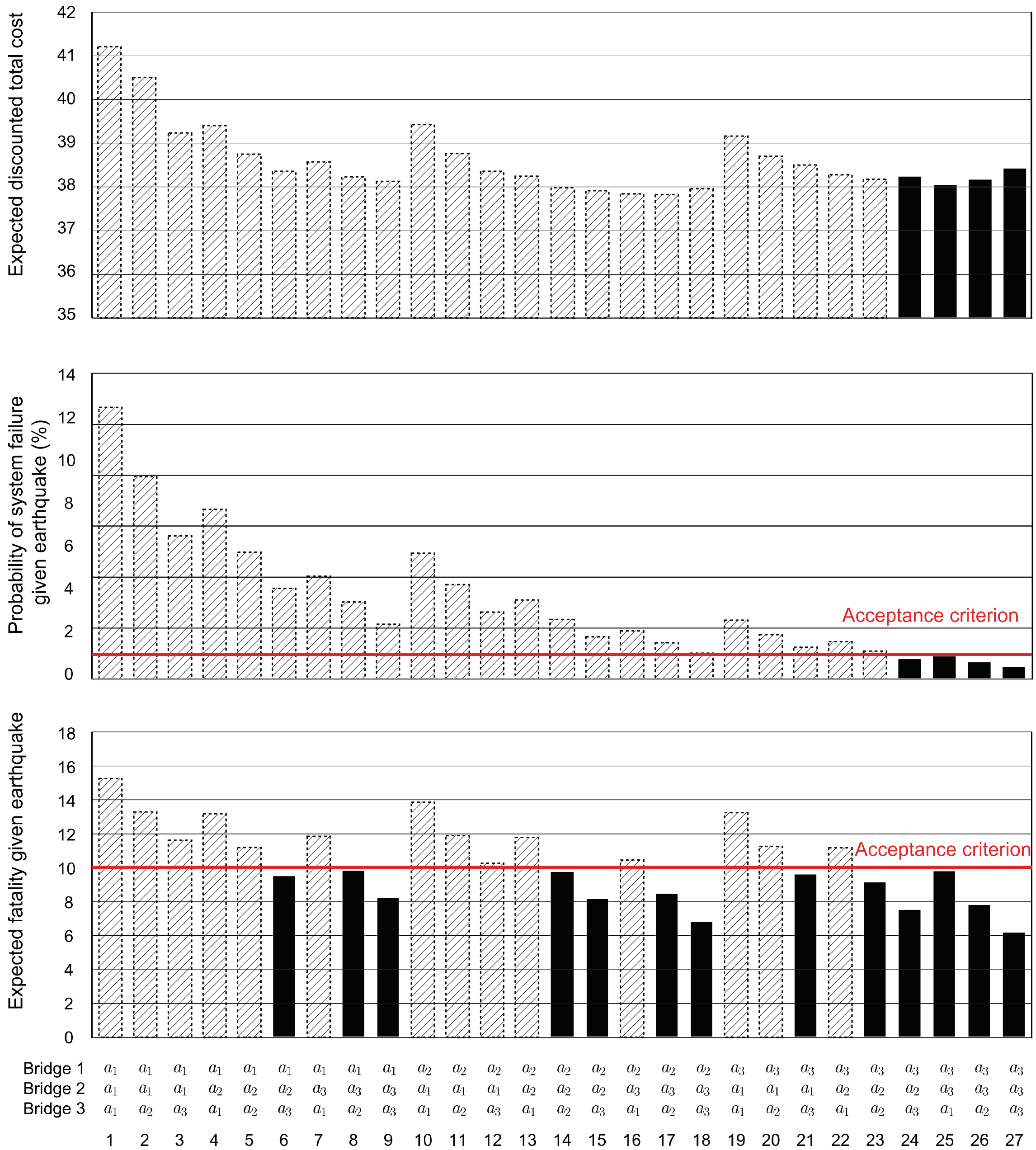


Figure 3.7. Expected discounted total cost, and expected fatality and probability of system failure given that an earthquake occurs.

3.4.3. Discussion

The hierarchical Bayesian approach provides a clear perspective of how the whole system should be built up using the modules representative of different levels of analyses. In this example, the transportation network system can be built up with four modules, i.e., earthquake module represented by the earthquake class BPN, a bridge module represented by the bridge class BPN, a design module and a consequence module, see Figure 3.5. These modules can be built up separately, whereas the interfaces between the modules must be specified. Such a module oriented modelling in the hierarchical Bayesian approach not only enhances the integration of the knowledge of different experts, experience and data available at different levels, but also increases the productivity of risk assessments, since the modules are re-useable.

Updating of the probabilistic characteristics in BPNs is of practical use, although this aspect is not emphasized in the example. For instance, when the data on the damage states of the bridges and the load effects are obtained after the occurrence of an earthquake, the uncertainties associated with the resistance of the bridges can be updated by conditioning the corresponding nodes. Hence, the updated probability can be used for future risk assessment.

While only a small number of discrete action alternatives are considered in this example, there are other cases where a large number of discrete action alternatives or continuous action alternatives are to be considered. In such cases it is not feasible to perform the ID analysis for every action, thus adaptation of efficient algorithms for solving optimization problems under constraints are needed. In this context, IDs serve as the function in the process of calculating the value of the expected utility and the values of the quantities for which acceptance criteria are defined which then in turn can be implemented into optimization algorithms. In the next example, it is shown how this may be realized using commonly available software tools.

3.5. Example 2

Optimal reliability for components in Floating Production Storage and Offloading Units (FPSOs) subject to fatigue deteriorations is considered in this example. The main function of FPSOs is to produce and store oil at offshore oil fields with given requirements to reliability in production and safety to persons and environment. Typically considered events of system failure for FPSOs are:

- Loss or damage of ship due to loss of buoyancy or explosions/fires.
- Loss of production due to reduced functionality.
- Loss of lives due to foundering or explosion/fires.
- Leaks and other damages to the quality of the environment.

Considering the hull as an assembly of components, the hull may be considered to comprise an assembly of tanks tied together with deck plates, tank partitions, and bottom and side plates. The individual components are furthermore stiffened by girders and web frames to ensure a sufficient structural integrity of the hull, see Figure 3.8. The corresponding hierarchical model representation is shown in Figure 3.9.

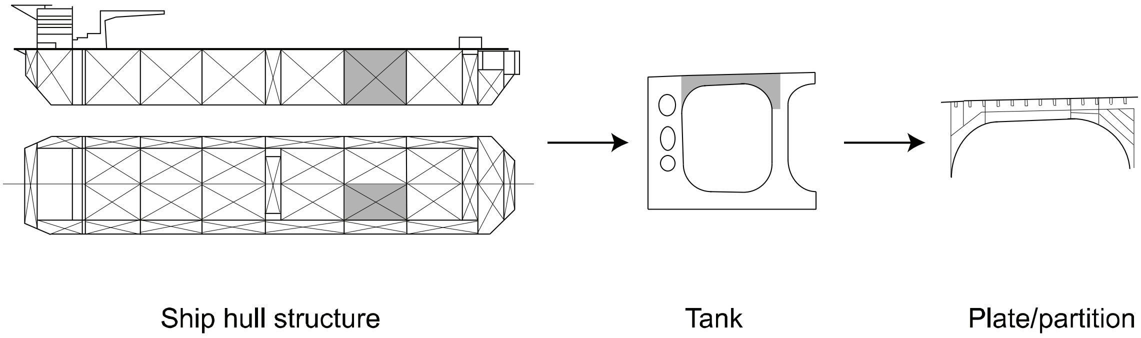


Figure 3.8. Hierarchy of ship hull structure considered.

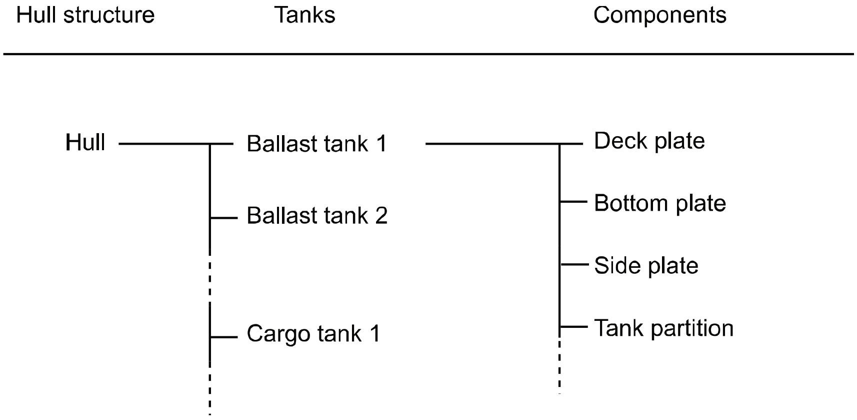


Figure 3.9. Hierarchical modelling of hull structure.

The hull components as described above have basically two functions, namely, to ensure that the overall ship has a sufficient structural integrity and provide the means for containing cargo and ballast. Failure of the components of the hull at this level can be assumed as the events of:

- Loss of or reduced structural integrity.
- Loss of containment due to explosion.
- Leaks of the individual tanks.

Considering now the individual components as outlined in the above these may be viewed upon as assembly of plates connected by welded joints. Failure of these components may lead to:

- Crack or pit through plate thickness.
- Reduced overall plate thickness.
- Joint stiffness reduction or failure.

Thus, the losses or damages at component level may lead to the hull failure or undesired economic and environmental losses as well as loss of lives given the way the components are interconnected. The problem in this example is to optimize the target reliabilities for the welded joints in plate and tank partition components given the requirements to the functionality/consequence of the ship hull, e.g. the probability of hull failure. It is emphasized in this example how commonly available software tools can be used in accordance with the proposed approach. For this purpose a software tool is developed using Hugin ® for BPN/ID representation and Microsoft Excel ® (hereafter Excel) for the optimization algorithm as well as the user interface. In the subsequent section, the overview of the software tool development is illustrated.

3.5.1. Optimization of target reliability for welded joints in components

The developed software tool provides an easy interface to obtain the optimal target reliabilities for welded joints subjected to fatigue deterioration. Excel is used as a platform for integrating the various computational modules and storing information required for calculations. The Excel platform is linked dynamically to the Hugin ActiveX server (hereafter Hugin). In order to use the software tool the user has to define, through Hugin files, the BPNs corresponding to the hierarchical model of the hull structure as described above. The outputs, i.e. optimized target reliabilities for all welded joints, are written into the Excel file.

In Figure 3.10, the illustration of the hierarchical Bayesian representation of the ship hull structure is given. Two BPNs in the top of the figure represent the performances of tanks. The tank performances are characterized by the states of the plates that constitute the tanks. As is described above, at this level the possible consequences due to component failures are capacity reduction, explosion and environmental damage due to leaks. The ID in the bottom of the figure concerns how the component failures may propagate and lead to further consequences at system level. Here, three attributes of the consequences are identified, i.e. economic loss, loss of lives and environmental damage measured in terms of leak intensities. These BPNs and ID are interconnected as shown in the figure. In the entire ID the conditional probability tables are assumed established with the help of experts, see e.g. Figure 3.11 (which is the conditional probability table for node "Explosion_1" as implemented into a Hugin file), whereas the nodes that represent the components serve as root nodes whose probabilities are represented in terms of unconditional probabilities, which are derived from the target reliabilities for welded joints in each components. Therefore, by changing the target reliabilities for the welded joints which are set in the Excel file, the unconditional probabilities for the components are changed accordingly. In turn, the corresponding probabilistic characteristics, e.g. expected total cost or probability of ship hull failure are changed and stored in the Excel file, see Figure 3.12. This process is made automatically through ActiveX. The design and service life maintenance cost for the different welded joints is in general a function of the target reliability in regard to fatigue failure, and this is implemented as a VBA code in the Excel file. For the assessment of the relationship between the reliability of the welded joints subjected to fatigue failure and the service life cost, the iPlan software described in Straub and Faber (2006) may be utilized. Finally, the optimal target reliabilities for welded joints are obtained using the Solver add-in provided in Excel – target reliabilities correspond to "changing cells", and acceptance criteria for the ship hull correspond to "constraints" in the Solver add-in.

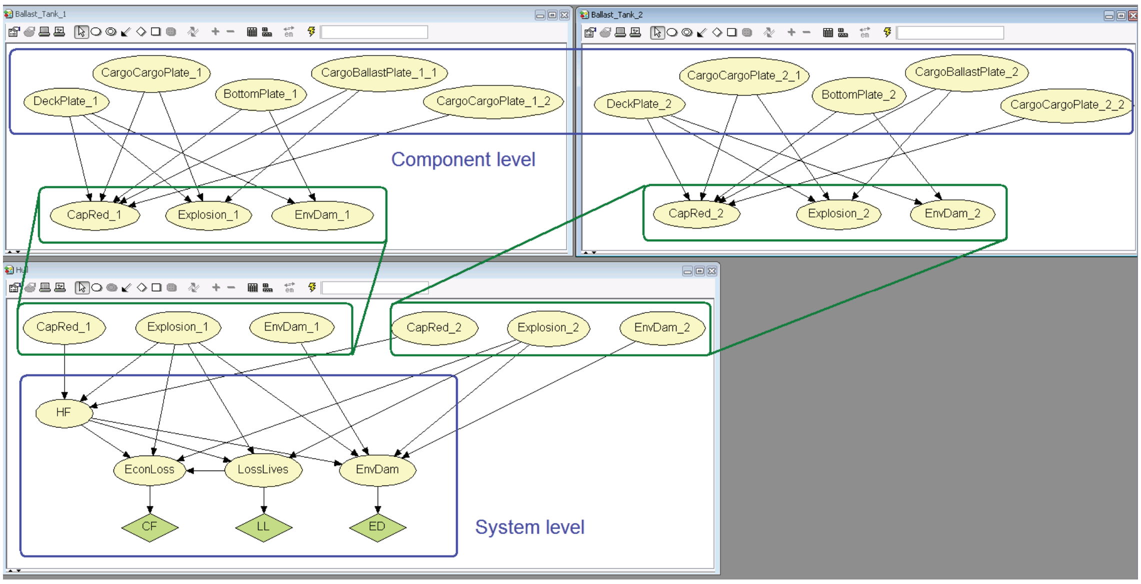


Figure 3.10. ID for the tanks and the hull structure.

Explosion_1

CargoBallastI	Fail				Survive			
CargoCargoF	Fail		Survive		Fail		Survive	
DeckPlate_1	Fail	Survive	Fail	Survive	Fail	Survive	Fail	Survive
No	0.4	0.5	0.8	0.8	0.4	0.5	0.6	1
Minor	0.5	0.45	0.1	0.19	0.55	0.49	0.35	0
Major	0.1	0.05	0.1	0.01	0.05	0.01	0.05	0

Figure 3.11. Illustration of conditional probability table.

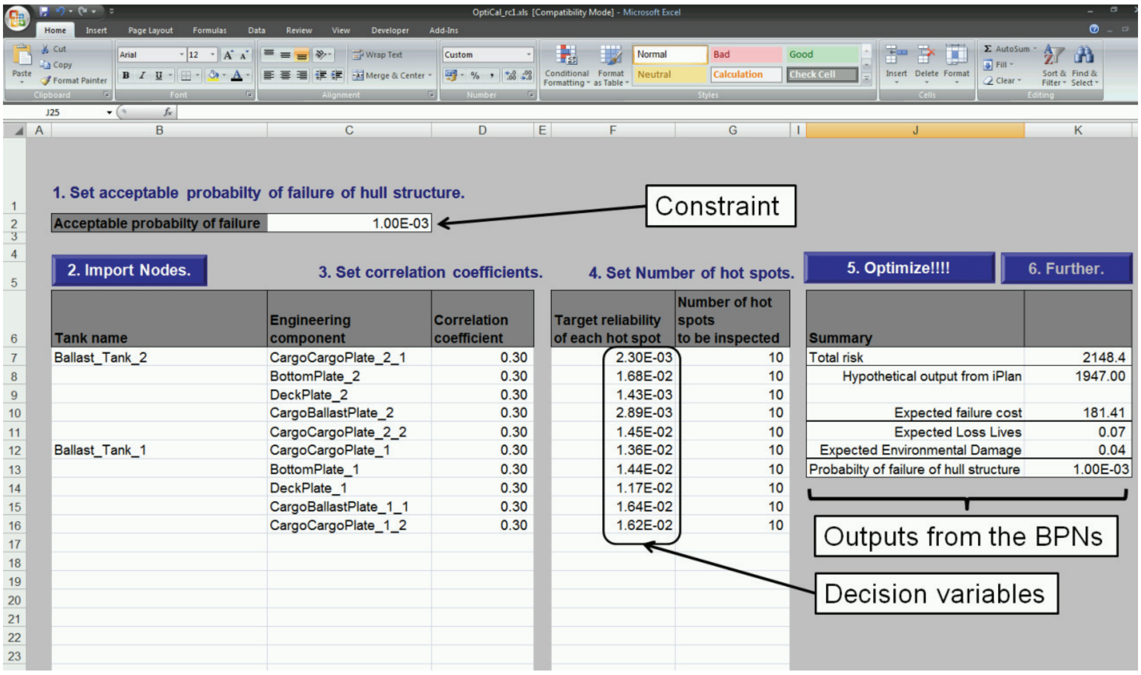


Figure 3.12. User interface of developed software tool.

3.5.2. Results and discussion

In this illustrative example, the acceptable probability of system failure is set as 10^{-3} per annum which constitutes the boundary condition in the optimization problem. The objective function is the expected total cost including the inspection cost, the repair cost and the failure cost due to ship hull failure. As is shown in Figure 3.12, different optimal target reliabilities are obtained for the components in different tanks, reflecting the different contribution to the system failure. The set of these optimal target reliabilities correspond to the set of the target reliabilities that satisfy the acceptance for the probability of system failure and that minimizes the expected total cost. Although in this example, the exposure to the ship hull structure, e.g. wave load, is not directly considered and thus the failures of the individual tanks are assumed to be independent, it is possible to take into account the exposures which may introduce the correlation between the failures of the components and/or subsystems by adding the node for the exposure scenario in the ID as is found in the previous example.

3.6. Conclusions

The present paper proposes a framework for the modeling and the optimization of reliabilities for components in complex engineered systems subject to requirements specified in terms of system performance. It is shown how the identification of the target component reliabilities that are optimal and consistent with given acceptance criteria for system performance can be treated as an optimization problem with constraints. Appreciating the perspective that engineered systems are built up by standardized components which through their connections with other components provide the desired functionality and that the system performance will depend on the way the components are interconnected, the proposed framework takes basis in a hierarchical system modelling facilitated by use of (object-oriented) BPNs and IDs. Using the established BPNs and IDs it is possible to calculate the objective function such as service life utility, and the quantities for which the acceptance criteria are given, both of which are required for solving the optimization problems with constraints. Two examples are shown: (1) optimization of the design of bridges in a transportation network subject to earthquake hazards, and (2) optimization of target reliabilities of welded joints in a ship hull structure subject to fatigue deterioration in the context of maintenance planning. The first example serves as the introduction how the proposed approach is implemented step by step. The second example illustrates how complex engineered system may be modelled and how the target component reliabilities may be optimized using commonly available software tools.

4. Inter-generational distribution of the life-cycle cost of an engineering facility (Paper III)

Kazuyoshi Nishijima

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 22.3, Zurich 8093, Switzerland.

Daniel Straub

Institute of Structural Engineering, ETH Zurich, Switzerland[8].

Michael Havbro Faber

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 23.3, Zurich 8093, Switzerland.

Journal of Reliability of Structures and Materials, Vol. 3, Issue 1, pp. 33-46, 2007.

8 As of the date of the paper acceptance.

Abstract

In decision making for civil engineering facilities, as well as other societal activities, the criteria for sustainability are inter-generational equity and optimality. Two challenging questions must be addressed in this context: How to compare the benefits and costs among different generations and how to compensate and adjust for the in-homogeneously distributed benefits and costs between the generations. To address and answer these questions for engineering facilities, first of all the temporal distribution of the life-cycle benefits must be assessed. To ensure optimality, the total life-cycle benefits for the facility must be maximized. In the present paper initially the normative criteria for sustainability are presented. Thereafter it is demonstrated how the criteria may be implemented for the purpose of optimization of structural design. The inter-generational distribution of benefits and the implications for sustainable decision-making are then illustrated by an example considering the optimal design of the concrete cover thickness of a RC structure subject to chloride-induced corrosion of the reinforcement.

Keywords

Sustainability, discounting, life-cycle cost, chloride-induced corrosion,
cover thickness.

4.1. Introduction

A significant amount of research has been devoted to life-cycle analysis for civil engineering facilities. In recognition of the significant uncertainties associated with the performance of structures over their service life, decision-theoretical approaches have been applied for the optimization of structural design, e.g., Rosenblueth and Mendoza (1971) and Rackwitz (2000). The developed methodological framework facilitates the optimization of the design of structures such that a balance is achieved between the benefits achieved through the facility and the costs associated with design and construction, future costs of inspection and maintenance as well as costs associated with possible repairs, replacements and failures. Recently, life-cycle analysis has been utilized to enhance a sustainable development of the built environment, e.g. Rackwitz et al. (2005), Faber and Rackwitz (2004), Nishijima et al. (2004) and Nishijima et al. (2005). In this context, focus is shifted from the facilities to a sequence of decision makers and stake holders, each of which represents a subsequent generation that benefits from the facility while paying the costs of maintenance, repair, replacement and other adverse consequences. Although life-cycle analysis is well advanced in the civil engineering field and has been applied within the context of sustainability, less attention has been paid to the distribution of costs over time. This distribution is essential, since it allows for assessing the burden of each generation, and thus indicates the necessity for an inter-generational compensation when the aggregation of benefits and costs is not uniformly distributed over time.

The present paper initially formulates the criteria for sustainability and thereafter sets up a multi-decision-maker framework for inter-generational sustainable decision making. As it will be discussed this framework may also provide a useful basis in any intra-generational context for organizations involved in decision making concerning activities with life times significantly exceeding the budgeting periods or the life time of the individuals responsible for the decision making within the organization. The optimization of structural design using the suggested framework is illustrated by an example considering the optimal design of the cover thickness for a RC structure subject to chloride-induced corrosion. Finally, the temporal distribution of the life-cycle costs is explicitly assessed, clearly illustrating how the benefits and costs are unevenly distributed over the generations.

4.2. Multi-decision-makers and criteria for sustainability

Sustainability is interpreted in accordance with the Rio convention in 1992, following the report by Brundtland (1987). To facilitate sustainable decision making, two criteria are provided: 1) inter-generational equity and 2) optimality. Inter-generational equity dictates equal treatment of the present and all future generations. Optimality can be interpreted as the maximization of an idealized utility function, considering all generations and their preferences. These two criteria are strongly interrelated and this must be taken into account in the decision making. In order to set up the utility function aggregating the benefits and costs for all generations, the equal treatment of the individual generations in accordance with the inter-generational equity criterion is required. Once optimality is obtained by maximizing the idealized utility function, the temporal in-homogeneity of the utilities among the different generations must be reconsidered to ensure inter-generational equity.

Basically any kind of activity at present has consequences for the future in terms of benefits and costs. The benefits and costs may not necessarily be expressed in monetary terms and there are controversial discussions on whether all societal and environmental consequences can be measured comprehensively in monetary terms, as discussed in Turner (1992) and Ayres et al. (1998). However, in the present paper, benefits and costs are assumed to be represented by monetary values for the convenience of discussion. The temporal distribution of consequences associated with different activities differs significantly; however, it is difficult to identify activities which do not have some effect for the future generations. In case of exploitation of natural resources the benefit is more or less immediate – but the resources exploited are no longer available for future generations. In case of disposal of toxic waste the situation is much the same – the benefit is achieved by the present generation but the potential adverse consequences are likely to be transferred to future generations. Sustainability is an issue which always has to be kept in mind.

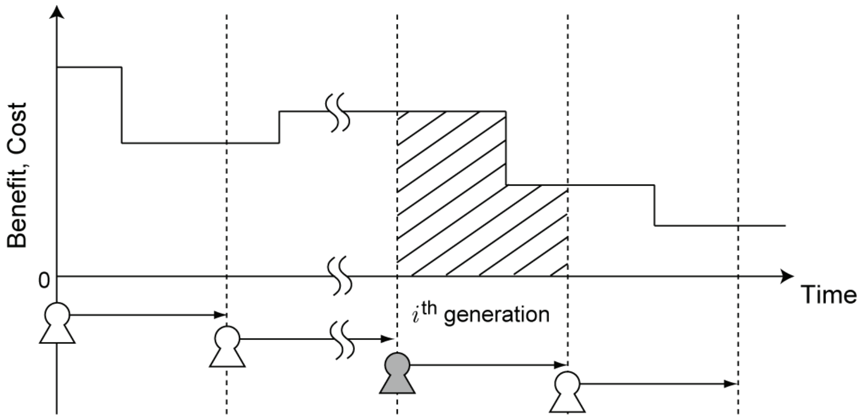


Figure 4.1. Schematic distribution of benefit or cost over time.

The schematic benefit or cost path is illustrated in Figure 4.1. A sequence of decision makers is assumed along with the time, each representing one generation. Since each generation considers the benefits and costs and makes decisions from its point of view, an explicit modeling of the different subsequent decision makers is indispensable, especially when the pure time preference or loss of life are considered in the utility function.

The benefits and the costs illustrated in Figure 4.1 correspond to the gross values at each point in time, i.e., they are not discounted. The i^{th} generation enjoys the benefit or carries the cost of the hatched area. Since this is the gross value, the same values at different points in time do not necessarily have the same perceived influence to different generations, mainly because of the economic growth. Therefore, benefits and costs should be discounted by the economic growth to ensure the equal treatment between generations in accordance with the inter-generational equity. Taking into account the economic growth and disregarding the effect of overlapping generations, the total utility aggregating benefits and costs can be expressed as:

$$U = \sum_{i=1}^{\infty} \delta(t_i) U_i \tag{4.1}$$

where U is the total utility for all generations, $\delta(\cdot)$ is the discounting factor representing economic growth and U_i is the utility for i^{th} generation which begins at $t = t_i$. Extension of Equation (4.1) to cover also the case of overlapping generations may be performed as shown in Bayer and Cansier (1999), Bayer (2003) and Rackwitz et al. (2005), however, the effect of this is of minor importance for the overall life-cycle benefit assessment. When decision making is subject to uncertainty, the utilities in Equation (4.1) should be interpreted as the expected utilities. The utility for the i^{th} generation may be written as:

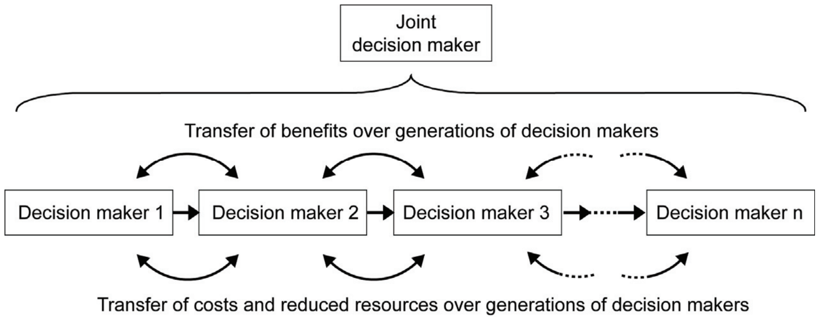


Figure 4.2. Transfer of benefits, costs and resources.

$$U_i = \int_{t_i}^{t_{i+1}} u(t)\gamma(t-t_i)dt \tag{4.2}$$

where $u(\cdot)$ is the utility per unit time and $\gamma(\cdot)$ is the discounting factor within one generation. The utility within one generation may be discounted by pure time preference as well as by economic growth, thus

$$\gamma(t) = \delta(t)\rho(t) \tag{4.3}$$

where $\rho(\cdot)$ is the discounting factor representing pure time preference. Note that the discounting factor is related to the discount rate, e.g., for $\delta(\cdot)$ as:

$$\delta(t) = \exp(-\delta t) \tag{4.4}$$

where δ is the discount rate per unit time.

Each decision in regard to a civil engineering facility results in one specific temporal distribution of expected utility and thus enables the calculation of the total utility according to Equation (4.1). To comply with the second criterion for sustainability, i.e., optimality, the total utility must be maximized, which in the case where the benefit function does not depend directly on the decision corresponds to a minimization of the total cost. However, even if the maximization is performed under consideration of inter-generational equity in terms of proper discounting as applied in Equations (4.1) - (4.3), it does not necessarily imply that each generation obtains the same utility from the facility, as illustrated in Figure 4.1. It is unlikely that each single activity optimized in the above sense results in a uniform distribution of the utility among the current and all future generations. Therefore, the transfer of the benefits in terms of, for instance, man-made capital or natural resources is essential to achieve inter-generational equity, see Figure 4.2. The distribution of costs over time provides the basic information required to achieve inter-generational equity, enabling a comparison and a compensation between the generations through societal activities which are not necessarily within the civil engineering field.

4.3. Equivalent sustainable discount rate

Classical life-cycle cost analysis approaches the discounting problem from the perspective of the anticipated duration of the considered activity, e.g. the anticipated service life when a given structure is considered. Furthermore, decision making in classical life-cycle analysis takes basis in a utility modeling where only the preferences of the present generation are directly accounted for. This includes also the aspects of valuation of future benefits and costs through discounting. For a given activity it is possible to assess a discount rate which if applied in a classical life-cycle analysis yields the same total expected utility as resulting from the proposed multi-decision-maker framework (Equations (4.1) - (4.3)). This discount rate is denoted the equivalent sustainable discount rate γ^* by:

$$\int_0^\infty e^{-\gamma^* t} u(t)dt = \sum_{n=1}^{\infty} e^{-\delta t_n} \int_{t_n}^{t_{n+1}} e^{-\gamma(t-t_n)} u(t)dt \tag{4.5}$$

where $u(t)$ is the (expected) utility per unit time at time t. The equivalent sustainable discount rate may be interpreted as the one which, if applied to a decision problem with the classical one-decision-maker perspective, yields the same total expected utility as when the decision problem is analyzed from the multi-decision-maker perspective. In general, it is not possible to obtain an analytical expression for γ^*. However, in the case where consequences are invariant at any time, the durations of generations $\tau = t_{n+1} - t_n$ ($n = 1,2,3,...$) are constant and the occurrences of events associated with consequences follow a stationary Poisson process, the equivalent sustainable discount rate is given as follows:

$$\gamma^* = \frac{1-e^{-\delta\tau}}{1-e^{-\gamma\tau}} \gamma \tag{4.6}$$

where δ is the discount rate per unit time by economic growth , ρ is the discount rate per unit time by pure time preference and $\gamma = \delta + \rho$, see Faber and Nishijima (2004). The equivalent sustainable discount rates for several cases are illustrated in Figure 4.3, where for ρ kept constant at 3% per year or 0% per year for comparison, the equivalent sustainable discount rates are given as functions of the duration of the generation τ for several values of δ. The equivalent sustainable discount rate γ^* is smaller than the total discount rate γ, except for the case where $\rho = 0$. If the discount rate consists only of pure time preference ($\delta = 0$), the equivalent sustainable discount rate is zero, i.e., within the classical framework, the benefits and costs should not be discounted at all to obtain the same utility function as with the multi-decision-maker framework. If the discount rate by pure time preference is set equal to zero and the discount rate by economic growth is set to equal to 5%, the equivalent sustainable discount rate is equal to 5%, regardless of the duration of the generation. This means that if the discount rate is only due to economic growth, the multi-decision-maker framework is identical to the classical framework. In general, the

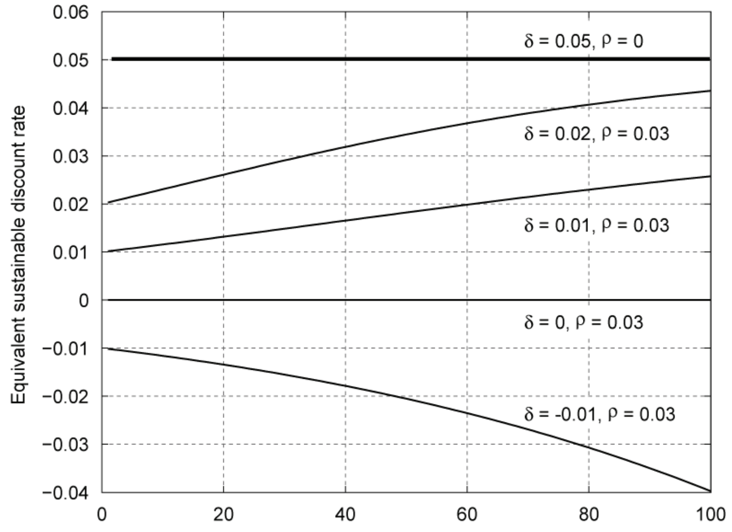


Figure 4.3. Equivalent sustainable discount rate γ^* for the case of constant utility per unit time.

discount rates which have been applied so far in the classical framework are too large, i.e., are leading to non-optimal solutions from the view-point of sustainability.

4.4. Example

Optimal life-cycle cost based design of the concrete cover thickness of a RC structure subject to chloride-induced corrosion of the reinforcement is considered. The intended service life time is assumed to be infinite, meaning that the desired function of the structure is unlimited in time. The applied probabilistic modeling of the degradation over time is included in Annex A for simple reference and more details are provided in Faber et al. (2005). The expected life-cycle costs are assumed to consist of the initial costs C_I, the expected repair costs $E[C_R]$ and the expected failure costs $E[C_F]$, which all depend on the optimization variable d_{nom}, i.e. the concrete cover thickness. It is assumed that visual inspections are made every $\Delta t_I = 5\text{yr}$ and that an indication of visible corrosion automatically triggers a repair. In accordance with the renewal-theoretical approach outlined in Faber and Rackwitz (2004), it is assumed that in case the structure fails, it is reconstructed. Following a repair or a reconstruction, the structure is assumed brought back to its original state, i.e., described using the same probabilistic model as a new structure. The realization of the structure after repair or reconstruction is assumed to be independent from previous structures. Furthermore, inspections are modeled as being perfect, i.e., visible corrosion is detected with probability 1 at an inspection. The costs of initial design, repairs and failures are modeled as:

$$C_I = (1 + a_I d_{nom})C_0 \tag{4.7}$$

$$C_R = a_R C_I \tag{4.8}$$

$$C_F = a_F C_I \tag{4.9}$$

with parameter values in Table 4.1, where also the assumed discount rates are summarized. The initial cost C_I is assumed to consist of a fixed cost and the cost depending on the cover thickness, and the repair cost C_R and the failure cost C_F are assumed to be proportional to the initial cost.

Table 4.1. Cost and discount model.

Discount rate for time preference: ρ	3% per year
Discount rate for economic growth: δ	2% per year
Normalizing cost C_0	1
Cost ratio for cover thickness a_I	0.002
Coefficient of repair cost a_R	0.5
Coefficient of failure cost a_F	5

4.4.1. Cost distribution over time

In order to calculate the distribution of life-cycle costs over time, an efficient algorithm is required, since the number of branches in the decision tree develops exponentially with time. In Nishijima et al. (2004), these costs are calculated by using a recursive formulation; in the following, a different recursive formulation is provided which facilitates the explicit calculation of the expected cost of repair and failure at each point in time. After specifying the decision rule, which defines in which situations a repair is made, the probability of repair $q_R(t)$ and the probability of failure $q_F(t)$ at time t ($t = 1\text{yr}, 2\text{yr}, 3\text{yr}, \ldots$) for a given realization of the structure are readily available, see e.g. Faber et al. (2005), Nishijima et al. (2004) and Nishijima et al. (2005). In accordance with the above, the decision rule adopted in this example is that the structure is repaired if and only if corrosion is visibly observed at the inspection. Whether or not this decision rule is optimal is beyond the scope of this paper, which focuses on the design optimization. However, by consideration of the deterioration model and the possible actions it is easily seen that, for the present example, there are only a few reasonable alternative rules. When optimizing the inspection/maintenance strategy, these alternatives can be compared and the one leading to minimal costs can be selected, e.g., Straub (2004). According to the probabilistic model in Annex A, $q_R(t)$ and $q_F(t)$ are estimated by Monte Carlo simulation with 10^{-6} samples for each cover thickness, see Figure 4.4 and Figure 4.5. Note that the probability of repair can be different from zero only at $t = i\Delta t_I$ ($i = 1, 2, 3, \ldots$), since the repairs are associated with inspections which are made at intervals of $\Delta t_I = 5\text{yr}$; failure can occur in any year, but its probability is increasing with time and thus more likely to occur when approaching the inspections. With the probabilities $q_R(t)$ and $q_F(t)$, the probability of repair $P_R(t)$ and the probability of failure $P_F(t)$ at time t are calculated based on the renewal theory (see, e.g., Feller (1966) in general and Rackwitz (2000) considering applications to civil engineering facilities) as:

$$P_R(t) = q_R(t) + \sum_{s=1}^{t-1} \left(P_R(s) + P_F(s) \right) \cdot q_R(t-s) \tag{4.10}$$

$$P_F(t) = q_F(t) + \sum_{s=1}^{t-1} \left(P_R(s) + P_F(s) \right) \cdot q_F(t-s) \tag{4.11}$$

for $t = 2\text{yr}, 3\text{yr}, 4\text{yr}, \ldots$ and

$$P_R(1\text{yr}) = q_R(1\text{yr}) \tag{4.12}$$

$$P_F(1\text{yr}) = q_F(1\text{yr}) \tag{4.13}$$

for $t = 1\text{yr}$.

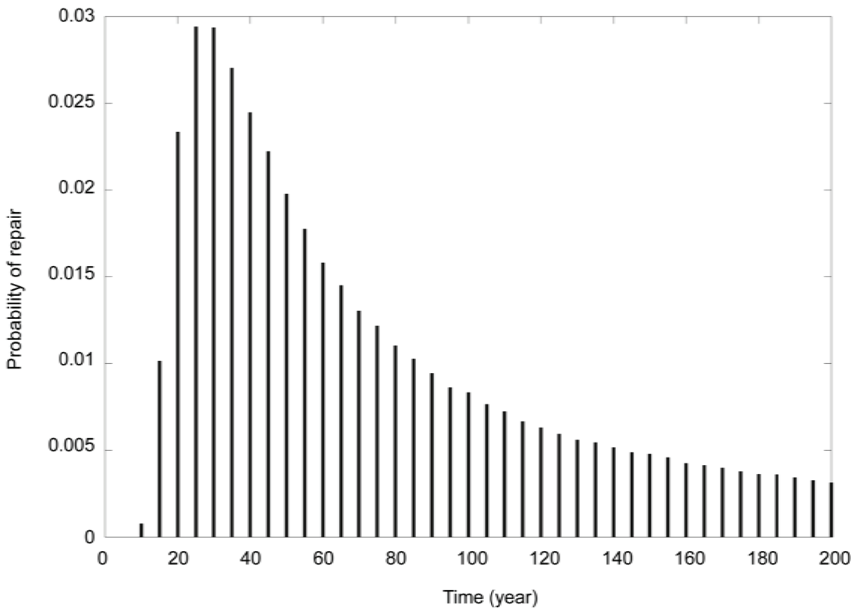


Figure 4.4. Probability of repair $q_R(t)$ at time t for a given realization of the structure (cover thickness = 50 mm).

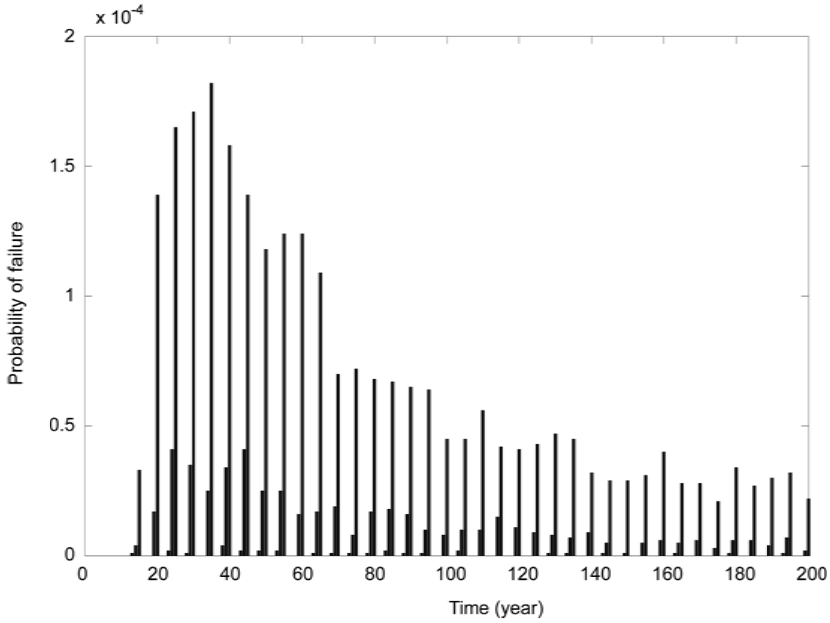


Figure 4.5. Probability of failure $q_F(t)$ at time t for a given realization of the structure (cover thickness = 50 mm).

The recursive formulations Equations (4.10) to (4.13) are obtained as follows. The set of possible different events leading to a repair at time t can be split into subsets: These subsets are differentiated by the time of the last repair or reconstruction, which can occur at times $t-1\text{yr}$, $t-2\text{yr}$, etc. until 0yr; the latter corresponding to the case where no repair or reconstruction has been performed previously. The probability of failure at time t is obtained analogously. As the decision rule just specifies $q_R(t)$ and $q_F(t)$, this recursive formulation can be applied for any kind of decision rule, as long as the structure is repaired at some point in time and reconstructed after failure, resulting in identical but stochastically independent structures.

Once the probability of repair and the probability of failure at each point in time are obtained, the calculation of the expected costs is straightforward. Repair and reconstruction after failure can be carried out at each inspection time, the inspection interval being 5 years. Figure 4.6 shows the distribution of costs over time for several cover thicknesses. These costs are not discounted. The expected cost for each point in time consists of the expected repair costs and the expected failure costs. The expected failure costs are much smaller than the expected repair costs in the present example, that is why the expected total costs are close to the expected repair costs. The (non-discounted) expected costs decrease with time for all cases in Figure 4.6. This tendency is due to the fact that the failure rate, which is the probability of failure per unit time conditional on survival up to time t, is decreasing with time for the considered deterioration mechanism. When the structure performs poorly (i.e., the realizations of the random variables are unfavorable), it will be repaired or reconstructed already after a few years. After each repair or reconstruction, the new structures are identical but stochastically independent of the old ones. A structure with an initially bad performance will thus eventually be replaced by one with a good performance. The expected value of the performance of the structure is therefore increasing with time and the expected costs of failures and repairs are decreasing. It should be realized that this tendency depends strongly on the assumed dependency between subsequent realizations of the structure as well as the characteristics of the failure rate function.

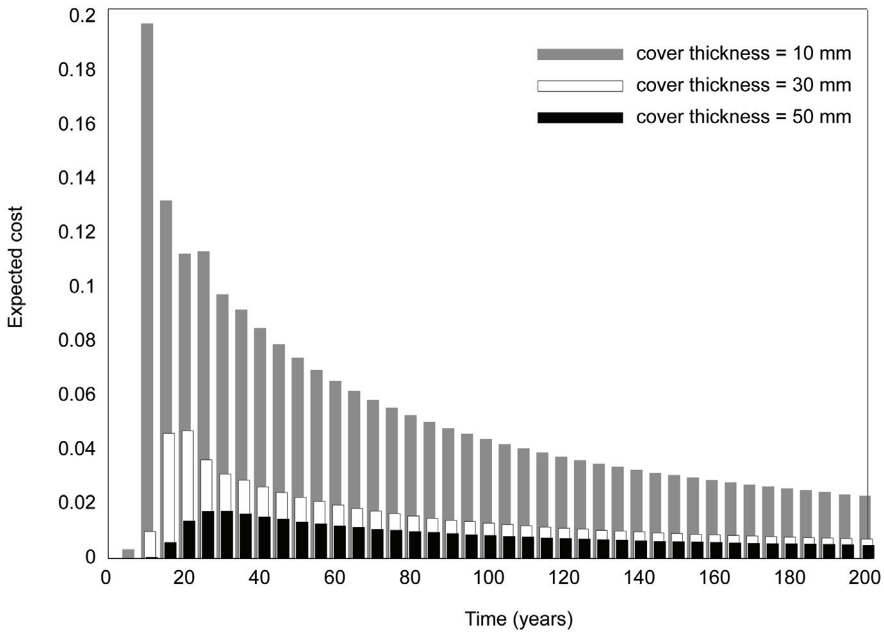


Figure 4.6. Temporal distribution of expected costs (at 5 year intervals), not discounted.

4.4.2. Optimization of the concrete cover thickness

Taking basis in the multi-decision-maker framework presented in the previous section, the total expected costs to be minimized are calculated for each decision alternative (i.e., for different cover thicknesses). For this example the total expected costs reduce to:

$$\begin{aligned}-E[U(d_{nom})] &= E[C(d_{nom})] \\ &= C_I(d_{nom}) + \sum_{i=1}^{\infty}\delta(t_i)\sum_{j=1}^{\tau/\Delta t_I}\left(E[C_{R,t_i+j\Delta t_I}(d_{nom})] + E[C_{F,t_i+j\Delta t_I}(d_{nom})]\right)\gamma(j\Delta t_I)\end{aligned} \quad (4.14)$$

which should be minimized. $C_{R,t}$ and $C_{F,t}$ are the costs of repair and failure at time t respectively. Since different discount rates are applied within the generations and between the generations, the duration of each generation τ must be specified. Figure 7 shows optimal cover thicknesses for different values of the durations of the generations. With increasing duration of generations, the optimal cover thickness becomes smaller. This is because the "equivalent sustainable discount rate" becomes larger as the duration of the generation becomes longer, see Equation (4.6) and Figure 4.3, and consequences in the future are, therefore, valued less. The case where the duration of a generation is infinite corresponds to the classical life-cycle analysis where only one decision maker is assumed. As observed in Figure 4.7, the optimal cover thickness varies significantly with the duration of the generations, pointing to the importance of considering the problem from the viewpoint of the multi-decision-makers.

In the following, the duration of a generation is assumed to be 25 years. When applying the multi-decision-maker framework, the optimal cover thickness is 52mm, in accordance with Figure 4.7. By applying the classical framework (the infinite duration of generations in Figure 4.7), the optimum is at 44mm. Figure 4.8 shows the corresponding expected costs with time. These costs are discounted to time $t = 0$, therefore, it is possible to compare these costs with each other. Within the classical framework the first generation pays less and the future generations pay more than within the multi-decision-maker framework. This is due to the fact that the classical framework weighs values in the future less than the multi-decision-maker framework through the relatively higher discount rate.

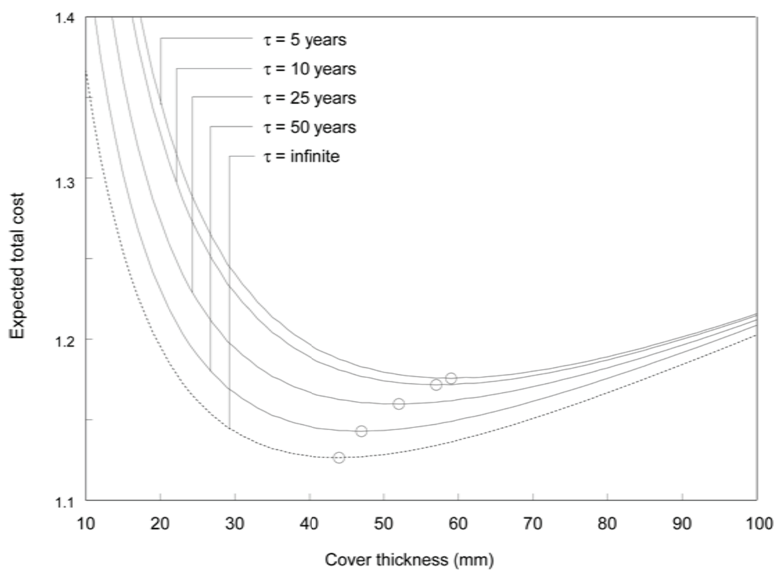


Figure 4.7. Optimal cover thickness for several durations of generation τ.

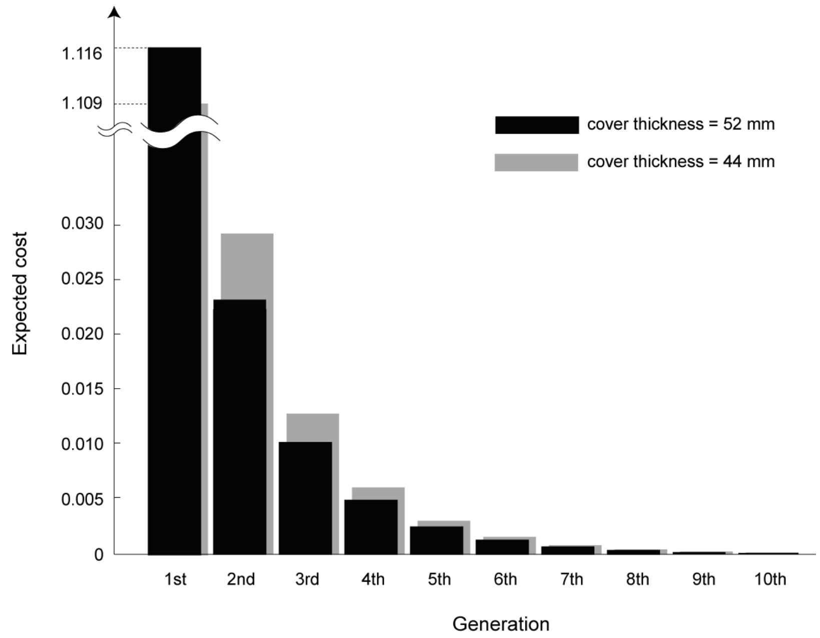


Figure 4.8. Discounted expected costs for each generation (with a duration 25 years).

4.5. Discussion

Figure 4.8 clearly shows the inhomogeneous distribution of costs among the generations. In particular the first generation pays much more than all following generations. In order to comply with the first criterion for sustainability, inter-generational equity, the temporal differences must be compensated by other means (e.g., by transferring the benefits on capital stocks and natural resources). Such compensation is beyond the scope of the analysis as presented in this paper, as it requires that all societal activities must be considered simultaneously within the multi-decision-maker framework. In this context it is reminded that, although in the presented example it is the first generation which pays most, many societal activities have large consequences in the future while only the current generation directly benefits from them.

The presented framework can be extended to portfolios of structures, which are distributed over time and space. The optimization of design and maintenance activities is performed in analogy to the case of the individual structure, but to ensure inter-generational equity through compensation, it is required to consider the cost distribution over time for all structures simultaneously.

The analysis presented here ensures that the second criterion of sustainability, optimality, is fulfilled in such a way that it is consistent with the first criterion. It seems paradoxical at first that by consideration of multi-decision-makers (which is required by the inter-generational equity criterion), the optimal design which fulfills the optimality criterion leads to an even more inhomogeneous distribution of costs among generations. For this reason it is crucial that the issue of compensation between the generations is also addressed.

The presented multi-decision-maker framework provides an analytical approach to the consideration of the preferences of all generations involved in the life-cycle of engineering structures. It allows for the assessment of the effect of postponing costs to the future through the use of large interest rates, which is a common tendency in societal decision making. In order to be sustainable, the equivalent sustainable discount rate presented in this paper must be applied.

Finally, it is important to note that whereas the present paper specifically addresses the problem of sustainable decision making in an inter-generational context the developed framework also may be valuable for the decision making in intra-generational contexts involving several decision makers and stakeholders as well as budgets over time. This is the situation when decision making is considered in organizations which are responsible for the design, construction and operation of engineering facilities such as high-way agencies. In such organizations both budgets as well as the persons involved in the decision making have a substantially shorter life time than the facilities they are responsible for. The multi-decision-maker framework may serve to set guidelines or rules for the decision making in such contexts, to help avoid decisions which for the fulfillment of preferences of individuals may yield a short term benefit but from an overall life-cycle perspective induce economical losses for the organization. Furthermore, the framework can be utilized as a rational basis for long term budgeting.

4.6. Conclusions

It is demonstrated how the inter-generational distribution of the life-cycle cost of an engineering facility can be assessed. This is of importance for ensuring sustainability of the facility, whereby the considered criteria for sustainability are inter-generational equity and optimality. It is shown how decisions regarding an engineering facility must be optimized in order to comply with these criteria and it is outlined that the results of the optimization may be used as a basis for a broader discussion regarding inter-generational equity taking into account all kinds of societal activities. Finally, it is highlighted that the developed framework also may provide a useful basis in any intra-generational context for organizations involved in decision making concerning activities with life times significantly exceeding the budgeting periods or the life time of the individuals responsible for the decision making within the organization.

The developed decision framework is illustrated by the optimization of the design of a RC structure subject to chloride-induced corrosion and is found to have a significant effect on the optimal design.

4.7. Annex A

For easy reference, the applied probabilistic model for deterioration of concrete structures subject to chloride-induced corrosion is presented in the following. The modeling corresponds to DuraCrete (2000) and here follows Faber et al. (2005), where additional details of the models are described.

Corrosion initiates at the reinforcement, when the chloride concentration has reached the critical chloride concentration C_{Cr}. The ingress of chlorides in the concrete is described by Fick's second law of diffusion. Based on this model, the random variable T_I representing the time until corrosion initiation is calculated as:

$$T_I = \left(\frac{d^2}{4k_e k_t k_c D_0 \left(t_0\right)^n} \left[erf^{-1}\left(1 - \frac{C_{cr}}{A_{C_S} \cdot (w/c) + \varepsilon_{C_S}} \right) \right]^{-2} \right)^{\frac{1}{1-n}} \tag{4.15}$$

The parameters of the model are given in Table 4.2.

The time until visible corrosion, corresponding to minor cracking and coloring of the concrete surface, can be determined based on experience. By adding the propagation time T_P to the initiation time T_I, the limit state function for visible corrosion is written as:

$$g_{VC}\left(t\right) = X_I T_I + T_P - t \tag{4.16}$$

The time between visible corrosion and failure is, for illustrative purposes, represented by the time T_P. The limit state function for failure is thus:

$$g_F\left(t\right) = X_I T_I + T_P + T_{P2} - t \tag{4.17}$$

Note that the model does not account for the dependency between the propagation time T_{P2} and the environmental parameters or the cover thickness.

The values of the distribution parameters for the random variables in Equations (4.15) to (4.17) can be obtained as functions of indicators, see Faber et al. (2005). For the considered example, they are stated in Table 4.2. These values are representative for a concrete with ordinary Portland cement in a splash environment.

The probabilities of the events visible corrosion and failure can be obtained by e.g. Structural Reliability Analysis (SRA) or simulation techniques.

Table 4.2. Example parameters for the deterioration model.

Parameter	Description	Distribution	Dimension			A	B
d	Cover thickness	Lognormal	mm				
k_e	Environmental factor	Gamma	-	0.924	0.155		
k_c	Curing factor	Beta	-	0.8	0.1	0.4	1.0
k_t	Test factor	Deterministic	-	1.0	-		
D_0	Diffusion coef.	Normal	mm2/yr	220.9	25.4		
t_0	Reference period	Deterministic	yr	0.077	-		
n	Age factor	Beta	-	0.362	0.245	0	0.98
C_{cr}	Critical chloride concentration	Normal	*	0.8	0.1		
w/c	Water/cement ratio	Deterministic	-	0.40	-		
A_{C_S}	Chloride surface concentration factor	Normal	*	7.758	1.36		
e_{C_S}	Chloride surface concentration factor	Normal	*	0	1.105		
X_I	Model uncertainty	Lognormal	-	1.0	0.05		
T_P	Propagation time	Lognormal	yr	7.5	1.9		
T_{P2}	Propagation time	Lognormal	yr	10.0	4.0		

* Mass-% of binder

5. A budget management approach for societal infrastructure projects (Paper IV)

Kazuyoshi Nishijima

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 22.3, Zurich 8093, Switzerland.

Michael Havbro Faber

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 23.3, Zurich 8093, Switzerland.

Structure and Infrastructure Engineering, Vol. 5, Issue 1, pp. 41-47, 2009.

Abstract

Life cycle costing analysis is broadly applied as a tool for decision support for civil engineering structures, whereby the expected total cost over the life cycle of the structure is advocated as the objective function to be minimized. The present paper takes the new perspective of considering the problem from a budgeting allocation problem where the aim is to optimize the allocation of budgets for the purpose of maintaining the operation of the portfolio of structures. Whereas all the consequences associated to the project must be taken into account in the life cycle costing analysis, it is important to distinguish the financial costs which must be paid from the user costs which represent the follow-up consequences, i.e., opportunity losses. This is because only the costs to be paid are related to the budget. The present paper proposes an approach to determine the optimal amount of budget and the optimal maintenance decisions, considering these two types of cost.

Keywords

Objective function, resource allocation, life cycle optimization.

5.1. Introduction

Over the last decade life cycle costing analysis has gained a widespread interest as a tool for decision support in civil engineering, e.g., Rosenblueth and Mendoza (1971), as well as in many other engineering fields. It has been appreciated in research and practice that the efficiency of engineering projects must be assessed with due consideration of all benefits and costs induced by the projects on time scales representative for the actual duration of the projects; only when the life-cycle benefits are larger than the corresponding costs can an engineering project be considered feasible, e.g. Rackwitz (2000). The feasibility of engineering projects such as societal infrastructure must thus be assessed considering all phases throughout their life-cycle – from the concept phase until the decommission.

As opposed to most private business initiatives, infrastructures built for the purpose of facilitating the development of society serve functions or are in other ways associated with benefits and/or costs which on the time scale reach well beyond the duration of the generations who decide to build them. To ensure a sustainable societal development, i.e., a development which aims to optimize the objectives of not only our own generation but also those of the future generations, the assessment of life-cycle costs must take into account the costs implied for future generations. To this end, life cycle costing analysis together with an appropriately chosen discounting function, see e.g., Rackwitz et al. (2005), provides a consistent rationale. The objective function to be minimized is, in many cases, the expected total cost under the assumption that the benefits from structures are independent of the decision variables, taking the follow-up consequences, e.g. reduced benefits due to unavailability, into account as user costs.

Turning our focus to practical situations, however, the decision makers which are responsible for the maintenance of portfolios of structures request budgets which are in excess of the expected total costs. The reason for this is obviously in part that their success as decision makers is measured in terms of whether they are able to meet their requested budgets and at the same time are able to keep their portfolio of structures in operation. It may well be that they request more if the lack of budget leads to serious consequences such as user costs associated with reduced functionality of roadway systems. Given this practical constraint, an optimal decision which minimizes the expected total cost does not necessarily lead to an optimal budgeting from a societal point of view, corresponding to a resource allocation of the society maximizing the societal net benefit. Thus the optimization of the decision and the total budget by maximizing the societal benefit becomes an issue in the context of optimal societal resource allocation.

The present paper proposes an approach to identify optimal decisions related to maintenance of structures and budget allocation by maximizing the expected net benefit, where the net benefit is composed of the benefit achieved through the operation of the considered portfolio, the allocated budget, the financial cost to be paid, the user cost and additional user cost which arises from the delay of maintenance activities due to the possible lack of budget. An example of the maintenance of a portfolio of RC structures subject to chloride-induced corrosion is given to illustrate how the proposed approach works in practical applications.

5.2. Budget management approach

5.2.1. Resource allocation

Optimal societal resource allocation has gained increased attention since the so-called Brundtland report (Brundtland (1987)) set focus on sustainability. In principle, resource allocation in a society is realized through allocation of the total available budget to the various sectors in the society. Maintenance and operation of civil engineering infrastructure represents one of the societal activities or sectors which must be allocated a budget to ensure the continued societal benefit from the structures. Within this sector, the budget is subdivided into smaller parts for sub-projects, typically to different groups of structures or e.g. segments of the roadway system. Despite the fact that a sector-wise or project-wise budgeting system can cause inefficiency of societal resource allocation, it is assumed in this paper to be a given constraint. The normative discussion whether or not the sector-wise or project-wise budget allocation is preferable is beyond the scope of the present paper.

Subject to significant uncertainties related to civil engineering projects, decision makers must decide on the amount of budget necessary and sufficient for successfully managing the projects. It has been widely accepted in life cycle costing analysis that the objective function to be minimized is the (discounted) expected total cost including follow-up consequences such as user costs. In this regard, it can be said that the optimization by the minimization of expected total cost implicitly assumes a "perfectly flexible budgeting", namely, a situation where the budget is always available when needed. For the purpose to assess the optimal amount of budget required for a project or projects, however, this may not be appropriate. A structure which has reduced availability due to failure or the need of repair works may not be rehabilitated due to insufficient budgets, which in turn may lead to additional user costs. The optimal budget allocation may not correspond to the expected total cost. In order to maximize the net benefit, the budget allocation, the financial costs and the user cost must be considered simultaneously.

In a broader sense, the objective function should be an aggregated utility, in which all the preferences of the decision maker are included, see e.g., Faber and Maes (2003). In practical situations, a decision maker may be precautious in a sense that he/she requests more budget than the expected cost in order to ensure a successful management of projects. In the following section, the net benefit is proposed as a utility function to represent the preferences of the decision maker.

5.2.2. Net benefit maximization

Life cycle costing analysis for civil engineering structures usually considers only the cost side, assuming that the benefit B is indifferent to the choice of the decision variables. The reduced benefits due to the loss of functions caused by adverse events are included in the cost term as user costs. However, the budget must also be taken into account in the analysis, since a failed structure or a structure with reduced availability cannot provide the desired benefits until it has been rehabilitated. If the budget is insufficient, the operation of such structures cannot be recovered until the budget is available. This can lead to additional user costs. Thus the evaluation of user costs is dependent on whether the budget for the recovery of the operation is available or not.

The net benefit NB induced by a structure may be written as:

$$NB = \begin{cases} B - K - \Delta B(e) & (C(a,e) \le K) \\ B - C(e,a) - \Delta B(e) - \Delta B(e,'C > K') & (C(a,e) > K) \end{cases} \qquad (5.1)$$

where K is the allocated budget, $\Delta B(e)$ is a user cost corresponding to the event e, $C(a,e)$ is the financial cost corresponding to (a,e), a is the decision variable and $\Delta B(e,'C > K')$ is a user cost induced by the possibly insufficient budget following the event e, see Figure 5.1. If the financial cost $C(a,e)$ does not exceed the budget K, the net benefit is the difference between the benefit and the sum of the budget and the user cost associated with the event e. Here it is assumed that the unused part of the budget within a budgeting period is not transferred to the next budgeting period which is a commonly known difficulty in the public sector. If the financial cost $C(a,e)$

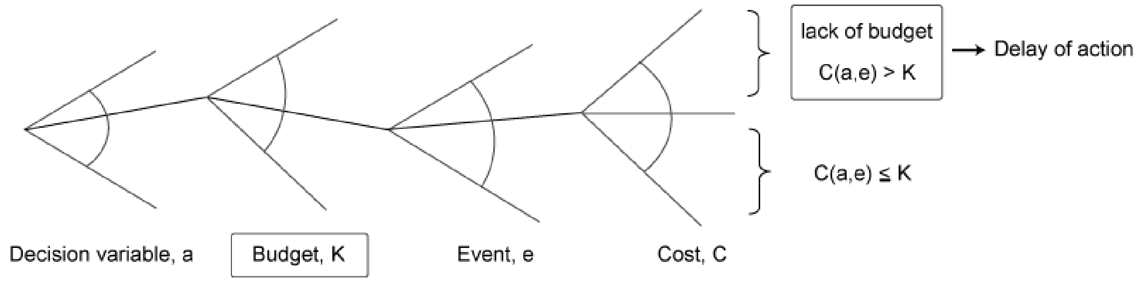


Figure 5.1. Decision event tree including budget.

exceeds the budget K (Budgeting failure), an extra budget must be asked for in order to reinstate the reduced availability, which will be provided at some later point in time, e.g. the subsequent budgeting period. Until the extra budget is obtained, the availability remains reduced, causing the additional user cost $\Delta B(e,'C > K')$.

As the amount of budget increases, the probability of budgeting failure $P(C > K)$ and the net benefit decreases, and vice versa. The optimal budget K^* and the optimal decision variable a^*, e.g. concerning inspection and maintenance activities are obtained by maximizing the expected net benefit $E[NB]$:

$$E[NB] = \int_E NB(a,e,K)dP(e;a) \tag{5.2}$$

where E is the set of possible events e and $P(e;a)$ is the probability of the occurrence of the event e given the decision variable a.

5.3. Example

5.3.1. Maintenance planning for a portfolio of RC structures

The maintenance planning for a portfolio of RC structures subject to deterioration due to corrosion is considered. The portfolio consists of 50 structures, each of which is composed of 100 elements. For illustrational purposes, they are all assumed to be 10 years old, operable but subject to deterioration. The objective of the maintenance planning is to find the optimal inspection interval and the optimal budget for each year so that the net benefit is maximized. The benefits induced by the portfolio are assumed to be independent of the inspection. Therefore, by letting $CT = B - NB$, the objective function to be minimized is written as:

$$E[CT] = \int_E CT(\Delta t_I, e, K)dP(e;\Delta t_I) \tag{5.3}$$

where Δt_I is the inspection interval (corresponding to the decision variable a in Section 5.2.2) and:

$$CT = \begin{cases} K + \Delta B(e) & (C(\Delta t_I, e) \le K) \\ C(e, \Delta t_I) + \Delta B(e) + \Delta B(e, 'C > K') & (C(\Delta t_I, e) > K) \end{cases} \tag{5.4}$$

CT can be considered the total cost including all consequences and is thus referred to as the "total cost" in the subsequent. It should, however, be noted that the total cost represented by Equation (5.4) differs from the typical definition in commonly applied life cycle analysis in the sense that it includes the budget and possible additional user cost due to insufficient budget. The term $\Delta B(e, 'C > K')$ accounts for the effect of insufficient budget to the reduction of the net benefit. Still, the net benefit defined by Equation (5.2) is maximized by minimizing the 'total cost' defined by Equation (5.3).

5.3.2. Inspection, repair and failure

Two states of visually observable corrosion for an element of a structure are considered, i.e., the state which will induce a repair e_R, and the state which corresponds to failure e_F. The former state requires the need of repair, e.g., replacement of concrete cover, while the latter state needs more serious action, e.g., replacement of reinforcement. Thus, the set of events E in this example is expressed as:

$$E = \{(e_0, e_R, e_F);\ \text{for all years and all elements in all structures}\} \tag{5.5}$$

where e_0 is the state when no action is required.

For the purpose of the illustration but with no effect on generality, it is assumed that the inspections are made visually and the probability of detection of corrosion is assumed to be equal to one, i.e., perfect inspections. As long as the budget is sufficient for performing the necessary repairs of the corroded elements, those are assumed performed in connection with the performed inspections. It is further assumed that the repaired elements are brought back to their original states, i.e., described using the same probabilistic models as those for new elements and the realization of the new elements are independent of the previous elements. The basic characteristics of the probabilistic modelling of deterioration are provided in the next section.

The number of structures to be inspected is uniformly distributed over time in accordance with the inspection planning – for instance, when the inspection period is 4 years, the number of structures to be inspected during 4 years is 13, 13, 12 and 12, respectively.

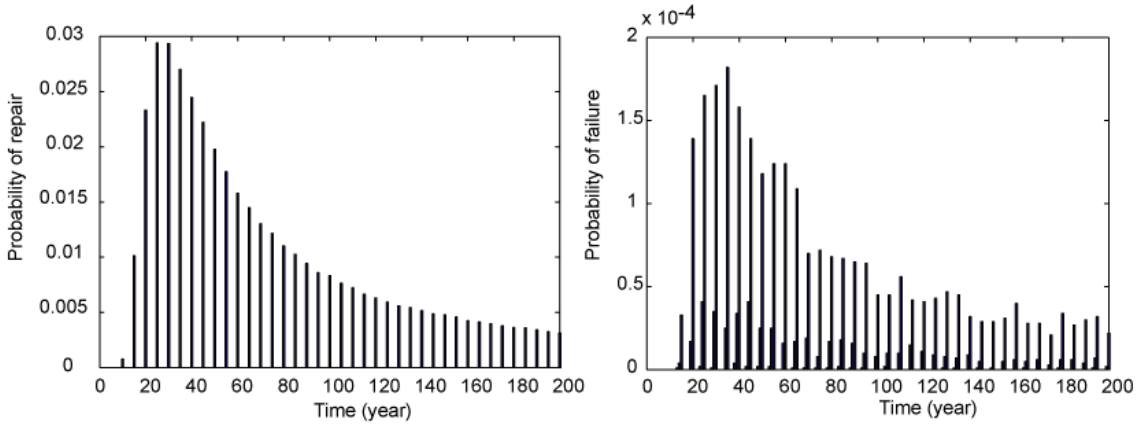


Figure 5.2. Probability of repair (left) and probability of failure (right) for a given realization of element.

5.3.3. Probabilistic corrosion model

The probabilistic model adopted here corresponds to DuraCrete (2000) and follows Faber et al. (2005). All the uncertain parameters are assumed to be independent between different elements. Two limit state functions are explicitly considered: One is related to the time until the realization of visual corrosion, which corresponds to the event e_R and the other is related to the time until the element fails, which corresponds to the event e_F .

An element is repaired if visual corrosion is observed at the time of inspection as long as the budget is sufficient. An element fails if and only if the degradation reaches the failure limit state between two subsequent inspections. Thus the probability of repair, $q_R(t)$, and the probability of failure, $q_F(t)$, at time t for a given realization of an element both depend on the inspection interval. In this example, the design or nominal cover thickness is assumed to be equal to 50mm.

Figure 5.2 shows the probability of repair and failure at time t after construction, repair or recovery due to failure for a given realization of an element in the case of $\Delta t_I = 5$. Both the probability of repair and failure are calculated by Monte Carlo simulation. The probability of repair is different from zero only at $i\Delta t_I$, ($i = 1, 2, 3, ...$). This is because the repair is made only if visual corrosion is observed at the inspection. On the other hand, failure can occur at any point in time. The probability of failure over time varies significantly as the inspection interval changes. When the inspection period is small, e.g., $\Delta t_I = 1$, the probability of failure is low, since more elements are already repaired. In contrast, if the inspection period is large, elements may fail more frequently before repair due to the few inspections. Thus, the inspection period affects both the probability of repair and the probability of failure. It should be mentioned that the time axis in the figure does not necessarily represent the structure age after the first installation, since a structure may have been repaired or replaced. When a structure performs poorly (i.e., the realizations of the random variables are unfavorable), it will be repaired or reconstructed relatively early. After each repair or replacement, the new structures are identical but stochastically independent of the old ones. A structure with an initially bad performance will thus eventually be replaced by one with a good performance. This is why the probabilities of failure and repair decrease after their peak.

5.3.4. Cost model

The financial maintenance cost C, consists of inspection cost C_I, repair cost C_R, and failure cost C_F.

$$C = C_I + C_R + C_F \tag{5.6}$$

These costs do not include any user costs associated with the repair actions. The user costs are considered separately in terms of the reduced benefits $\Delta B(e)$. It is assumed that the reduced benefits $\Delta B(e)$ are additive and proportional to the number of repaired elements. Due to the uncertainties associated with the physical process of the deterioration, the maintenance costs can be considered as random variables. As the inspection interval decreases, the repair cost increases, while the probability of failure decreases, and vise versa. The additional user cost due to the lack of budget is assumed to be proportional to the user cost for repair:

$$\Delta B(e,'C > K') = g\Delta B(e) \tag{5.7}$$

where g is a multiplying factor. In order to see the significance of the additional user cost to the optimal inspection interval, g is set to 2, 10 and 100.

The total life cycle period considered for the maintenance planning is 200 years and budgeting is assumed to be made annually. The discount rate is assumed to be 2% per year equivalent to the economic growth per capita. The discount rate by time preference is neglected in this paper for simplicity. This may be justified for short budgeting periods, i.e., 1 year, by the result in Nishijima et al. (2007). The cost parameters assumed in this example are summarized in Table 5.1 together with other parameters.

Table 5.1. Parameters assumed in the example.

Number of structures	50
Number of elements in each structure	100
Total life cycle time to be considered	200 years
Discount rate by economic growth	2% per year
Inspection cost for each structure	1
Repair cost for each element	1
Failure cost for each element	10
User cost for each repair	1
User cost for each failure	10
Multiplying factor g	2,10 and 100

5.3.5. Numerical results

The optimization of the amount of budget and inspection interval is made based on Monte Carlo simulations in accordance with the probabilistic model for corrosion and cost model. Figure 5.3 shows the probabilities of the number of elements to be repaired each year. For the purpose of simplicity it is assumed that repairs are made immediately after inspections if necessary whether or not the budget is available. The differences between the case where the budget is available for repair and the case where the budget is not available are considered through the additional user cost. This assumption significantly simplifies the analysis, while there is little difference in the assessment of probabilities of the number of elements to be repaired each year. The probabilities of the number of failed elements each year are also simulated. The deviation of the numbers of elements to be repaired and the number of failed elements is of relevance for the optimization of the budget for each year. If the budget is insufficient additional user costs may be implied. On the other hand, if the requested budget is too large, the net benefit decreases.

First, the optimization is made for the optimal budget for each year, for a given inspection interval. Figure 5.4 (left) shows an example of the optimization of the budget. The expected total cost $E[CT]$, which is defined in Equation (5.4), becomes large as the multiplying factor, g, becomes large. Accordingly, the optimal budget which minimizes the expected total cost becomes large as g becomes large. After the budget for each year is optimized, the expected total costs for all years are summed up weighted with the corresponding discounting factors. In Figure 5.4 (right), the discounted expected total costs are shown for each inspection interval. The optimal inspection interval is obtained as the one which minimizes the discounted expected total cost. As the multiplying factor increases, the corresponding discounted expected total cost increases and the optimal inspection interval decreases. Since the optimal budget for each year for a given inspection interval is already obtained, the optimal combination of budget and inspection time is derived, see Figure 5.5. The higher “penalty” due to the lack of budget, which is represented by the multiplying factor g, is reflected in the optimal amount of budget in Figure 5.5. In both cases of $g = 10$ and $g = 100$, the expected financial costs remain the same, while the optimal budget is higher in the case of $g = 100$ than in the case of $g = 10$, reflecting the precautionary attitude toward larger consequences due to the lack of budget. It should be mentioned that the periodic fluctuations of the expected financial cost and the optimal budget come from the different number of structures to be inspected as mentioned in Section 5.3.2.

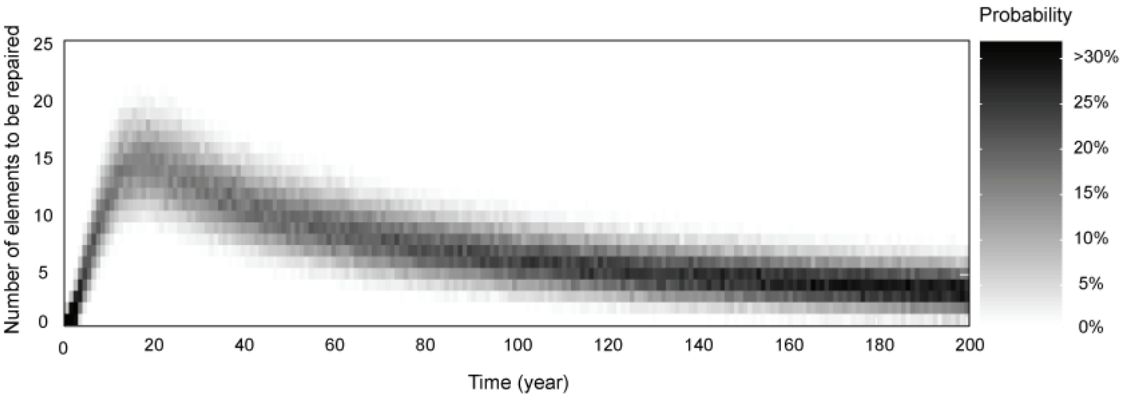


Figure 5.3. Probability of number of elements to be repaired each year (inspection interval: 5 years).

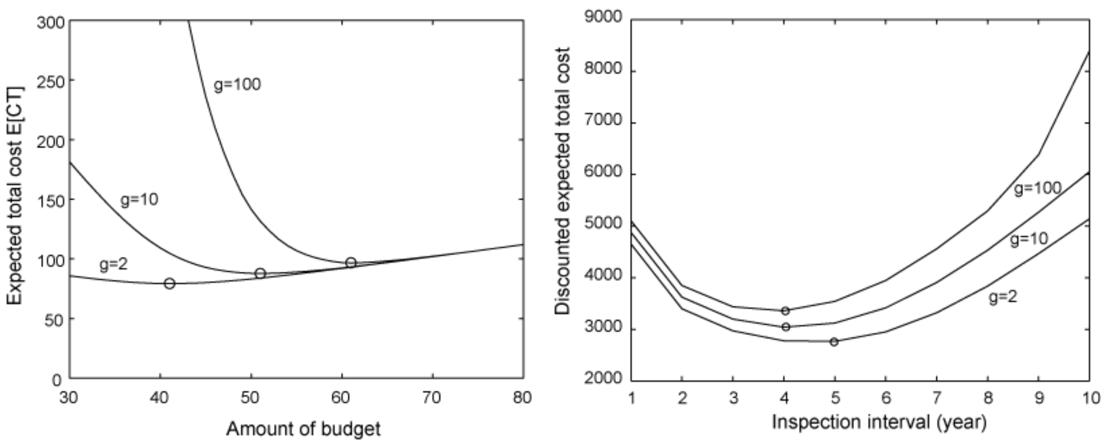


Figure 5.4. Optimal amount of budget at 20th year in the case of inspection interval: 5 years (left) and optimal inspection intervals (right).

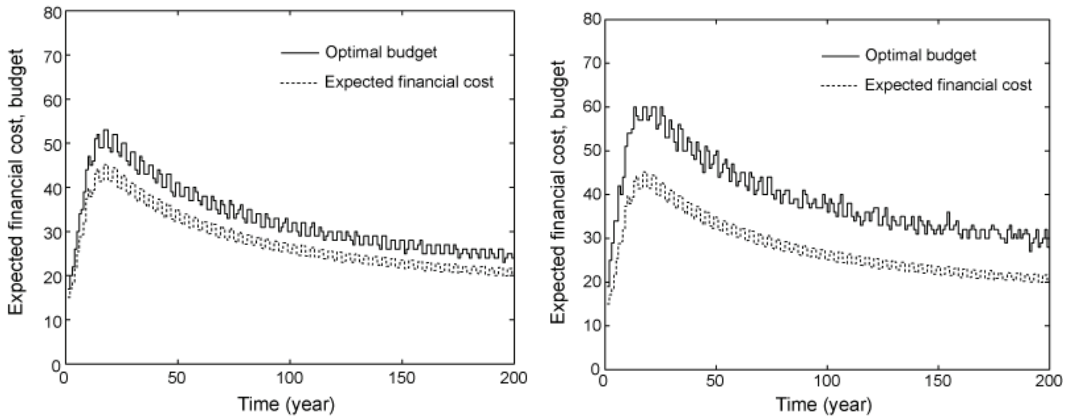


Figure 5.5. Optimal budget and expected financial cost at each year for g = 10 (left) and g = 100 (right), (not discounted).

5.4. Discussions

In the example the features and advantages of the proposed approach are shown considering maintenance planning for RC structures. The approach works especially well in the case of relatively high probability of occurrence of adverse events and relatively low consequences. For the case where the occurrence probability of adverse event is relatively small and the consequence is relatively large, e.g., floods or earthquakes, the annual budget approach may not work well. In such situations, establishment of a fund shared by projects, which corresponds to the integration of projects into one portfolio, may be a good strategy. However, the basic idea in the proposed approach, namely, optimization of budgeting by maximization of the net benefit still works in these situations. In the present example, optimal budget distribution over time has a sharp peak, which is inconvenient in practical budgeting. However, the budget distribution will be averaged out by considering a portfolio which is composed of structures of different ages. Thus, the budget distribution over time shown in the example is due to the fact that only structures whose ages are identical are considered; not indicating a limitation of the present approach.

In regard to the net benefit induced by a project, it is assumed that the unused portion of the budget in the case where the cost does not exceed the budget is lost. However, this can underestimate the net benefit, since the unused portion of the budget can be spent for relevant activities: in the case of the example, for instance, it could be used for additional (unplanned) inspections. In applications, this aspect should be properly taken into account.

Finally, the assumption made in the simulation that repairs are made immediately after inspections if necessary whether or not the budget is available, may not be suitable if the repair time is crucial. The repair time is, in general, dependent on when the budget is available, therefore, the budget for one year does affect the time for repair, which must be reflected in the simulation of deterioration. The repair time also affects the user cost associated with the delay of repair due to a possible insufficient budget. As the delay increases, the user cost increases.

5.5. Conclusions

Optimal decision making for maintenance of structures is addressed from a societal perspective as an optimal budget allocation problem. An approach to find the optimal budget to be allocated and the corresponding optimal inspection and maintenance strategy is proposed. Thereby the expected net benefit is adopted as the objective function to be maximized. In addition to the user costs associated with repair activities the user cost which might result from postponed repair and consequential reduced availability due to insufficient budget is taken into account.

The proposed approach provides a rational framework for decision makers responsible for the budgeting and planning of maintenance activities for portfolios of structures and leads to optimal budgets which are consistent with the adverse consequences of possible insufficient budgets. For the purpose of illustrating the application of the proposed approach the problem of maintenance planning for a portfolio of RC structures subject to chloride-induced deterioration is considered. The example clearly shows that the optimal budgets differ from the commonly applied expected total costs and this also has an effect on the optimal choice of inspection plans.

6. Societal performance of infrastructure subject to natural hazards (Paper V)

Kazuyoshi Nishijima

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 22.3, Zurich 8093, Switzerland.

Michael Havbro Faber

Institute of Structural Engineering, ETH Zurich, ETH Hönggerberg, HIL E 23.3, Zurich 8093, Switzerland.

Australian Journal of Structural Engineering, Vol. 9, No.1, pp. 9-16, 2009.

Abstract

The present paper proposes a methodology for assessing the effect of different design and maintenance policies for infrastructure on societal economic growth. The approach adopted takes basis in the general economic theories and economic models, and provides an interface between economics and civil engineering with which the engineering knowledge can be reflected in the economic models. The proposed methodology can be utilized by societal decision makers to identify the optimal investments into infrastructure for ensuring sustainable societal development. An illustrative example is provided considering sustainable decision making in regard to design and maintenance of infrastructure subject to natural hazards. Thereby the advantage of the proposed methodology is shown; it enables one to analyze the economic growth and the associated uncertainties corresponding to different design and maintenance policies for infrastructure.

Keywords

Sustainability, societal decision making, reliability theory, economic theory.

6.1. Introduction

Sustainable societal development has become an issue of increased and wide spread societal attention especially during the last two decades. The tremendous economic developments of former third world nations such as China and India and the general impact of globalization have put even larger pressures on our limited natural resources and fragile environment. Faced with an ever increasing amount of evidence that the activities of our own generation might actually impair the possibilities for future generations to meet their needs it has become a political concern that societal development must be sustainable. The issuing of the famous Brundtland report "Our Common Future" (Brundtland (1987)) forms a milestone on the political arena. This important event has enhanced the public awareness that substantial changes of consumption patterns are called for and has further significantly influenced the research agendas worldwide; it is fair to state that "sustainable development" has strongly influenced the consciousness and the moral setting in society.

Recent disasters caused by natural hazard events, e.g. the tsunamis in Southeast Asia in 2004 and the flood induced by the hurricanes in the United States of America in 2005, have proven the importance of infrastructure in society and revealed how societies in both developing countries and developed countries supported by infrastructure are vulnerable to natural hazards. Recognizing the lesson learned from these recent disasters it is necessary to reconsider the framework for identifying the optimal level of reliability of infrastructure in regard to the performance with due consideration of the role that the infrastructure plays for societies.

Infrastructure such as road networks, water and electricity distribution systems assists economic growth. Aschauer (1989) has reinforced this perception by showing that investment into infrastructure has a strong explanatory power for societal productivity taking up the case of the United States of America. A number of studies have confirmed and generalized this observation; some of these studies, however, claim that the estimated return rates of investment into infrastructure might be biased, see the review paper by Gramlich (1994).

In the field of civil engineering, the life cycle cost (LCC) optimization concept has gained a reputation as being a means for identifying optimal designs as well as maintenance strategies for infrastructure with due consideration of possible consequences and proper discounting for future expenses. More recently, the LCC optimization concept often has been applied in the context of sustainable decision making for infrastructure projects. However, the application of the LCC optimization concept in this context may not be appropriate since it tends to focus on the marginal analysis of the benefits and the costs of projects. For instance, the LCC optimization concept implicitly assumes that the necessary budget is available whenever it is needed, which in practice is not necessarily true. Nishijima and Faber (2006) discuss this issue taking into account the opportunity costs that the lack of budget may incur. Furthermore, the LCC optimization concept does not aim to identify how to optimally allocate limited resources into different projects; it primarily addresses the optimization of each individual project or a portfolio of projects assuming these projects are in any case undertaken. This is especially problematic in the context of sustainable decision making, since sustainability fundamentally concerns the issue of allocation of limited resource in different societal sectors and projects. From this perspective, the optimization problem in the context of sustainable decision making should be formulated as: 1) given the amount of investment into the civil engineering sector, how much of the investment should be directed to new construction and maintenance works respectively and then 2) at the level of societal decision making how much of the investment should be allocated to the civil engineering sector. Whereas the latter optimization is conducted from the perspective of societal decision makers, the former optimization is a civil engineering issue. However, these two optimizations have never been discussed jointly due to the lack of the interface between civil engineering and economics.

Economics plays the central role in analyzing the development of society in the most aggregated way. It considers not only economic development but also environmental issues, societal preferences regarding e.g. issues of human safety and inter- and intra-generational equity etc. The general discussion on the implications of sustainability is also ongoing in the field of economics, although no agreement is yet established, see e.g. Perman et al. (2003); the present paper assumes that the agreement on implications of sustainability should be made in the general economics. Therefore, the present paper does not aim at defining the objective function and constraints concerning sustainable societal development but at providing a methodology with which economic output, which is one of the relevant indicators concerning sustainable societal development, can be evaluated as a function of the amount of investments allocated to the civil engineering sector.

The main problem in employing general economic models in sustainable decision making in connection with civil engineering is that they do not account for the performance of infrastructure based on scientific and engineering knowledge; mostly they are based on aggregated statistical analysis using historical data. Therefore, it is difficult to study the effects of different design and maintenance policies on societal economic growth.

With this background the present paper proposes an interface between the general economic theories and civil engineering. The proposed methodology takes basis in the methodology proposed by Nishijima and Faber (2007c). However, an extension is made such that the losses of infrastructure capital due to natural hazards can be considered in an explicit probabilistic manner. After formulating the methodology this is applied for the investigation of the effect of different target reliabilities for infrastructure facilities on the economic growth and the degree of uncertainties associated with the economic growth.

6.2. Problem setting

Public infrastructure is the primary concern in the present paper, e.g., road networks, water and electricity distribution systems, for which societal decision makers, to a large extent, can decide the amount of investment for design and maintenance policies. The methodology presented in Section 6.4 can be partly applied to private infrastructure, e.g. machinery and residential houses. The question, however, remains whether sustainable decision making can be expected from private stakeholders; societal policy measures, e.g. imposing taxes and giving subsidies, may be required in order to lead decision makers in private sectors to societal optimal actions.

Two issues are addressed in the present paper. The first concerns the reliability of infrastructure facilities. Economic models must be able to account for the different reliabilities of infrastructure facilities resulting from different design and maintenance policies. In general, the deterioration rate of infrastructure depends on the target reliability in regard to any type of reduction of performance of the infrastructure and thus depends on the policy in regard to design and maintenance. Usually, in the field of economics the deterioration rate is estimated directly or indirectly based on historical data, see e.g. Aschauer (1989), Gramlich (1994) and Greenwood et al. (2000). Using historical data as a basis, however, there is no possibility to reflect the effect of new policies on the future deterioration of infrastructure facilities. The proposed idea in the present paper is that the reliabilities of infrastructure facilities in the engineering sense can be related to the deterioration rates in an economic sense. Secondly, two different types of investments into infrastructure should be differentiated; 1) the investment into new construction of infrastructure facilities, which will increase the economic output through the increased infrastructure capital stock, and 2) the investment for achieving higher reliability of infrastructure facilities, which does not directly increase the economic productivity but improves the durability of the structures and prolongs their lifetime. The distinction between these two different types of investments is realized by assessing the infrastructure capital stock by physical units, as opposed to monetary units.

The necessity of increased investments into infrastructure in terms of maintenance works has been appreciated for both developing and developed countries for different reasons. In developed countries the investments into maintenance are considered as an urgent issue in the light of the severe deterioration of aged infrastructure. For developing counties on the other hand the necessary investments into maintenance works have been considered as a potential opportunity for increasing the investment efficiency of expenditures into the built environment; the investment for deteriorating infrastructure into maintenance works may be more efficient than the investment into construction of new infrastructure, see e.g. World Bank (1994), Rioja (2003) and Kalaitzidakis and Kalyvitis (2004). The present methodology is formulated allowing for considering both types of investments.

6.3. Role of infrastructure in economic context

The role of infrastructure in an economic context is illustrated in Figure 6.1. The performance of infrastructure must reflect societal needs in regard to productivity and societal preferences, for instance, concerning life safety and damages to the qualities of the environment. Societal preferences in regard to life safety have been discussed in the context of economic output and consumption through the recently developed concept of the Life Quality Index, see e.g. Nathwani et al. (1997) and Rackwitz (2002). These considerations fall into the category of how to define a utility function and/or constraints in the context of decision making. The present paper, in contrast, focuses on the relation between the productivity of societal infrastructure versus investments into new and existing infrastructure.

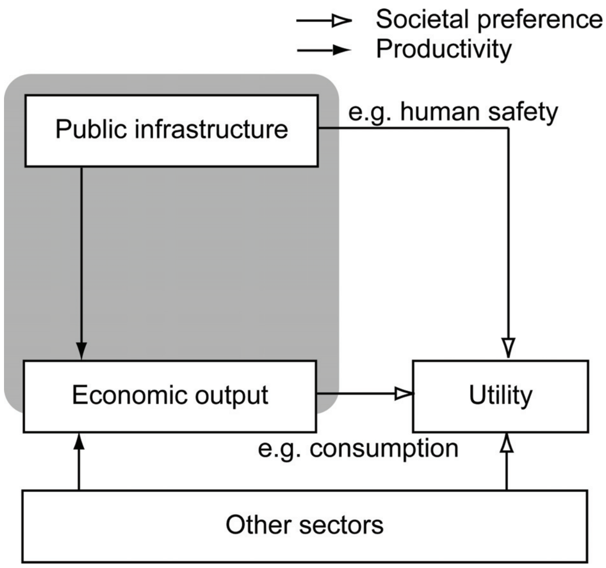


Figure 6.1. Focused role of infrastructure in an economic context, after Nishijima and Faber (2007c).

The technology currently available determines the level of economic output given the amounts of different types of capitals, e.g. human capital, physical capital etc. This relation is in the field of macroeconomics often represented by a production function as:

$$Y = f(K^{(1)}, K^{(2)}, ...) \tag{6.1}$$

where Y is the economic output in a given period, $K^{(i)}$ $(i = 1, 2, ...)$ represents the amounts of different types of capitals. The level of differentiation of capitals depends on the level of analysis and data available. For instance, $K^{(n)}$ may represent the amount of the aggregated capital of infrastructure including different types of infrastructure or may represent one specific type of infrastructure. It is also possible that $K^{(m)}$ represents the amount of a capital differentiated according to its relevance that, however, belongs to the same type of infrastructure, e.g. road networks that connect large cities versus road networks in remote areas. In what follows, however, only one type of the capitals is considered and it is abbreviated as $K^{(1)} = K$ for the purpose to make clear the concept of the proposed methodology, and it is not the limitation of the proposed methodology.

The production function can be estimated using historical data by time-series analysis and/or cross-sectional studies. Thereby the capitals are measured in physical terms e.g. kilowatts of electricity generating capacity or length of road or in monetary terms (by multiplying the amount measured in physical units with the corresponding prices). However, for the present purpose it is important to measure the infrastructure capital in physical terms since otherwise the investment for achieving higher reliability and the investment for increasing the amount of infrastructure cannot be distinguished. Several datasets and estimated production functions in regard to several types of infrastructures are available, e.g. Canning (1998) and Canning and Bennathan (2000).

The equation of capital accumulation is often written in the following form:

$$\Delta K_t = K_t^{new} - \delta K_t \tag{6.2}$$

where the subscript t in K_t represents that the amount of capital K is evaluated at time t, ΔK_t is the net increment of the amount of the capital between time t and $t+\Delta t$, Δt is the increment of time, K_t^{new} is the amount of infrastructure capital constructed between time t and $t+\Delta t$, and δ is the deterioration rate. Note that the amount of the capital is measured in physical units and Δt is often chosen as $\Delta t = 1$ year. As mentioned previously, the deterioration rate δ is usually estimated using historical data and it is often represented as a deterministic value. The exceptions for this are Bulow and Summers (1984) and Zeira (1987), who consider the uncertainty in depreciation of capital[9]. From a civil engineering point of view the amount of deteriorated capital between time t and $t+\Delta t$ represented by δK_t in Equation (6.2) indeed is a function of design and maintenance policies. In general, the amount of deteriorated capital should be considered as a random variable unless it may be assumed to converge to its expected value; this may not be the case for infrastructure facilities subject to natural hazards when the geographical sizes of the hazard events are relatively large compared to the sizes of societies. In the following section a methodology for solving these issues is proposed.

6.4. Proposed methodology

6.4.1. Definition of infrastructure failure

Let R_t denote a set of states that represent the performance of an infrastructure facility at time t. R_t may consist of not only physical states of infrastructure facilities but also societal states of relevance that are related to the use of the infrastructure facilities. The infrastructure facility is considered to have failed if the performance of the infrastructure facility does not satisfy the societal requirements. The failure of an infrastructure facility may occur e.g. due to natural hazards, physical deterioration and societal obsolescence. The societal requirements to the infrastructure facility are assumed expressed through the failure domain $\Omega_{F,t}$. The failure domain $\Omega_{F,t}$ can be a composite set of single-failure events, each of which relates to different

9 They consider the uncertainty of capital depreciation in terms of uncertain changes of the monetary value of capitals. Thus, the depreciation therein does not concern the change of the amount of physical capital due to e.g. natural hazards.

societal requirements. Examples hereof include collapse of a structure, severe deterioration where repair actions are not feasible as well as situations where the safety of a structure does not fulfill given acceptance criteria and must be demolished and/or replaced. Then failure may be defined as:

$$R_t \in \Omega_{F,t} \tag{6.3}$$

The conditional probability $p_{F,t}$ of failure of infrastructure in time period $(t, t+\Delta t]$ is defined as:

$$p_{F,t} = P\left[R_t \in \Omega_{F,t}; (t, t+\Delta t] \mid R_t \notin \Omega_{F,t}; [0,t]\right] \tag{6.4}$$

In cases where the failure domain consists of m independent failure event sets $\Omega_{F,t}{}^{(i)}$ ($i = 1, 2, ..., m$) and $\Omega_{F,t} = \bigcup_i \Omega_{F,t}{}^{(i)}$, the conditional probability of failure can be written as:

$$\begin{aligned} p_{F,t} &= P\left[R_t \in \bigcup_i \Omega_{F,t}{}^{(i)}; (t, t+\Delta t] \mid R_t \notin \Omega_{F,t}; [0,t]\right] \\ &= 1 - \prod_i \left(1 - P\left[R_t \in \Omega_{F,t}{}^{(i)}; (t, t+\Delta t] \mid R_t \notin \Omega_{F,t}; [0,t]\right]\right) \end{aligned} \tag{6.5}$$

In this way, the conditional probability of failure defined in terms of Equation (6.4) can be regarded as a generalized measure of capital deterioration. The advantage of the definition of infrastructure failure in this manner is that it enables the use of the reliability theory for the calculation of the probabilities corresponding to the structural design and maintenance policies for infrastructure facilities whenever probabilistic models are available. Otherwise the probabilities estimated by expert judgments can be partly integrated into probabilistic terms in Equation (6.5), hence, it is possible to combine objective and subjective evaluations in order to quantify the conditional probability of failure.

6.4.2. Equation of capital accumulation

The increment ΔK_t of the infrastructure capital from time t to $t+\Delta t$ can be generally written taking basis in Equation (6.2) as:

$$\Delta K_t = K_t^{new}(a_t) - X_t \tag{6.6}$$

where $K_t^{new}(\cdot)$ is the amount of new infrastructure constructed at time t and X_t is the amount of failed infrastructure. In general, X_t should be considered as a random variable. Note that applying expectation operation Equation (6.6) is reduced to

Equation (6.2). $K_t^{new}(a_t)$ is a function of the design policy a_t at time t and is written as:

$$K_t^{new}(a_t) = \frac{I_t}{q_t(a_t)} \tag{6.7}$$

where I_t is the budget allocated to construction of new infrastructure at time t and $q_t(a_t)$ is the unit cost of the construction corresponding to the design policy a_t. The probability distribution of the amount of failed infrastructure X_t is characterized by the amount of the capital K_t at time t, the sequence of design policies $\{a_i\}_{i=1}^t = \{a_1, a_2, ..., a_t\}$ and the sequence of maintenance policies $\{b_i\}_{i=1}^t = \{b_1, b_2, ..., b_t\}$ for the infrastructure until time t. In cases where large-scale hazards, e.g. earthquakes and hurricanes, are of concern the geographical distribution of the infrastructure is also a relevant factor. Finally, since the budget allocated to infrastructure is divided into the investments into new construction and maintenance works the following equation must hold:

$$G_t = I_t + M_t(\{a_i\}_{i=1}^t, \{b_i\}_{i=1}^t) \tag{6.8}$$

where G_t is the allocated budget for the civil engineering sector at time t and M_t is the budget necessary for maintenance works. M_t is a function of $\{a_i\}_{i=1}^t$ and $\{b_i\}_{i=1}^t$. With these settings it is possible to identify the optimal design policies $\{a_i^*\}_{i=1}^t$ and maintenance policies $\{b_i^*\}_{i=1}^t$ given the budget sequence $\{G_i\}_{i=1}^t = \{G_1, G_2, ..., G_t\}$.

The methodology proposed above requires as an input parameter the amount of investment into infrastructure, considers the design policy and the maintenance policy to be decision variables to be controlled and provides as outputs the sequence of the amount of capital K_t and the corresponding economic growth Y_t.

6.5. Illustrative example

The economic model in the following example assumes that infrastructure is the only capital that affects the economic production in society. The production function is assumed to be written as:

$$Y_t = AK_t^{\alpha} \tag{6.9}$$

which is a special form of Equation (6.1). Therein A is the factor that represents the technology in the society, which is assumed constant. The exponent α represents the marginal increase of the economic output with respect to the infrastructure capital. It is assumed that the infrastructure capital is exposed to natural hazards and that the infrastructure capital can be geographically divided into n segments within which the failures of infrastructure facilities are perfectly correlated and between which the failures are independent. Namely, the parameter n represents the relative geographically affected size of the natural hazards compared with the size of the society. Furthermore, the occurrence of natural hazards is assumed temporary independent. Under these assumptions the amount of capital which is lost at time t can be expressed as:

$$X_t = \frac{N_t}{n} K_t \tag{6.10}$$

where N_t represents the number of failed segments among n independent segments with the probability of failure p_f and follows the binomial distribution with n n trials and the probability of failure being equal to p_f. p_f is the probability of failure within the duration $\Delta t = 1$ year. Note that as n becomes large X_t converges to its expected value, $E[N_t]/n \cdot K_t = p_f K_t$ thus the equation of capital accumulation is reduced to the form of Equation (6.2). By substituting Equation (6.10) into Equation (6.6), the equation of capital accumulation is written as:

$$\Delta K_t = \frac{I_t}{q_t(a_t)} - \frac{N_t}{n} K_t \tag{6.11}$$

The values of the parameters assumed in this example are shown in Table 6.1. These values are postulated for illustrative purposes, however, in practice these values can and should be determined by economic as well as engineering analyses.

Table 6.1. Assumed parameters in the example.

Investment ratio into infrastructure	$\lambda = 0.05$	
Exponent in production function	$\alpha = 0.2$	
Factor in production function	$\mathrm{A} = 10$	
Independent segments of infrastructure	$n = 5, 50$	
	Policy1	Policy2
Probability of failure per year	$p_f = 0.01$	$p_f = 0.001$
Construction cost per unit	$q_t = 1$	$q_t = 2$
Maintenance cost	$0.01K_t$	$0.01K_t$

The probability of failure p_f is a function of the policy in regard to design and maintenance. Here, two policies are considered, each of which targets the probability of failure shown in Table 6.1. The corresponding construction costs and maintenance costs are also shown in the table. In practical situations, the probability of failure and the associated costs can be identified using the definition of infrastructure failure represented by Equation (6.3) employing the civil engineering knowledge and the structural reliability theory for the calculation of Equation (6.4).

The analyzed economic output paths as the function of the policies and different numbers of independent segments are shown in Figure 6.2. The figure shows the median, 5% and 95% of the economic output as a function of time. Figure 6.2 (left) shows the economic output paths when the number of independent segment is relatively small ($n=5$). The economic growth is faster when policy 1 (lower reliability associated with lower construction cost) is adopted. However, in a long run the economy grows more when policy 2 (higher reliability associated with higher construction cost) is adopted. It should be mentioned that the economic growth path under policy 1 is associated with larger uncertainty, i.e. results in a less stable economic growth, compared to the economic growth path under policy 2. The economic growth paths are more stable in a sense that the uncertainty on the economic output is smaller when the number of independent segments is larger ($n=50$), see Figure 6.2 (right). The results shown in Figure 6.2 and the interpretation of the results stated above are coherent with engineering understanding. Furthermore, it is possible with the proposed methodology to evaluate in a quantitative manner the effects of different policies on the economic growth within a general economic model framework.

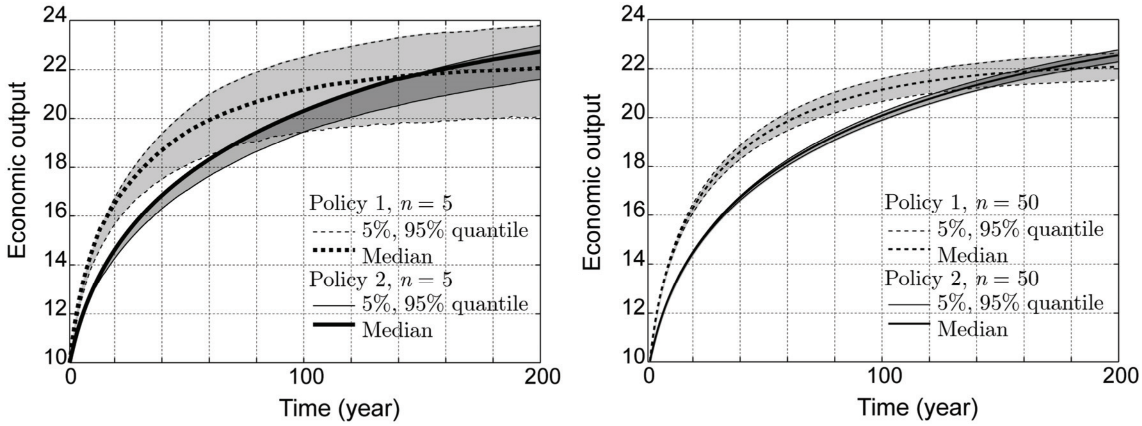


Figure 6.2. Economic output paths for different policies and different numbers of independent segments $n=5$ (left) and $n=50$ (right).

6.6. Discussion

In practical applications, the amount of infrastructure capital losses due to natural hazards can be readily assessed by risk analysis together with Geographical Information Systems (GIS), see Bayraktarli and Faber (2007). The design costs and maintenance costs for infrastructure facilities corresponding to different policies can be optimally identified using the framework proposed by Nishijima et al. (2008). This framework models the infrastructure using hierarchical Bayesian networks and formulates the problem as a constrained optimization problem where the expected costs are considered as the objective function and the requirements to the performance of infrastructure, e.g. target reliabilities and acceptance criteria for fatalities, are accounted for by constraints. These techniques can be incorporated into the proposed methodology. The assumption that failures of structures are independent generally does not hold even if the hazard events that affect each structure are independent. This is because of the presence of modeling uncertainties, e.g. on the resistance of structures, that may commonly affect all considered structures, see Faber et al. (2007a). Thus the proposed methodology and the analysis in the example of this paper should be considered as being conditional on the modeling uncertainties. In general the integration with respect to the modeling uncertainties is necessary in the analyses of losses of infrastructure facilities and societal economic growth.

6.7. Conclusion

The present paper proposes a methodology for assessing the effect of different design and maintenance policies for infrastructure on societal economic growth. The proposed methodology can serve as a component of a general decision making framework for optimal resource allocation in the context of sustainable societal development. The proposed methodology requires the amount of investments into infrastructure as an input parameter. It incorporates the design policy and the maintenance policy as decision variables. It provides the sequence of the amount of capital together with the corresponding economic growth as outputs. In an example the advantage of the proposed methodology is illustrated; it enables one to analyze in a quantitative manner the economic growth and the economic stability corresponding to different design and maintenance policies for infrastructure.

7. Optimal design and maintenance policy on infrastructure from a macroeconomic perspective

7.1. Introduction

In Chapters 3, 4 and 5, the optimization problem of the reliability of individual structures or groups of structures is addressed. In these chapters, the reliability or decision variables related to structural performance are optimized based on the life-cycle cost optimization concept. Strictly speaking, the life-cycle cost optimization concept can be applied only if the benefit and cost of the project concerned are assumed marginal in the economy; that is, the economic growth is not affected by whether or not or how the project is undertaken. Thus, the life-cycle cost optimization concept may not be appropriate as the philosophical principle for decision making in cases where the consequences of the decisions are considered as non-marginal.

In practice, there are many situations where the consequences of the decisions are considered as non-marginal. Such decision situations include, for example, code making in which the acceptable reliability of structures is controlled, and design and maintenance strategies on nationwide infrastructure projects. These decisions affect the capital accumulation of infrastructure and thus, in turn, the long-term development of the economy. Therefore, in these decision situations a non-marginal economic framework has to be adopted.

As a first step to develop a general decision framework for facilitating these decision situations, this chapter examines how the optimal reliability of infrastructure may be identified within the economic growth theoretical framework. For this, a simplistic economic model is developed, employing the approach proposed in the previous chapter for incorporating the reliability of infrastructure in economic models. Using the developed economic model, it is investigated how the reliability of infrastructure affects the economic growth and the optimal reliability at each point in time depends on the economic level. The aim of this chapter is to show the potential that such a general framework can provide the optimization principle for non-marginal decision analysis.

The structure of this chapter is as follows. First, the principle of the life-cycle cost optimization concept is reviewed. Then, the assumptions and limitations of the concept are pointed out. Second, previous research work on the role of civil infrastructure within the economic growth theoretic framework are introduced briefly, followed by some critical reviews on the assumptions made in these works. Third, a simplistic economic model is presented. Finally, the optimal reliability of infrastructure is examined within the model, and the results are discussed.

7.2. Principle of life-cycle cost optimization concept

The life-cycle cost optimization concept is considered as an extension of cost-benefit analysis. Thus, before deriving the life-cycle cost optimization concept, the derivation of the principle of cost-benefit analysis is introduced. The derivation introduced here is based on Stern (2006)[10].

7.2.1. Derivation[11]

A project is socially profitable if the social welfare is increased through the project. This is expressed as:

$$\Delta W = W^1 - W^0 > 0 \tag{7.1}$$

where W is the social welfare function, and W^0 and W^1 are the social welfares when the project is not undertaken and the project is undertaken respectively. In general, the social welfare function is a function of many variables that concern the utilities of all members in the society. However, here it is assumed that the social welfare function consists of the utility function of a representative household and a discount factor, and the utility is a function only of the consumption of the household. Under these assumptions, the social welfare function can be written as:

$$W = \int_0^\infty u(c_t) e^{-\rho t} dt \tag{7.2}$$

where $u(c_t)$ is the utility function of the representative household, c_t is the consumption at time t, and ρ is the discount rate for pure-time preference. Assuming that the change of the consumption in Equation (7.2) is small, and substituting Equation (7.2) into Equation (7.1), ΔW can be written as:

$$\Delta W = \int_0^\infty \frac{\partial u(c_t)}{\partial c_t} \Delta c_t e^{-\rho t} dt = \int_0^\infty \lambda_t \Delta c_t dt \tag{7.3}$$

where Δc_t is the perturbation of the consumption from a baseline consumption path, and λ_t is the discount factor and is written as:

10 Other derivations can be found in e.g. Ramsey (1928) and Koopmans (1965) in the context of the economic growth theory.

11 For simplicity, here it is assumed that a representative individual lives for an infinite time. However, the derivation can be extended for the case where many generations live for finite lifetimes, which is the situation assumed in the generation-adjusted discounting concept introduced in Chapter 4. Furthermore, it is assumed that the population is constant over time.

$$\lambda_t = \frac{\partial u(c_t)}{\partial c_t} e^{-\rho t} \tag{7.4}$$

Here, the increase/decrease of consumption at each point in time corresponds to the benefit/cost from the project. The rate of the temporal change of the discount factor $\dot{\lambda}_t / \lambda_t$ is obtained as[12]:

$$\begin{aligned}\frac{\dot{\lambda}_t}{\lambda_t} &= \frac{\{u''(c_t)\dot{c}_t - \rho u'(c_t)\} e^{-\rho t}}{u'(c_t) e^{-\rho t}} \\ &= \frac{c_t u''(c_t)}{u'(c_t)} \frac{\dot{c}_t}{c_t} - \rho \\ &= -(\eta \delta_t + \rho)\end{aligned} \tag{7.5}$$

where $u'(c_t) = \partial u / \partial c_t$, $u''(c_t) = \partial^2 u / \partial c_t^2$, $\eta = c_t u''(c_t) / u'(c_t)$ and $\delta_t = \dot{c}_t / c_t$. η is the elasticity of the marginal utility of consumption. δ_t is the growth rate of consumption, which is assumed to be exogenously given.

If the growth rate of consumption is assumed constant and given as $\delta_t = \delta$, then the discount factor is obtained as[13]:

$$\lambda_t = e^{-(\eta\delta + \rho)t} \tag{7.6}$$

By substituting Equation (7.6) into Equation (7.3), the criterion for the project appraisal is finally obtained as:

$$\Delta W = \int_0^\infty \Delta c_t e^{-(\eta\delta + \rho)t} dt > 0 \tag{7.7}$$

In cases where several decision alternatives are available for the project, the above criterion should be applied for the decision alternative that maximizes ΔW.

The life-cycle cost optimization concept typically employed in civil engineering decision analysis is derived by further assuming that the perturbation of the consumption $\Delta c_t(a)$ is a function of the decision variable a regarding structural performance and this is equal to the benefit B_t less the cost $C_t(a)$, which is also a function of the decision variable a as:

$$\Delta c_t(a) = B_t - C_t(a) \tag{7.8}$$

12 The dot "$\cdot$" on the top of symbols represents the derivative with respect to time.
13 The choice of the constant λ_0 is arbitrary, thus here it is chosen as $\lambda_0 = 1$.

Note that the benefit B_t may vary as a function of time but is often assumed to be independent of the decision variable a. By substituting Equation (7.8) into Equation (7.7), neglecting the constant benefit term, and taking the negative sign, the objective function, i.e. the life-cycle cost $C_T(a)$ is obtained as a function of the decision variable a as:

$$C_T(a) = \int_0^\infty C_t(a) e^{-(\eta\delta+\rho)t} dt \qquad (7.9)^{14}$$

Whenever uncertainty is involved in the cost term, the expectation should be taken as:

$$\overline{C}_T(a) = \int_0^\infty E[C_t(a)] e^{-(\eta\delta+\rho)t} dt \qquad (7.10)$$

where $\overline{C}_T(a)$ is the expected life-cycle cost as a function of the decision variable a, and this should be employed as the objective function in the optimization.

7.2.2. Assumption and limitation

The fundamental assumption in the derivation of the life-cycle cost optimization concept shown in the above is that the growth rate δ_t of the consumption is exogenously given and is not affected by the benefits and costs from the project; namely, the benefit from the project at each point in time is consumed at the time (i.e. not invested for capital accumulation), and costs incurred by the project at each point in time are compensated by the decrease of consumption at the time (thus the amount of investment remains unchanged). Note that the stock losses of the infrastructure capital due to failure (direct consequence of failure) and the economic flow losses associated with the capital losses (indirect consequence of failure) in case of failures should be interpreted as reduced benefits, which are also assumed not to affect the growth rate of consumption. The application of the life-cycle cost optimization concept should be limited to the extent that the assumption can be considered as reasonable.

As is clear from the above derivation, the growth rate δ_t of consumption does not need to be constant, although in practice it is often assumed constant. It may be also worth mentioning that whenever uncertainty is involved in the discount rates δ and ρ, the expectation should be taken as[15]:

$$\overline{C}_T(a) = \int_0^\infty E[C_t(a)] E[e^{-(\eta\delta+\rho)t}] dt \qquad (7.11)$$

14 If $u(c_t) = \ln c_t$, then $\eta = 1$, and it coincides with the formulation of the objective function in Chapter 4.

15 Here, the cost term is assumed to be independent of the term of the discount factor. However, if they are not considered as being independent, the expectation operator should be only applied to the product of the two terms; this may be the case when some of the costs included in the cost term, which are measured in real terms, may change in accordance with the economic growth.

Note that the expectation operator is applied to the discount factor, not to the discount rates, see e.g. Newell and Pizer (2004) for more discussion.

7.3. Available economic models for infrastructure

In order to describe the role of infrastructure within the economic growth theoretical framework, two types of component economic models are required; one for describing the contribution of the infrastructure capital to the economic productivity, and the other for describing the accumulation of the infrastructure capital. The former is represented in terms of a production function, and the latter is represented in terms of so-called "equation of motion," which describes how the capital is accumulated as a function of the investment into new construction of infrastructure and the deterioration rate of the infrastructure capital.

Concerning the production function that incorporates the infrastructure capital, there are a number of research works available both theoretically (e.g. Glomm and Ravikumar (1994) and Duggal et al. (1999)) and empirically (e.g. Aschauer (1989), Easterly and Rebelo (1993) and Canning and Bennathan (2000)). There are also some research works on the estimation of the deterioration rate of infrastructure capital, see e.g. Gramlich (1994) and Greenwood et al. (2000). However, only a few research works are available that explicitly treat the deterioration rate of infrastructure capital as a variable which can be controlled in terms of maintenance policy on infrastructure, e.g. Rioja (2003) and Kalaitzidakis and Kalyvitis (2004).

For example, Rioja (2003) considers the amount of investment in maintenance work for the infrastructure (relative to economic output) as a control variable, and the optimal investment ratio in maintenance work is derived. Kalaitzidakis and Kalyvitis (2004) extend Rioja's economic model by endogenizing the decision of budget allocation into both investment in the construction of new infrastructure and investment in maintenance work for existing infrastructure.

These pioneering works are remarkable in the sense that the deterioration rate is considered as a variable and can be optimized through the investment ratio into maintenance work. However, the relations between the deterioration rate and the investment ratio assumed in the models are not realistic. One of the drawbacks of these assumptions is that the deterioration rate at any time is dependent only on the current investment ratio in maintenance work; the current deterioration rate is not a function of past maintenance policies, and the current maintenance work does not affect the future deterioration rate. Furthermore, the effect of differing design policies on the deterioration of infrastructure is not considered.

However, in civil engineering it is commonly agreed that a slight increase of initial cost for the purpose of increasing the durability of infrastructure would significantly reduce the future costs for maintenance work. Similarly, undertaking maintenance work at an earlier stage of deterioration would reduce additional maintenance costs in the future. Thus, the investment in construction and maintenance works for reducing the deterioration rate can be considered at least partly as an investment into the future. However, the economic models proposed by those pioneering works may fail to capture this nature of the investment.

In order to overcome these drawbacks of the economic models proposed previously, in the following section, a simplistic economic model that enables one to capture this type of investment is developed, and the economic model is examined.

7.4. Analysis with simplistic economic model

7.4.1. Economic model

The aggregated output $Y(t)$ is assumed to be produced by means of capital $K(t)$ and labor $L(t)$ at time t. This relation is assumed to be represented by the neoclassical production function[16]:

$$Y(t) = F(K(t), L(t)) \tag{7.12}$$

Herein, it is furthermore assumed that the capital $K(t)$ consists only of infrastructure capital. Assuming that the production function exhibits constant return to scale, the production function can be reformulated in terms of variables per capita[17] as:

$$y(t) = \frac{Y(t)}{L(t)} = \frac{F(K(t), L(t))}{L(t)} = F(K(t)/L(t), 1) = f(k(t)) \tag{7.13}$$

where $y(t)$ and $k(t)$ denote the output and capital per capita at time t, and $f(\cdot)$ represents the production function in terms of the variables per capita. In addition to these assumptions, it is assumed that the saving rate of the household is exogenously given as e ($0 < e < 1$) and the amount of labor is constant over time[18].

The important difference in the economic model assumed here from the models employed in Rioja (2003) and Kalaitzidakis and Kalyvitis (2004) appears in the equation of motion for the capital accumulation, especially on the way of modelling infrastructure deterioration.

Consider the infrastructure constructed at time s. The expected service life time $\bar{T}^s$ of the infrastructure and associated costs q^s for construction and maintenance work

16 Namely, $\partial F/\partial K > 0, \partial F/\partial L > 0, \partial^2 F/\partial K^2 < 0, \partial^2 F/\partial L^2 < 0$.

17 Here, it is assumed that the population is equal to the amount of labor capital.

18 Thus, the analysis can be made only in terms of variables per capita. In what follows, the small symbols for the corresponding variables represent the variables per capita.

are assumed to be a function of the design and maintenance policy a^s at time s, i.e. $\bar{T}^s = \bar{T}(a^s)$ and $q^s = q(a^s)$. Herein, the associated costs refer to all the costs that are required in order to realize the target expected service life $\bar{T}(a^s)$. Whereas the service life time of infrastructure is in general a random variable, it is assumed here for simplicity that the service life time is deterministically represented by its expected value, $\bar{T}^s$. Furthermore, it is assumed that the infrastructure provides full functionality until it exceeds the expected service life time and does not provide any functionality when it exceeds the expected service life time. Note that for the assessment of the expected service life time the approach presented in Section 6.4.1 is useful. In this setting the failure of the infrastructure should be interpreted in a broader sense; the relevant failure modes include not only physical collapse but also unavailability of required functionality for any reasons, e.g. severe deterioration, failure to satisfy given acceptable safety level, and even societal obsolescence[19]. The expected service life $\bar{T}^s$ thus defined can be interpreted to represent the reliability of infrastructure; the longer the expected service life of infrastructure, the higher the reliability of the infrastructure.

To consider these properties in the economic model, the following function is introduced;

$$g(t;s,\bar{T}^s) = \begin{cases} 0 & (t < s, s+\bar{T}^s < t) \\ 1 & (s \le t \le s+\bar{T}^s) \end{cases} \tag{7.14}$$

Then the contribution $\kappa^s(t)$ of the infrastructure constructed at time s to the capital accumulation is written as:

$$\kappa^s(t) = k^s \cdot g(t;s,\bar{T}^s) \tag{7.15}$$

where k^s represents the per-capita amount of the infrastructure constructed at time s, see Figure 7.1. The effect of the design and maintenance policy a^s at time s on the deterioration of the infrastructure in the future can be represented through the function $g(t;s,\bar{T}^s)$ in terms of the expected service life time $\bar{T}^s$ of infrastructure.

Assume that all the current and future costs for construction and maintenance work for the infrastructure constructed at time s are invested at time s, and denote the overall cost per unit capital by q^s[20]. Since the amount of investment is assumed given exogenously as $i(s) = ey(s)$, the increment k^s of the capital due to the investment at time s is given as:

19 See Section 6.4.1 for the definition of the generalized capital deterioration.

20 If the variables in the economic model are measured in terms of a physical unit, this cost per unit capital should be interpreted as the multiplying factor for adjusting the difference of the required amount of resources for different design and maintenance policies.

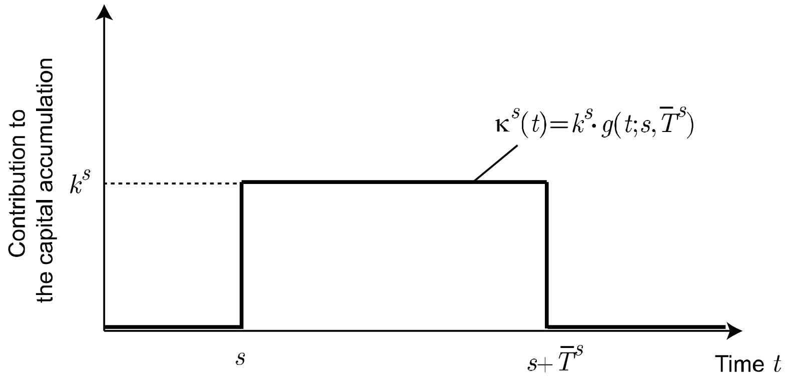


Figure 7.1. Increment of capital due to investment into infrastructure at time s.

$$k^s = \frac{i(s)}{q^s} = \frac{ey(s)}{q^s} \tag{7.16}$$

Finally, the amount $k(t)$ of capital at any given time t is represented as:

$$k(t) = \int_0^t \kappa^s(t)ds + k_0(t) = \int_0^t k^s g(t;s,\bar{T}^s)ds + k_0(t) \tag{7.17}$$

where $k_0(t)$ represents the amount of the initial capital remaining at time t.

The objective function of the dynamic optimization problem here is the social welfare function, which is defined as:

$$W = \int_0^\infty U(c(t))e^{-\rho t}dt \tag{7.18}$$

where $U(c(t))$ is the utility function of the representative household in the economy and ρ is the discount rate for pure-time preference. Note that the consumption $c(t) = (1-e)y(t)$ in the utility function is implicitly a function of the set of the decision variable $\{a_s\}_{s=0}^t$ until time t. The dynamic optimization problem for the design and maintenance policy on infrastructure is thus formulated as:

$$\max_{\{a^s\}_{s=0}^\infty} \quad W = \int_0^\infty U(c(t))e^{-\rho t}dt \tag{7.19}$$

subject to:

$$c(t) = (1-e)f(k(t)) \tag{7.20}$$

$$k^s = \frac{ef(k(s))}{q(a^s)} \tag*{(7.16)'}$$

$$k(t) = \int_0^t k^s f(t; s, \overline{T}(a^s))ds + k_0(t) \qquad (7.17)'$$

with the initial conditions: the initial amount $k_0(0) = k_0$ of the infrastructure capital, and the expected service life time of the infrastructure initially available. The set of decision variables that should be optimized by societal decision-makers is the set of design and maintenance policies $\{a_s\}_{s=0}^{\infty}$.

7.4.2. Steady state analysis

First, the steady state is analyzed where the amount of capital is constant, i.e. the state of no economic growth, which is characterized by $\dot{k} = 0, \ (k > 0)$. For this state, the increment of the capital due to the investment in the infrastructure capital should exactly compensate the decrease of capital due to deterioration. This is represented as, see also Figure 7.2:

$$\frac{ef(k^*)}{q^*} = \frac{k^*}{\overline{T}^*} \qquad (7.21)$$

where the superscript "*" on the symbols signifies that the quantities are the quantities at the steady state. The left hand side comes from Equation (7.16), and the right hand side is obtained from the assumptions made on the deterioration of the infrastructure.

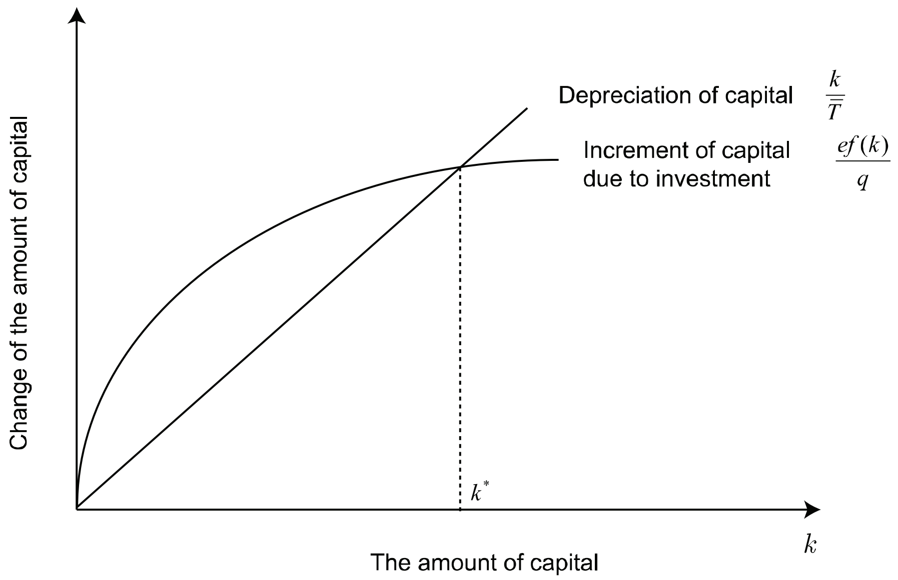


Figure 7.2. Steady state where the increment of the capital exactly compensates the depreciation.

Reformulating Equation (7.21),

$$ef(k^*) = \frac{q^*}{\overline{T}^*} k^* \tag{7.22}$$

From the assumed properties of the production function ($df/dk > 0, d^2 f/dk^2 < 0$), k^* is maximized when q^*/T^* is minimized. Since the highest production level leads to the highest consumption level for a given saving rate, the optimal policy at the steady state is the policy a^* that minimizes q^*/T^*.

Note, however, that this steady state does not necessarily correspond to the optimal state in the sense that the consumption is maximized. This is because of the assumption that the saving rate e is exogenously given. Since the saving rate e corresponds one-to-one with the amount k^* of the capital at the steady state through the relation given by Equation (7.22), the optimal saving rate that maximizes consumption at the steady state is characterized by k^* as:

$$\max_{k^*} \quad c^* = (1-e) f(k^*) = f(k^*) - \frac{q^*}{\overline{T}^*} k^* \tag{7.23}$$

Thus, the optimal amount k_{opt} of the capital that maximizes the consumption at the steady state is obtained as the amount that satisfies the following equation:

$$f'(k_{opt}) = \frac{q^*}{\overline{T}^*} \tag{7.24}$$

where f' represents the derivative with respect to k. This corresponds to *the golden rule of accumulation* for the Solow-Swan model, see Phelps (1961).

The optimization principle obtained from Equation (7.22) for the design and maintenance policy on infrastructure shows that the sum of the initial cost and maintenance cost of infrastructure divided by the service life time, i.e. average cost per unit time, should be minimized. This is intuitively appealing. In order to investigate this principle further, an illustrative relationship between the expected service life time $\overline{T}$ and the average cost $\tilde{q}(\overline{T})/\overline{T}$ [21] per unit time is shown in Figure 7.3. For a smaller expected service life time, the average cost per unit time is higher. This is because the overall cost divided by a shorter expected life time is disproportionately large. On the other hand, infrastructure with a very long expected service life may be very costly due to technical reasons and/or may not even be feasible for other reasons, e.g. societal obsolescence. This is why in Figure 7.3 the average cost per unit time increases sharply

[21] Both the cost $q(a)$ and the expected service life $\overline{T}(a)$ are functions of the decision variable a. However, since the expected service life corresponds to the decision variable one-to-one, the cost is considered as a function of the expected service life time, i.e. there exists such a function as $\tilde{q}(\overline{T}) = q(a)$. The variables without the superscript s represent the variables at any arbitrary time.

for a very long expected life time. Between these two extremes, the average cost per unit time moderately decreases as the expected service life time increases.

One of the most relevant differences between the optimization principle obtained here and the life-cycle cost optimization commonly utilized in engineering decision-making is that the principle obtained here does not involve failure cost terms, which in the life-cycle cost optimization play an important role. The explanation for this is: first, in the economic model considered here (also in most economic models), the loss of infrastructure due to failure is considered in the deterioration terms in the equation of the capital accumulation (see Equation (6.2) or Equation (7.17), although in Equation (7.17) the deterioration term is implicit); second, the reduction of the economic output associated with the loss of capital is considered through the production function by substituting a smaller amount of capital due to the loss of capital. Namely, possible consequences due to the loss of infrastructure are already taken into account. However, note that although the objective function in the optimization principle obtained above and the objective function in the life-cycle cost optimization principle are not the same[22], this is not contradictory. In fact, the contexts in which these two principles are assumed to be applied are different; the life-cycle cost optimization principle is suitable for marginal decision analysis, and the principle obtained above is suitable for non-marginal decision analysis.

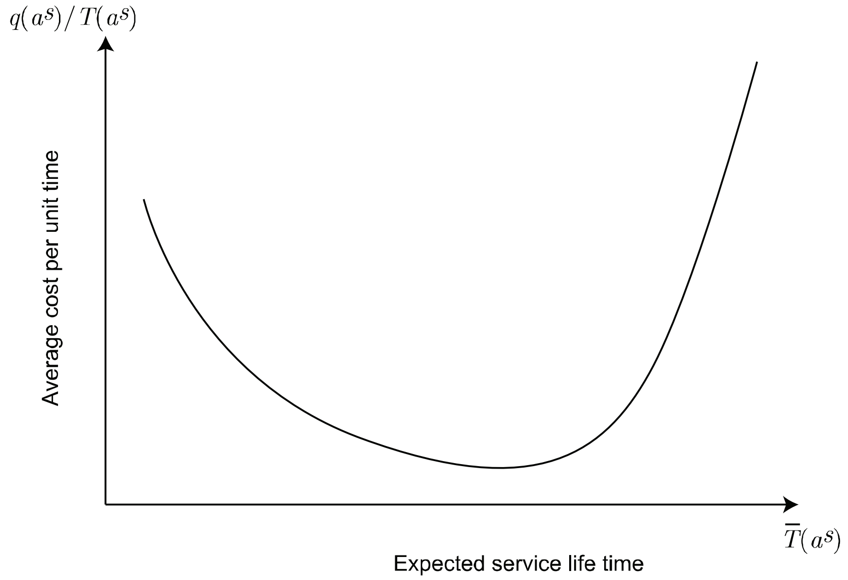


Figure 7.3. Relationship between expected service life time $\overline{T}(a^s)$ and average cost per unit time $q(a^s)/\overline{T}(a^s)$ at any given time s.

22 Since the objective function derived in this section is the objective function at the steady state, the objective function in the life-cycle cost optimization should assume zero discount rate for economic growth, $\delta = \dot{c}/c = 0$, in order for the comparison to be meaningful.

7.4.3. Transition state analysis

In the previous section, the optimal design and maintenance policy on infrastructure is considered at the steady state. However, the application of this optimal policy in a transition state (i.e. economy is under development, $\dot{y} > 0$) may not be optimal in the sense that the overall social welfare for all relevant generations defined in Equation (7.19) may not be maximized. In order to answer this question, the dynamic optimization problem defined by Equations (7.16) to (7.20) in Section 7.4.1 is considered.

The dynamic optimization problem is solved numerically here, since it appears difficult to apply commonly available algorithms for the analytical solution of dynamic optimization problems, e.g. the variation methods or the maximum principle (see Chiang (1999)). For this reason, the parameters required for solving the problem are postulated as shown in Table 7.1. The functional forms of the utility function, production function, cost function, and the function that represents the deterioration of the initial capital are also shown. Note that the values of the parameters are assumed only for performing the numerical calculation, thus the values themselves are not relevant.

In the dynamic optimization problem, the equations are discretized on a multi-annual basis. It is assumed that the design and maintenance policy can be changed every 10 years for the first 100 years, and the same policy is taken after 100 years. The reasons for this assumption are 1) that a more frequent change of the policy, e.g. every year, may be feasible but not realistic in practice, 2) that a more frequent change of the policy increases the number of the variables to be optimized in the optimization problem, which makes the optimization cumbersome, and 3) the optimization of the policies in the distant future is computationally more demanding because the contribution of the change of the policies in the distant future to the objective function is much less due to discounting. Thus, the optimization variables are effectively eleven expected service life times for the infrastructure that is constructed in each respective period.

It should be mentioned that this optimization problem is reduced to identify the optimal balance between the quality of infrastructure (measured in terms of the expected service life time $\overline{T}^t$) and the quantity of infrastructure (the amount k^t of new construction), since the size of budget available is limited to $ef(k(t))$; the constraint $ef(k(t)) = q(\overline{T}^t)k^t$ must be satisfied each time.

Note that in this assumption the optimal policy at the steady state corresponds to $\overline{T}^* = 100$ years, because the annual average cost, $q(\overline{T})/\overline{T} = 1/\overline{T} + \overline{T}/100^2$ (see Table 7.1), is minimized at $\overline{T}^* = 100$.

Table 7.1. Functional forms and parameters postulated in the optimization problem.

Utility function	$u(c(t)) = \ln c(t)$
Discount rate for pure-time preference	$\rho = 0.02$ [1/year]
Production function	$y(k(t)) = Ak(t)^{\alpha}$, $A = 10$, $\alpha = 0.2$
Design and maintenance cost	$q(\overline{T}) = 1 + (\overline{T}/100)^2$
Amount of initial capital	$k_0(0) = 10$
Deterioration of initial capital	$k_0(t) = k_0(0)\left(1 - t/30\right)$, ($0 < t \le 30$)
Saving rate	$e = 0.2$

The optimized[23] service life time in each period and corresponding economic growth path (denoted by a dynamically optimized policy) are shown in Figure 7.4. It is seen that the optimal policy, which maximizes the social welfare, is to choose a shorter expected service life time at an earlier stage of the economy and then to switch to a longer expected service life time later. It should be mentioned that the optimized expected service life time after 100 years is not $\overline{T}^* = 100$ years, which could lead to the highest steady state. This is because the contribution of the utility of future generations to the social welfare is small so that higher consumption of earlier generations is more important to reach a higher social welfare. For comparison purposes, two other economic paths for different policies are calculated; the economic growth path in the case where the expected service life time is fixed at 100 years for all periods (denoted by "$\overline{T}^s = 100$ years, fixed" in the figure) and the economic growth path in the case where the expected service life time is incrementally increased from 40 years to 100 years (denoted by "step" in the figure). It is clearly seen that with the "fixed" policy the economy suffers lower economic output in earlier years, although in the long run the economic output can reach the highest value. Under the "step" policy the economy can grow as fast as the economy under the dynamically optimized policy in the earlier years. However, the economic growth becomes slower in later years because of higher design and maintenance costs for the infrastructure with the longer expected service life time. The calculated social welfare is highest in the case of the dynamically optimized policy, the second highest in the case of the "step" policy and the lowest in the case of the "fixed" policy. That is, from the viewpoint of the social welfare maximization the application of the optimal policy at the steady state in the transition state is suboptimal.

23 Note that this dynamic optimization result is only approximate. The reason is that the time horizon is truncated at a finite time (in this calculation, 200 years). This might be problematic because with this truncation a sound strategy for a future generation living just before the 200 years time limit is to construct infrastructure with a very short expected service life time to increase the economic output for a short term, without considering severe deterioration of the infrastructure which could occur after 200 years. However, the main conclusion in this section is valid, i.e. that the application of the optimal policy at the steady state in the transition state is suboptimal, because the social welfare that corresponds to the "dynamically optimized" policy is calculated using the obtained expected service life times, and this is larger than the social welfare that corresponds to the policy whereby the expected service life time of $\overline{T}^s = 100$ [year] is taken for the whole period of time.

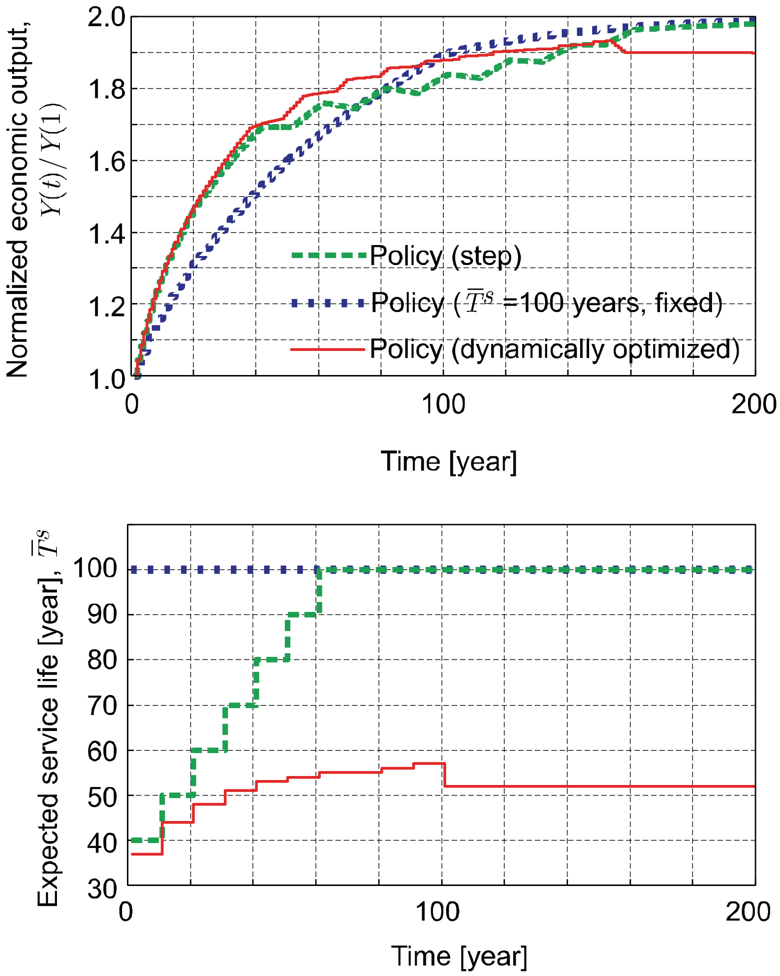


Figure 7.4. Economic growth paths (top) and expected service life times (bottom).

7.4.4. Discussion and conclusion

In this chapter, first the derivation of the life-cycle cost optimization concept from the more general principle, i.e. the social welfare maximization concept, is introduced. Then, the assumptions and limitations of the life-cycle cost optimization concept are pointed out. Thereafter, an optimization principle for the design and maintenance policy on infrastructure is presented in a macroeconomic context based on the economic growth theory. The optimization principle derived here can only be applied at steady states of the economy. Finally, the dynamic optimization problem is considered where the policy on the target expected service life time of infrastructure is optimized. With the assumptions made in Sections 7.4.1 and 7.4.3, it is shown that a better policy in the sense that it leads to a higher social welfare is to choose a shorter expected service life time when the economic output level is relatively low, and then to shift to a longer expected service life time when the economy grows enough to afford the high cost but highly reliable infrastructure; the optimal policy on the reliability of civil infrastructure at each time depends on the current economic output level.

The economic model considered in this chapter is simplistic. There are many possibilities to extend the economic model; including private capitals and other types of capitals in the economy; considering technological development, or employing endogenous economic models; modeling the deterioration of infrastructure in different probabilistic ways; using a more realistic budget framework especially for maintenance costs. These extensions are addressed as future research tasks.

8. Conclusions and outlook

8.1. Conclusions

In the present thesis, the issues of sustainable decision-making in civil engineering, especially design and maintenance strategies for structures, are addressed. These issues are examined from two perspectives, i.e. marginal decision analysis and non-marginal decision analysis. Within the context of marginal decision analysis, sustainable decision problems can be formulated as constrained optimization problems. Therein, the objective function is the expected discounted life-cycle cost associated with the projects concerned, and the constraints correspond to the societal preferences with respect to different aspects of sustainability, which are usually represented in terms of acceptance criteria. In the context of non-marginal decision analysis, sustainable policy-making on the design and maintenance of civil infrastructure can be discussed within a macroeconomic framework. Focusing on individual issues in marginal decision analysis as well as non-marginal decision analysis, the present thesis proposes methods useful to formulate and solve the constrained optimization problems, and a methodological approach for implementing the structural performance of infrastructure in terms of reliability in the economic growth theoretical framework.

In the context of marginal decision analysis, the main constituents of the objective function are: probability of failure; discount factors; cost terms such as initial cost, maintenance cost, cost of failure and indirect cost beyond the direct cost associated with structural failure. In Chapters 2, 4 and 5, these constituents are individually addressed and investigated from a sustainability perspective. On the other hand, in Chapter 3 a computational method is presented for formulating and solving the constrained optimization problems integrating these constituents.

Chapter 2 considers the treatment of aleatory and epistemic uncertainties in the probabilistic assessments of events. The motivation for this chapter is to emphasize the importance of the consistent treatment of these types of uncertainty in probabilistic assessment in general, and in the probabilistic assessment of extreme events during longer periods which usually requires the extrapolation of knowledge concerning e.g. the probabilistic characteristics of events during shorter periods in particular. Its importance is emphasized by introducing three practical examples in which the uncertainties are integrated in an inconsistent manner, and then by showing that such inconsistent treatments can lead to highly biased estimates of the probabilistic characteristics of extreme events. The principle presented in this chapter serves as a philosophical basis for the treatment of uncertainties in sustainable decision analysis for civil infrastructure.

Chapter 3 presents a platform on which the constituents of the constrained optimization problems are fully represented, and the calculations required for optimizations can be performed in a generic manner. For this, the Bayesian probabilistic networks and influence diagrams are adopted as the probabilistic representation platforms. Such representation directly allows for calculating any conditional probabilities and conditional expected values of variables of interest by use of the generic algorithms developed for such calculations as a function of any given decision alternative. Furthermore, by linking the networks/diagrams to the generic algorithms available for solving constrained optimization problems, the constrained optimization problems of interest can be solved quasi automatically once the networks/diagrams corresponding to the problems are established. The decision-makers can thus focus on the development of such networks/diagrams, which is highly useful in practical applications. Another practical advantage of employing the Bayesian probabilistic networks or influence diagrams is that they can be facilitated as communication tools among experts as well as between experts and non-experts. The use of the networks/diagrams as a communication tool is especially useful in decision analysis for civil infrastructure, since civil infrastructure is in general a complex system composed of components at different levels and therefore the modelling of these components and their possible consequences require the collaboration of experts from different disciplines.

Chapter 4 addresses the issue of discounting. The motivation of this chapter is to reconsider the formulation of the life-cycle cost optimization problem from an intergenerational-equity perspective. The focus is on discounting for pure-time preference. Because discounting for pure-time preference reflects the myopic nature of individuals, the application of discounting for pure-time preference can be logically justified only within individual generations. However, often in life-cycle optimization problems, discounting for pure-time preference has been applied without considering the finite duration of the generations (or as if one generation lives for ever). In this chapter, based on the consideration similar but independent from the generation-adjusted discounting concept proposed by Bayer and Cansier (1999), a formula is proposed to calculate an equivalent discount rate. The equivalent discount rate is the rate which, if applied to a decision problem with the classical perspective in which one generation that is assumed to have an infinite lifetime, yields the same total expected utility as when the decision problem is analyzed in accordance with the consistent consideration of discounting over generations. The use of the formula can thus avoid tedious calculations required for the assessment of the discounted life-cycle costs if the generation-adjusted discounting concept was applied in a straightforward manner. Furthermore, from the formula it directly turns out that the classical perspective tends to put more burden on future generations by applying a higher discount rate than the rate that is logically consistent with the implication of discounting for pure-time preference and the finite duration of individual generations.

Chapter 5 reformulates the life-cycle cost optimization concept from a budget allocation perspective. The background of this reformulation is that the life-cycle cost optimization concept implicitly assumes that the necessary budget is available whenever it is needed, which is not realistic in practice. Since the failure to acquire the necessary budget on time incurs additional indirect costs due to the delay of actions, an explicit consideration of such costs is required to formulate the objective function in the life-cycle cost optimization concept. Thereby, because the probability of occurrence of the non-availability of the budget is a function of the size of the budget allocated to the project, the size of the budget to be allocated should be one of the decision variables in the optimization problem. This shift of perspective is especially useful for societal decision-makers who have to decide how to optimally allocate the limited budget to different projects.

In Chapters 3, 4 and 5, it is implicitly assumed that the decision analyses are marginal. That is, these analyses are only valid if the consequences of the decisions are reasonably assumed not to influence long-term economic growth. However, there have been some cases where the life-cycle cost optimization concept, which is the typical concept for the marginal decision analysis, seems to have been applied beyond its limitation[24]. Therefore, in order to clarify its underlying assumption and limitation, the derivation of the life-cycle cost optimization concept from a broader decision principle is introduced in Section 7.2.

In contrast to Chapters 3, 4 and 5, the non-marginal decision analysis is addressed within the framework of economic growth theory in Chapters 6 and 7. The original contribution here is that a methodological approach is proposed in regard to how the reliability of structures defined in terms of limit state functions, and initial cost and maintenance cost can be integrated into the framework. The direct benefit of the proposed approach is that sustainable design and maintenance policies on infrastructure can be discussed in the macroeconomic context, thereby enabling one to assess the effects of the policies on long-term economic development. In Chapter 7, an optimization principle for the design and maintenance policy on infrastructure is derived in a macroeconomic context, which can apply at steady states of the economy. Its objective function takes a different form from that of the objective function in the life-cycle cost optimization concept. However, this is not contradictory; the optimization principle derived in Chapter 7 should be applied for non-marginal decision analysis, and the life-cycle cost optimization concept should be applied for marginal decision analysis. It is also shown that the presented methodological approach enables one to identify the optimal reliability (represented by the expected service life time) of infrastructure as a function of the economic states, which could be difficult within the marginal decision analysis framework.

24 In reality, it is often very difficult to check the marginality. Hence, the marginality is often merely an assumption for the decision analysis. Still in such cases, it is important that the interpretation and application of the analysis results should be consistent with the assumption.

The methods proposed for the marginal decision analysis are directly useful in present practical decision situations where due consideration of sustainability is required. On the other hand, the proposed approach for non-marginal decision analysis will serve as the first step for a further development of the general framework for sustainable policy-making on civil infrastructure.

8.2. Scientific achievements and limitations

The scientific achievements of the present thesis work are: (1) introduction of the concept of marginality of engineering decision-making; (2) adaptation of the classical life-cycle cost optimization concept to sustainable decision-making in the context of a marginal decision analysis; (3) development of a non-marginal engineering decision analysis framework.

However, these scientific achievements are limited. Regarding (1), the marginality of decisions introduced in the present thesis is difficult to assess in practice; strictly speaking, any engineering decisions can affect economic growth. Hence, by definition any engineering decision can be non-marginal. Thus, marginal decision analysis can be considered as an approximation. Therefore, the concept of marginality should be used in practice to check if the assumption of the marginality is reasonable in given decision situations; only when the assumption is reasonable can the marginal decision analysis be performed. Otherwise, a non-marginal decision analysis should be undertaken. Concerning (2), it is assumed that the objective function in the decision problems can be represented by or otherwise converted to a monetary term. However, it is not clear if the objective function can be fully described by the monetary term. Even if it were possible, it is still not obvious how the values of different actions and consequences are objectively quantified in the monetary term. The present thesis does not provide any justification for this assumption and ways to quantify them. Furthermore, the boundary conditions in the decision-making process are assumed to be given. However, since the choice of the boundary conditions affects the optimization of the decision-making, these boundary conditions should be carefully assessed and chosen. The way in which they are assessed and chosen is addressed as a future task. Finally, in regard to (3), the proposed framework is under development in the sense that the role of civil infrastructure in the economy is considered only in terms of productivity; civil infrastructure plays other important roles in society such as amenity for leisure and safety measures to mitigate natural hazards. At the same time, the operation of civil infrastructure impacts on the quality of the environment. These aspects are not considered in the proposed framework and the consideration of these aspects is addressed as an additional future task.

8.3. Outlook

8.3.1. Assessment of the boundary conditions in marginal decision analysis

Throughout the present thesis, the research focus is mainly on theoretical aspects of the issues mentioned in the introduction. That is, whereas the method for solving the constrained optimization problems given the constraints is presented (Chapter 3), the constraints are assumed to be given in terms of the acceptance criteria related to the aspects of sustainability, e.g. human safety and environmental impact; whereas the formula for obtaining sustainable equivalent discount rates is presented (Chapter 4), it is not investigated which values should be assumed for the discount rates for economic growth and pure-time preference. In practice, these may be equally or even more relevant in decision-making.

Concerning the acceptance criteria for human safety in the context of engineering project appraisal, a number of approaches have been proposed and utilized in practice. One common approach in practice is the use of the Farmer diagram, often called the F-N curve[25]. In this approach, the F-N curve concerning a considered project is compared with the criterion F-N curve, which is usually provided by regulatory authorities; the considered project is acceptable if the F-N curve concerning the project is below the criterion F-N curve. However, several inconsistencies in the use of F-N curves for project appraisal concerning human safety have been pointed out. Among others, it is possible that a project that associates higher expected fatality due to possible accidents in a given time period is accepted, whereas another project that associates lower expected fatality is rejected. This is because the F-N curve-based project appraisal essentially concentrates on one extreme feature of the distribution of fatalities due to different possible accidents disregarding the overall characteristics of the distribution of the fatalities, see Evans and Verlander (1997). In Evans and Verlander (1997), it is also shown that the F-N curve-based project appraisal fails to pass a logical test for a prescriptive criterion. Recently, a promising approach has been developed based on the life quality index (LQI) proposed by Nathwani et al. (1997). The LQI is a social indicator that is composed of the gross domestic product per capita, the life expectancy and the fraction of lifetime spent in working for a living. In this approach, the LQI is considered to represent the indifference between the increase/decrease of life expectancy and the decrease/increase of consumption per capita. Thus, the willingness to pay for life-saving measures can be derived from this index, see e.g. Skjong and Ronold (1998) and Rackwitz (2003). Further development of this approach is necessary and is on-going, see e.g. Ditlevsen (2004), Kübler and Faber (2005), Pandey et al. (2006), and Ditlevsen and Friis-Hansen (2008).

25 An F-N curve represents, for different n, the mean absolute frequency $F(n)$ of the accidents in a reference time period which associate n or more fatalities in a considered project. Normally, in the diagram the horizontal axis corresponds to the number n of fatalities and the vertical axis corresponds to the mean absolute frequency $F(n)$.

The acceptance criteria for environmental impacts, e.g. targets in pollution control, use of non-renewable resources, and recycling of partly-renewable materials, have been intensively discussed in environmental sciences and economics, see Perman et al. (2003) for an overview. Among others, what is relevant to the design and maintenance on civil infrastructure are: the recycling of the construction materials, e.g. cements, aggregates in concrete and steel; emissions of carbon dioxide in the construction and operation of the infrastructure. These should be addressed in the context of life-cycle optimization problems as well as the life-cycle cost. Thereby, a controversial issue arises; how to identify the decision alternatives among a set of the Pareto-optimal solutions in the optimization problem if formulated as a multi-objective optimization problem, or otherwise which attribute should be taken as the (scalar) objective function and which should be considered as constraints in constraint optimization problems. This should be addressed as a challenging research task.

Different choices of the values of discount rates often lead to different conclusions. One of the examples can be seen in the debate between Nordhaus (2007) and Stern and Taylor (2007) on the necessity for urgent countermeasures to global warming. Nordhaus (2007) criticizes the choice of the values of the discount rates in the Stern Report (2006) (the discount rate for pure-time preference: $\rho = 0.001/yr$, the discount rate for consumption growth: $\delta = 0.013/yr$, and the elasticity of the marginal utility of consumption: $\eta = 1$) by arguing that the resulting real return rate, $r = \eta\delta + \rho = 0.014/yr$, is far smaller than the real return rate observed in the capital market. For this criticism, Stern and Taylor (2007) justify their choice by claiming that 1) the discount rate for pure-time preference should be significantly smaller when the consequences of the decision affect both current and future generations[26], and 2) the capital market is imperfect in the sense that those who do not or cannot participate in the market (i.e. the young, the poor and the future generations) have little or no influence on current market behavior. The underlying issue in this debate is the choice of the perspective to be followed in societal decision-making; normative or descriptive. The normative perspective seems reasonable for societal decision-making. However, with this approach it is difficult to directly obtain the value of the discount rate for pure-time preferences without relying on statistics that may be affected by the capital market. Furthermore, the justification to assume a positive discount rate for pure-time preference is a controversial issue, see e.g. Price (1993); the choice of the value of the discount rate can be subjective.

Decision-making in civil engineering often encounters the similar situation in which the consequences of the decisions affect current and future generations, e.g. construction of nuclear power plants and nationwide infrastructure projects. In these decision situations, whereas it is difficult to choose the commonly agreed discount rates, it is important that the process leading to the choice is clear and transparent, and

26 For this, the discussions in Chapter 4 may provide a rationale.

consistent with the state of the art of the philosophical discussions on discounting; continuous literature reviews and dissemination of the review results to decision-makers is important.

8.3.2. Further development of non-marginal decision framework

The proposed macroeconomic decision framework is a promising platform, on which the effects of policies in different sectors concerning the long-term development of the economy can be examined individually as well as jointly. This is possible because the framework defines the generic format for the component models based on the economic growth theory: the social welfare function, the production function, and the equation of motion that governs the changes in the amount of capitals. Thereby, it is readily possible that the quality of different types of capitals can be modeled based on the corresponding engineering knowledge through the limit state representation, the reliability in the generalized sense is calculated using the structural reliability theory and it is implemented into economic models in similar manners as illustrated for civil infrastructure in Chapters 6 and 7. Here, it is mentioned that the policy making concerning the global warming is one of such possible applications of the proposed framework.

However, as is pointed out in Chapter 7 the equation of motion typically employed in the economic growth theory is too simplistic for some types of capitals, including infrastructure capital. On the other hand, realistic equations of motion as presented in Chapter 7 significantly complicate dynamic optimization problems, so that the application of the maximum principle is not feasible. Therefore, either a simplification of the model that does not lose the relevant characteristics or an efficient algorithm for solving the complicated dynamic optimization problem is required.

At the same time, more sophisticated, realistic economic models need to be developed to fully capture the interaction between infrastructure capital and other capital, and the socio-economic roles of infrastructure. It is also required to extend the framework in order to enable one to take in account environmental aspects such as exploitation, recycling and reuse of non-renewable resources and protection of the biodiversity. The goal in the development of a non-marginal decision framework is to obtain a framework that can identify the sustainable policies on infrastructure taking into account all relevant aspects of sustainability including economic growth, the socio-economic role and environmental issues jointly.

References

Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies*, 29, 155-173.

ASCE7-98. (2000). Minimum design loads for buildings and other structures, Revision of ANSI/ASCE.

ASCE. (2005). Report card for America's infrastructure.

Aschauer, D. A. (1989). Is Public Expenditure Productive? *Journal of Monetary Economics*, 23, 177-200.

Ayres, R. U., van den Bergh, J. C. J. M., and Gowdy, J. M. (1998). Viewpoint: Weak versus strong sustainability. Tinbergen Institute Discussion papers.

Baker, J. W., Schubert, M., and Faber, M. H. (2007). On the assessment of robustness. *Structural Safety*, In Press, Corrected Proof.

Bangso, O., Flores, M. J., and Jensen, F. V. (2003). Plug & Play OOBNs. *Lecture Notes in Artificial Intelligence*, 3040, 457-467.

Bangso, O., and Olesen, K. G. (2003). Applying Object Oriented Bayesian Networks to Large (Medical) Decision Support Systems. *Proceedings of 8th Scandinavian Conference on Artificial Intelligence, SCAI'03*, Bergen, Norway.

Barro, R. J., and Sala-i-Martin, X. (2004). *Economic Growth Second Edition*, The MIT Press, Cambridge, Massachusetts.

Bayer, S. (2003). Generation-adjusted discounting in long-term decision-making. *International Journal of sustainable development*, 6(1), 133-149.

Bayer, S., and Cansier, D. (1999). Intergenerational Discounting: A New Approach. *Journal of International Planning Literature*, 14(3), 301-325.

Bayraktarli, Y., and Faber, M. H. (2007). Bayesian network approach for managing earthquake risks. *International Forum on Engineering Decision Making, IFED3*, Shoal Bay, Australia.

Benjamin, J. R., and Cornell, C. A. (1970). *Probability, Statistics and Decision for Civil Engineers*, McGraw-Hill, New York.

Bobbio, A., Ciancamerla, E., Franceschinis, G., Gaeta, R., Minichino, M., and Portinale, L. (2003). Sequential application of heterogeneous models for the safety analysis of a control system: a case study. *Reliability Engineering & System Safety*, 81(3), 269-280.

Bobbio, A., Portinale, L., Minichino, M., and Ciancamerla, E. (2001). Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. *Reliability Engineering and System Safety*, 71(3), 249-260.

Brundtland, G. H. (1987). *Our Common Future*, Oxford University Press.

Bulow, J. I., and Summers, L. H. (1984). The Taxation of Risky Assets. *The Journal of Political Economy*, 92(1), 20-39.

Canning, D. (1998). A Database of World Infrastructure Stocks, 1950-95. World Bank Policy Research Working Paper, 1929, World Bank.

Canning, D. (1999). Infrastructure's contribution to aggregate output. World Bank Policy Research Working Paper, World Bank.

Canning, D., and Bennathan, E. (2000). The Social Rate of Return on Infrastructure Investments. World Bank Policy Research Working Paper, 2390, World Bank.

Cass, D. (1965). Optimum Growth in an Aggregative Models of Capital Accumulation. *Review of Economic Studies*, 32, 233-240.

Clark, J. S., and Gelfand, A. E. (2006). *Hierarchical Modelling for the Environmental Sciences*, Oxford University Press, New York.

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

Coles, S., Pericchi, L. R., and Sisson, S. (2003). A fully probabilistic approach to extreme rainfall modeling. *Journal of Hydrology*, 273, 35-50.

Cornell, A. (1969). A Probability-Based Structural Code. *ACI Journal*, 974-985.

Dasgupta, P. S., and Heal, G. M. (1974). The optimal depletion of exhaustible resources. *Review of Economic Studies*, 41, 3-28.

De Groot, H. (1970). *Optimal Statistical Decisions*, John Wiley&Sons, Inc.

Der Kiureghian, A., and Ditlevsen, O. (2007). Aleatory or epistemic? Does it matter? *Special workshop on risk acceptance and risk communication*, Stanford University, California, USA.

Der Kiureghian, A., and Ditlevsen, O. (2008). Aleatory or epistemic? Does it matter? *Structural Safety*, available online.

Der Kiureghian, A., Haukaas, T., and Fujimura, K. (2006). Structural reliability software at the University of California, Berkeley. *Structural Safety*, 28(1-2), 44-67.

Der Kiureghian, A., and Moghtaderi-Zadeh, M. (1982). An integrated approach to the reliability of engineering systems. *Nuclear Engineering and Design*, 71(3), 349-354.

Der Kiureghian, A., and Song, J. (2008). Multi-scale reliability analysis and updating of complex systems by use of linear programming. *Reliability Engineering & System Safety*, 93(2), 288-297.

Ditlevsen, O. (2004). Life quality index revisited. *Structural Safety*, 26(4), 443-451.

Ditlevsen, O., and Bjerager, P. (1986). Methods of structural systems reliability. *Structural Safety*, 3(3-4), 195-229.

Ditlevsen, O., and Friis-Hansen, P. (2008). Cost and benefit including value of life, health and environmental damage measured in time units. *Structural Safety*, In Press, Corrected Proof.

Ditlevsen, O., and Madsen, H. O. (2005). *Structural Reliability Methods*.

Duffy-Deno, K. T., and Eberts, R. W. (1991). Public infrastructure and regional economic development: A simultaneous equations approach. *Journal of Urban Economics*, 30, 329-343.

Duggal, V. G., Saltzman, C., and Klein, L. R. (1999). Infrastructure and productivity: a nonlinear approach. *Journal of Econometrics*, 92, 47-74.

DuraCrete. (2000). Statistical Quantification of the Variables in the Limit State Functions. Final DuraCrete report, European Union.

Easterly, W., and Rebelo, S. (1993). Fiscal policy and economic growth. *Journal of Monetary Economics*, 32(3), 417-458.

Engle, R. F., and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251-276.

EUROCONSTRUCT. (2007). European construction market trends to 2010, Country report.

Evans, A. W., and Verlander, N. Q. (1997). What Is Wrong with Criterion FN-Lines for Judging the Tolerability of Risk? *Risk Analysis*, 17(2), 157-168.

Faber, M. H. (2003). Uncertainty Modeling and Probabilities in Engineering Decision Analysis. *Proceedings of the 22nd International Conference on Offshore Mechanics and Arctic Engineering, OMAE2003*, Cancun, Mexico.

Faber, M. H., Bayraktarli, Y., and Nishijima, K. (2007a). Recent Developments in the Management of Risks Due to Large Scale Natural Hazards. *XVI Congreso Nacional Ingenieria Sismica*, Ixtapa-Zihuatanejo, Mexico.

Faber, M. H., Engelund, S., Sørensen, J. D., and Bloch, A. (2000). Simplified and Generic Risk Based Inspection Planning. *Proceedings OMAE2000, 19th Conference on Offshore Mechanics and Arctic Engineering*, New Orleans, Louisiana, USA, [OMAE2000/S&R6143].

Faber, M. H., Maes, M., Baker, J., Vrouwenvelder, T., and Takada, T. (2007b). Principles of risk assessment of engineered systems. ICASP 10, Tokyo, Japan.

Faber, M. H., and Maes, M. A. (2003). Modeling of Risk Perception in Engineering Decision Analysis. *Proceedings of 11th IFIP WG7.5 Working Conference on Reliability and Optimization of Structural Systems*, 113-122.

Faber, M. H., and Maes, M. A. (2005). Epistemic Uncertainties in Decision Making. *Proceedings of the 24th International Conference on Offshore Mechanics and Arctic Engineering, OMAE2005*, Halkidiki, Greece.

Faber, M. H., and Nishijima, K. (2004). Aspects of Sustainability in Engineering Decision Analysis. *Proceedings 9th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability*, Albuquerque, New Mexico, USA.

Faber, M. H., and Rackwitz, R. (2004). Sustainable decision making in civil engineering. *Structural Engineering International*, 14(3), 237-242.

Faber, M. H., Straub, D., and Maes, M. A. (2005). A Computational Framework for Risk Assessment of RC structures Using Indicators. *Computer-Aided Civil and Infrastructure Engineering, accepted for publication*.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*, John Wiley & Sons Inc., New York.

Ferry-Borges, J., and Castanheta, M. (1971). Structural Safety, 2nd Edition. National Civil Engineering Laboratory, Lisboa, Portugal.

Freudenthal, A. M. (1947). The safety of structures. *Transactions of American Society of Civil Engineering*, 112, 125-159.

Freudenthal, A. M. (1954). Safety and Probability of Structural Failure. *Transactions of American Society of Civil Engineering*, 1337-1397.

George, D., and Hawkins, J. (2005). A Hierarchical Bayesian Model of Invariant Pattern Recognition in the Visual Cortex. *Neural Networks*, 3, 1812-1817.

Glomm, G., and Ravikumar, B. (1994). Public investment in infrastructure in a simple growth model. *Journal of Economic Dynamics and Control*, 18(6), 1173-1187.

Gramlich, E. M. (1994). Infrastructure Investment: A Review Essay. *Journal of Economic Literature*, 32(3), 1176-1196.

Greenwood, J., Hercowitz, Z., and Krusell, P. (2000). The role of investment-specific technological change in the business cycle. *European Economic Review*, 44, 91-115.

Guha-Sapir, D., Hargitt, D., and Hoyois, P. (2004). Thirty years of natural disasters 1974-2003: The numbers. Centre for Research on the epidemiology of Disasters.

Guikema, S. D., and Pate-Cornell, M. E. (2002). Component choice for managing risk in engineered systems with generalized risk/cost functions. *Reliability Engineering & System Safety*, 78(3), 227-238.

Hartwick, J. M. (1977). Intergenerational equity and the investing of rents from exhaustible resources. *American Economic Review*, 67, 972-974.

Hartwick, J. M. (1978). Substitution among exhaustible resources and intergenerational equity. *Review of Economic Studies*, 45(347-354).

Hasofer, A. M., and Lind, N. C. (1974). Exact and Invariant Second-Moment Code Format. *Journal of Engineering Mechanics Division, Proceedings of the American Society of Civil Engineers*, 100, 111-121.

Hohenbichler, M., and Rackwitz, R. (1982). First-order concepts in system reliability. *Structural Safety*, 1(3), 177-188.

Holtz-Eakin, D., and Schwartz, A. E. (1995). Infrastructure in a structural model of economic growth. *Regional Science and Urban Economics*, 25(2), 131-151.

JCSS. (2001a). *Probabilistic Assessment of Existing Structures, A publication of the Joint Committee on Structural Safety (JCSS)*, RILEM Publications S.A.R.L The publishing Company of RILEM.

JCSS. (2001b). Probabilistic Model Code. The Joint Committee on Structural Safety.

Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*, Springer, New York.

Johnson, V. E., Graves, T. L., Hamada, M., and Reese, C. S. (2002). A Hierarchical Model for Estimating the Reliability of Complex Systems. *Bayesian Statistics*, 7, 199-214.

JSCE. (2008). わが国におけるインフラの現状と評価 インフラ国勢調査 2007 -体力測定と健康診断-.

Kübler, O., and Faber, M. H. (2005). LQI: On the correlation between life expectanca and the gross domestic product per capita. Proceedings ICOSSAR2005, 9th International Conference on Structural Safety and Reliability, Rome, Italy, 3513-3517.

Kalaitzidakis, P., and Kalyvitis, S. (2004). On the macroeconomic implications of maintenance in public capital. *Journal of Public Economics*, 88, 695-712.

Kjaerulff, U. (1995). dHugin: A computational system for dynamic time-sliced Bayesian networks. International Journal of Forecasting.
Koopmans, T. C. (1965). *On the concept of optimal economic growth*, Amsterdam: North Holland.
Korb, K. B., and Nicholson, A. E. (2004). *Bayesian Artificial Intelligence*, Champman & Hall/CRC.
Krautkraemer, J. A. (1999). On sustainability and intergenerational transfers with a renewable resource. *Land Economics*, 75, 167-184.
Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Series*, Springer Verlag, New York.
Li, F.-F., and Pietro, P. (2005). A Bayesian Hierarchical Model for Learning Natural Scene Categories. *IEEE Conference on Computer Vision and Pattern Recognition* San Diego, USA.
Lin, Y. K. (1967). *Probabilistic theory of structural dynamics*, McGraw-Hill.
Lindley, D. V. (1965). *Introduction to probability and statistics : from a Bayesian viewpoint*, Cambridge University Press, Cambridge.
Lindley, D. V. (1980). *Introduction to Probability & Statistics - from a Bayesian viewpoint, Part.1-Probability*, Cambridge University Press, Cambridge.
Lindley, D. V., and Smith, A. F. M. (1972). Bayes Estimates for the Linear Model. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(1), 1-41.
Lucas, R. E. J. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22, 3-42.
Maes, M. A. (1990). The Influence of Uncertainties on the Selection of Extreme Values of Environmental Loads and Events. *Civil Engineering Systems*, 7(2), 115-124.
Maes, M. A., and Jordaan, I. J. (1985). Extremal Analysis of Loads Using Exchangeability. *Proceedings of the 4th International Conference on Structural Safety and Reliability, ICOSSAR 1985*, Tokyo, Japan, 579-589.
Mayer, M. (1926). *Die Sicherheit der Bauwerke*, Springer, Berlin.
Munich Re. (2005). Annual review: Natural catastrophes 2005. Topics Geo.
Munnell, A. H. (1992). Policy Watch: Infrastructure Investment and Economic Growth. *The Journal of Economic Perspectives*, 6(4), 189-198.
Nathwani, J. S., Lind, N. C., and Pandey, M. D. (1997). *Affordable Safety by Choice: The Life Quality Method*, University of Waterloo, Waterloo.
Newell, R. G., and Pizer, W. A. (2004). Uncertain discount rates in climate policy analysis. *Energy Policy*, 32(4), 519-529.
Nijkamp, P., and Poot, J. (2004). Meta-analysis of the effect of fiscal policies on long-run growth. *European Journal of Political Economy*, 20(1), 91-124.
Nishijima, K., and Faber, M. H. (2006). A Budget Management Approach for Societal Infrastructure Projects. *IABMAS'06, 3rd International Conference on Bridge Maintenance, Safety and Management*, Porto, Portugal.
Nishijima, K., and Faber, M. H. (2007a). Bayesian approach to proof loading of quasi-identical multi-components structural systems. *Civil Engineering and Environmental Systems*, 24(2), 111-121.
Nishijima, K., and Faber, M. H. (2007b). A Bayesian framework for typhoon risk management. *12th International Conference on Wind Engineering, 12ICWE*, Cairns, Australia.
Nishijima, K., and Faber, M. H. (2007c). On Structural Performance vs. Societal Economic Growth. *10th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP10*, Kashiwa, Japan.
Nishijima, K., Maes, M. A., Goyet, J., and Faber, M. H. (2008). Constrained optimization of component reliabilities in complex systems. *Structural Safety*.
Nishijima, K., Straub, D., and Faber, M. H. (2004). Sustainable decisions for Life-Cycle Based Design and Maintenance. *First Forum on Engineering Decision Making, IFED*, Stoos, Switzerland.

Nishijima, K., Straub, D., and Faber, M. H. (2005). The Effect of Changing Decision Makers on the Optimal Service Life Design of Concrete Structures. *Proceedings of the 4th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructures Systems*, Cocoa Beach, Florida, 325-333.

Nishijima, K., Straub, D., and Faber, M. H. (2007). Inter-generational distribution of the life-cycle cost of an engineering facility. *Journal of Reliability of Structures and Materials*, 3(1), 33-46.

Nordhaus, W. (2007). Critical Assumptions in the Stern Review on Climate Change. *Science*, 317, 201-202.

O'Hagan, A., and Oakley, J. E. (2004). Probability is perfect, but we can't elicit it perfectly. *Reliability Engineering & System Safety*, 85, 239-248.

Pandey, M. D., Nathwani, J. S., and Lind, N. C. (2006). The derivation and calibration of the life-quality index (LQI) from economic principles. *Structural Safety*, 28(4), 341-360.

Pate-Cornell, M. E. (1996). Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering & System Safety*, 54, 95-111.

Perman, R., Ma, Y., McGilvray, J., and Common, M. (2003). *Natural Resource and Environmental Economics (3rd edition)*, Pearson Education Limited, Harlow Essex UK.

Pezzey, J. C. V. (1992). Sustainability: an interdisciplinary guide. *Environmental Values*, 1, 321-362.

Pezzey, J. C. V. (1997). Sustainability constraints versus optimality versus intertemporal concern, and axioms versus data. *Land Economics*, 73(4), 448-466.

Pezzey, J. C. V., and Withagen, C. A. (1998). The rise, fall and sustainability of capital-resource economics. *Scandinavian Journal of Economics*, 100(2), 513-527.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1988). *Numerical Recipes in C*, Cambridge University Press.

Price, C. (1993). *Time, Discounting & Values*, Blackwell Publishers, Oxford, United Kingdom.

Rackwitz, R. (2000). Optimization - the basis of code-making and reliability verification. *Structural Safety*, 22(1), 27-60.

Rackwitz, R. (2002). Optimization and Risk Acceptability Based on the Life Quality Index. *Structural Safety*, 24(2-4), 297-332.

Rackwitz, R. (2003). Acceptable Risks and Affordable Risk Control for Technical Facilities and Optimization. *Reliability Engineering and Systems Safety*.

Rackwitz, R., Lentz, A., and Faber, M. (2005). Socio-economically sustainable civil engineering infrastructures by optimization. *Structural Safety*, 27(3), 187-229.

Raiffa, H., and Schlaifer, R. (1961). *Applied Statistical Decision Theory*, Cambridge University Press, Cambridge.

Ramsey, F. (1928). A mathematical theory for saving. *Economic Journal*, 38, 543-559.

Raudenbush, S., and Bryk, A. S. (1986). A Hierarchical Model for Studying School Effects. *Sociology of Education*, 59(1), 1-17.

Rioja, F. K. (2003). Filling potholes: macroeconomic effects of maintenance versus new investments in public infrastructure. *Journal of Public Economics*, 87, 2281-2304.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94, 1002-1037.

Rosenblueth, E., and Mendoza, E. (1971). Reliability Optimization in Isostatic Structures. *Journal of the Engineering Mechanics Division*, 97(6), 1625-1648.

Royset, J. O., Der Kiureghian, A., and Polak, E. (2003). Reliability-based optional design: problem formulations, algorithms and application. *Proceedings of the 11th IFIP WG7.5 Working conference on reliability and optimization of structural systems*, Banff, Canada, 1-12.

Salazar, D., Rocco, C. M., and Galvan, B. J. (2006). Optimization of constrained multiple-objective reliability problems using evolutionary algorithms. *Reliability Engineering & System Safety*, 91(9), 1057-1070.

Skjong, R., and Ronold, K. (1998). Societal Indicators and Risk Acceptance. *17th International Conference on Offshore Mechanics and Arctic Engineering*, Lisbon, Portugal.

Smyth, P. (1997). Belief networks, hidden Markov models, and Markov random fields: A unifying view. *Pattern Recognition Letters*, 18(11-13), 1261-1268.
Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70, 65-94.
Solow, R. M. (1974). Intergenerational equity and exhaustible resources. *Review of Economic Studies*, 41(22-46).
Solow, R. M. (1986). On the intergenerational allocation of natural resources. *Scandinavian Journal of Economics*, 88(1), 141-149.
Song, J., and Kang, W.-H. (2008). System reliability and sensitivity under statistical dependence by matrix-based system reliability method. *Structural Safety*, In Press, Corrected Proof.
Stern, N. (2006). Stern Review: The Economics of Climate Change. HM Treasury.
Stern, N., and Taylor, C. (2007). Climate Change: Risk, Ethics, and the Stern Review. *Science*, 317, 203-204.
Stiglitz, J. E. (1974). Growth with exhaustible natural resources: efficient and optimal growth path. *Review of Economic Studies*, 41, 123-137.
Straub, D. (2004). Generic approaches to risk based inspection planning for steel structures, PhD thesis, ETH Zurich, Zurich.
Straub, D., and Der Kiureghian, A. (2008). Improved seismic fragility modeling from empirical data. *Structural Safety*, 30(4), 320-336.
Straub, D., and Faber, M. H. (2005). Risk based inspection planning for structural systems. *Structural Safety*, 27(4), 335-355.
Straub, D., and Faber, M. H. (2006). Computational Aspects of Risk-Based Inspection Planning. *Computer-Aided Civil and Infrastructure Engineering*, 21(3), 179-192.
Swan, T. W. (1956). Economic growth and capital accumulation. *Economic Record*, 32, 334-361.
Tang, W. H. (1973). Probabilistic Updating of Flaw Information. *Journal oft Testing and Evaluation*, 1, 459-467.
Thoft-Christensen, P., and Sørensen, J. D. (1987). Optimal Strategy for Inspection and Repair of Structural Systems. *Civil Engineering Systems*, 4, 94-100.
Turner, R. K. (1992). Speculations on Weak and Strong Sustainability. Centre for Social and Economic Research on the Global Environment (CSERGE).
USNRC. (1975). Reactor Safety Study - An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants, WASH-1400 (NUREG-75/014). U.S. Nuclear Regulatory Commission.
USNRC. (1990). Severe Accident Risks: An Assessment for Five U.S. Nuclear Power Plants (NUREG-1150). U.S. Nuclear Regulatory Commission.
Valente, S. (2005). Sustainable development, renewable resources and technological progress. *Environmental and Resource Economics*, 30(1), 1573-1502.
Vanmarcke, E. (1983). *Random Fields: Analysis and Synthesis*, MIT Press, Cambridge, Massachusetts.
Vesely, W. E., Goldberg, F. F., Roberts, N. H., and Haasl, D. F. (1981). Fault Tree Handbook (NUREG-0492). U.S. Nuclear Regulatory Commission.
Volovoi, V. (2004). Modeling of system reliability Petri nets with aging tokens. *Reliability Engineering & System Safety*, 84(2), 149-161.
Wen, Y. K., Ellingwood, B. R., Veneziano, D., and Bracci, J. (2003). Uncertainty Modeling in Earthquake Engineering. MAE Center Project FD-2 Report.
Wierzbicky, W. (1936). La sécurité des constructions comme un problème de probabilité. *Annales de l'academie des sciences techniques*, 7, 63-74.
Withagen, C. A. A. M. (1996). Sustainability and investment rules. *Economic Letters*, 53, 1-6.
World Bank. (1994). World development report 1994: infrastructure for development. World Bank.
Zeira, J. (1987). Risk and Capital Accumulation in a Small Open Economy. *The Quarterly Journal of Economics*, 102(2), 265-280.

Curriculum vitae

PERSONAL DETAILS

Name	Kazuyoshi Nishijima
Address	Hofwiesenstr. 200, 8057 Zurich, Switzerland
Telephone	044 633 43 16
E-mail	nishijima@ibk.baug.ethz.ch
Date of birth	13 August, 1978
Citizenship	Japan

EDUCATION

2004-	ETH Zurich, Institute of Structural Engineering (IBK), D-BAUG Ph.D. student in the Group of Risk and Safety
2003-2004	University of Tokyo, Institute of Environmental Studies Ph.D. student
2001-2003	University of Tokyo, Institute of Environmental Studies **Master of Environmental Studies**
1997-2001	University of Tokyo, Faculty of Engineering (Architecture) **Bachelor of Engineering**

PROFILE

Awards

2005	Japan Association for Wind Engineering Award (shourei-sho) from the Japan Association for Wind Engineering
2003	Award for excellent master thesis from department of environmental studies, courses of socio-cultural and socio-physical environment, the University of Tokyo

Scholarship

2003-2004	Research fellowship for young scientists of the Japan Society for the Promotion of Science (JSCS), DC1
1997-2001	Scholarship student of the Kinoshita Scholarship Foundation

Language

Japanese	Native
English	Very good
German	Good