FP7 2013 Cooperation Work Programme
Theme 6: Environment (Including Climate Change)

Novel indicators for identifying critical *INFRA*structure at *RISK* from Natural Hazards

Deliverable D6.1

Stress Test Methodologies

Primary Authors	Yuliya Avdeeva and Pieter van Gelder/Probabilistic Solutions Consult and Training BV (PSCT)
WP	6
Submission Date	26/May/2014
Primary Reviewer	Eugene O' Brien/Roughan & O' Donovan Limited (ROD)
Dissemination Level	PU

This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 603960.

Project Information

Project Duration: 1/10/2013 - 30/09/2016

Project Coordinator: Professor Eugene O' Brien
Roughan & O' Donovan Limited
eugene.obrien@rod.ie

Work Programme: 2013 Cooperation Theme 6:
Environment (Including Climate Change).

Call Topic: Env.2013.6.4-4 Towards Stress Testing of Critical Infrastructure Against Natural Hazards-FP7-ENV-2013-two stage.

Project Website: www.infrarisk-fp7.eu

Partners:

Roughan & O' Donovan Limited, Ireland

Eidgenössische Technische Hochschule Zürich, Switzerland.

Dragados SA, Spain.

Gavin and Doherty Geosolutions Ltd., Ireland.

Probabilistic Solutions Consult and Training BV, Netherlands.

Agencia Estatal Consejo Superior de Investigaciones Científicas, Spain.

University College London, United Kingdom.

PSJ, Netherlands.

Stiftelsen SINTEF, Norway.

Ritchey Consulting AB, Sweden.

University of Southampton (IT Innovation Centre), United Kingdom.

Document Information

Version	Date	Description	Primary Author
Rev01	6/4/2014	Draft for comment	Avdeeva and Van Gelder
Rev02	8/4/2014	Revised	Avdeeva and Van Gelder
Rev03	26/5/2014	Final	E. O’ Brien

Executive Summary

The aim of this deliverable is to give an overview of stress testing methodologies and to review existing stress test methodologies in other domains than critical infrastructural domains. In particular this review report will address the methodologies in finance, mechanics and the nuclear domain. An overview of the variety of definitions of stress tests is given, followed by a general methodology to model stress tests. The uncertainty reduction in the failure probability of structures, which is one of the outcomes of a stress test, is investigated. This can be the failure probability with respect to ultimate limit states or serviceability limit states. Finally, the potential for stress tests of critical infrastructure is presented.

Table of Contents

1.0 INTRODUCTION

The aim of this report is to give an overview of stress testing methodologies and to review existing stress test methodologies in other domains than critical infrastructural domains. In particular this review report will address the methodologies in finance, mechanics and the nuclear domain. This report is organised as follows. First, an overview of the variety of definitions of stress tests is given, in Section 2. This is followed by a general methodology to model stress tests in Section 3. In Section 4, the uncertainty reduction which is a consequence of stress tests is investigated. In Sections 5, 6 and 7, three sectors are briefly reviewed in which stress tests are being applied. In Section 8 the potential for stress tests of critical infrastructure is presented.

2.0 DEFINITION OF STRESS TESTS

The definition of stress testing which is given in the different literature sources highlights that it is testing of the system/structure under abnormal (exceptional) conditions.

According to the Merriam -Webster Dictionary (http://www.merriam-webster.com/dictionary/stress%20test), it is a test to show how strong something is.

According to Dictionary.com (http://dictionary.reference.com/browse/stress+test?s=t) it is;

1. Generally; A test, especially one conducted in a laboratory, to determine how much pressure, tension, wear, or the like that a given product or material can withstand.
2. Medicine/Medical; a test of cardiovascular health made by recording heart rate, blood pressure, electrocardiogram, and other parameters while a person undergoes physical exertion.
3. A simulation test to determine how a given institution, system, etc., would perform under greater than usual stresses or pressures: the government's stress test for big banks; mandatory stress tests for nuclear plants; a website stress test that simulates peak traffic.

According to Investopedia (http://www.investopedia.com/terms/s/stresstesting.asp) it is a simulation technique used on assets to determine their reactions to different situations. Stress tests are also used to gauge how certain stressors will affect a structure, company or industry. They are usually computer-generated simulation models that test hypothetical scenarios. The Monte Carlo simulation is one of the most widely used methods of stress testing.

According to BusinessDictionary.com (http://www.businessdictionary.com/definition/stress-testing.html) it is a test conducted on some equipment to determine how much of a load the system can handle before it breaks or reaches its limit. A stress test is often used to determine the maximum load on a system, and to evaluate if the system will be able to operate correctly in certain common situations.

Sometimes, stress tests can be 'free of charge'. Consider a flood defence structure which protects a low lying area against storm surges. Every storm that was survived by the structure improves the knowledge of the resistance of the structure (pushing the lower tail of the resistance distribution to the right). Every year the knowledge of the owner of the structure grows and the probability of failure of the structure falls, caused by these natural stress tests, generated by nature.

3.0 A METHODOLOGICAL FRAMEWORK FOR STRESS TESTING

When studying the safety of critical infrastructure and the use of stress tests, it is important to develop a philosophy to discern between inherent uncertainty in time and in space. In this section a proposal for this philosophy will be described and applied to a fictitious example.

Inherent uncertainty in time means that the realizations of the process in the future remain uncertain. For instance, in a dike design, we have inherent uncertainty in time from the individual wave heights and water levels in front of the dike. Unlimited data will not reduce this inherent uncertainty. Inherent uncertainty in space is from a different kind. The properties of the foundation and the strength of the dike have only one realization per lifetime. An important aspect of this type of uncertainty is that without further investigations, the knowledge of the foundation increases with time. During the life of the structure information will be gained as each storm exceeding the previous that is survived by the structure, pushes the lower limit of the strength upward (Bayesian updating of the strength).

The above principles will be illustrated in this section by calculating the probability distribution of the life time to failure of a structure with a resistance with inherent uncertainty in space subjected to a yearly maximal load with inherent uncertainty in time. From this distribution the conditional failure rate will be derived. Two examples will be discerned in which the standard deviation of the resistance will be varied against the standard deviation of the yearly maximal load. The consequences on the conditional failure rate will be analysed.

3.1 Inherent Uncertainty in Time

Stochastic processes running in time (individual wave heights, significant wave heights, water levels, discharges, etc.) are examples of the class of inherent uncertainty. Unlimited data will not reduce this uncertainty. The realisations of the process in the future stay uncertain. The probability density function (p.d.f.) or the cumulative probability distribution function (c.d.f.) and the auto-correlation function describe the process.

In the case of a periodic stationary process like a wave field, the autocorrelation function will have a sinusoidal form and the spectrum as the Fourier-transform of the autocorrelation function gives an adequate description of the process. Attention should be paid to the fact that the well known wave energy spectra such as Pierson-Moskowitz and Jonswap are not always able to represent the wave field at a site. In quite some practical cases, swell and wind wave form a wave field together. The presence of two energy sources may be clearly reflected in the double peaked form of the wave energy spectrum.

An attractive aspect of the spectral approach is that the inherent uncertainty can be easily transferred through linear systems by means of transfer functions. By means of the linear wave theory, the incoming wave spectrum can be transformed into the spectrum of wave loads on a flood defence structure. The probability density function (p.d.f.) of wave loads can be derived from this wave load spectrum. Of course it is assumed here that no wave breaking takes place in the vicinity of the structure. In the case of non-stationary processes, that are governed by meteorological and atmospheric cycles (sign. wave height, discharges) the p.d.f. and the autocorrelation function are needed. Here the autocorrelation function gives an impression of the persistence of the phenomenon. The persistence of rough and calm conditions is of utmost importance in workability and serviceability analyses.

If the interest is directed to the analysis of ultimate limit states e.g. sliding of the structure, the autocorrelation is eliminated by selecting only independent maxima for the statistical analysis. If this selection method does not guarantee a set of homogeneous and independent observations, physical or meteorological insights may be used to homogenise the dataset. For instance if the fetch in the NW-direction is clearly maximal, the dataset of maximum significant wave height could be limited to NW-storms. If such insight fails, one could take only the observations exceeding a certain threshold (peaks over threshold or P.O.T.) into account hoping that this will lead to the desired result. In the case of a clear yearly seasonal cycle, the statistical analysis can be limited to the yearly maxima.

Special attention should be given to the joint occurrence of significant wave height Hs and spectral peak period Tp. A general description of the joint p.d.f. of Hs and Tp is not known. A practical solution for extreme conditions considers the significant wave height Hs and the wave steepness sp as independent stochastic variables to describe the dependence structure between Hs and Tp. This is a conservative approach as extreme wave heights are more easily realised than extreme peak periods. For the practical description of daily conditions (SLS) the independence of the wave steepness sp and peak period Tp seems sometimes a better approximation. Also the dependence of water levels and significant wave height should be explored because the depth limitation to waves can be reduced by wind setup. Here the statistical analysis should be clearly supported by physical insight. Moreover it should not be forgotten that shoals could be eroded or accreted due to changes in current or wave regime induced by the construction of the flood defence structures.

3.2 Inherent Uncertainty in Space

Soil properties can be described as stochastic processes in space. From a number of field tests the p.d.f. of the soil property and the (three-dimensional) autocorrelation function can be fixed for each homogeneous soil layer. Here the theory is more developed than practical knowledge. Numerous mathematical expressions are proposed in the literature to describe the autocorrelation. No clear preference has however emerged yet as to which functions describe the fluctuation pattern of the soil properties best. Moreover the correlation length (distance where correlation becomes zero) seems to be of the order of 30 to 100m while the spacing of traditional soil mechanical investigations for flood defence structures is of the order of 500m. So it seems that the intensity of the soil mechanical investigations has to be increased considerably if reliable estimates are to be made of the autocorrelation function.

The acquisition of more data has a different effect in the case of stochastic processes in space than in time. As breakwater structures are immobile, there is only one single realisation of the field of soil properties. Therefore the soil properties at the location could be exactly known if sufficient soil investigations were done. Consequently the actual soil properties are fixed after construction, although not completely known to man. The uncertainty can be described by the distribution and the autocorrelation function, but it is in fact a case of lack of information.

An important aspect of this type of uncertainty is that without further investigations, the knowledge of the foundation increases with time. During the life of the structure, information will be gained as each storm exceeding the previous that is survived by the structure, pushes the lower limit of the strength upward (Bayesian updating of the strength).

3.3 Effect of Uncertainty on the Lifetime Reliability

From an engineering point of view the Bayesian approach that takes all uncertainties into account as p.d.f.'s, reflects the designer's intuition very well. Keeping the physical structure equal, an increase in uncertainty of any variable increases the formal probability of failure too

From this point of view there is no difference between inherent, statistical and model uncertainty; all have to be incorporated in the probabilistic calculations. In the probabilistic calculations however a difference occurs between uncertainties that have many (e.g. yearly) realizations during the lifetime of the structure and those that have only one connected to the specific structure and the site. Every storm season shows an independent maximum, Hs every year. The properties of the foundation and the strength of the flood defence structure have only one realisation per structure. Consequently the probability of failure is not solely a property of the structure but also a result of our lack of knowledge.

The effect of uncertainty on the lifetime reliability will be illustrated by the following hypothetical example of the probability of failure of a flood defence structure. In the example, one type of load function and three types of resistance function are considered.

For the load function one can think of the exponential p.d.f. of the wave heights in front of the structure. For the three types of resistance function one can think of the normal p.d.f.'s of the (quite certain; uncertain; very uncertain, resp.), crest height of the structure.

Failure is defined by Z<0, in which Z=R-S with R the resistance function and S the load function. A series of observations can be made.

<u>First Observation</u>

Let the failure at time i be defined as {Zi<0}. Consider the correlation in the reliability in two subsequent years i and i+1.

$$\rho(Zi, Zi+1) \; = \; {\sigma_R}^2/({\sigma_R}^2 + {\sigma_S}^2)$$

This result can be derived from the calculation of

$$cov(Zi, Zi+1) = E(((R - Si) - (\mu_R - \mu_S))((R - Si + 1) - (\mu_R - \mu_S))).$$

It can be noted from the above expression that if σ_R increases, then ρ converges to one. In words: if the standard deviation of the yearly maximal wave load is large in relation to the standard deviation of the resistance, the dependence between failures in two subsequent years is low. If however, keeping the standard deviation of the reliability function equal, the opposite is true, failures in subsequent years are dependent.

<u>Second Observation</u>

Consider the probability of failure at a certain moment (year number N), assuming that no failure occurred before that time. This probability is given by:

$$h(N) \; = \; f_L(N)/(1 - F_L(N))$$

where h(N) is called the conditional failure rate (or hazard function) and $F_L(N)$ is the cdf of N (subscript L for lifetime). It can be noted that if σ_R = 0, then h(N) remains constant. If σ_R increases, then h(N) converges to 0.

Conditional failure rates are usually decreasing functions in the beginning of the lifetime of the structure. This phenomenon is also known as infant mortality; i.e. early failure of the structure that is attributable to construction defects (see Hoeg, 1996 for some very interesting statistics of dam failures occurring, almost all in the beginning of their lifetimes, Figure 3.1)). Conditional failure rates are increasing functions at the end of the lifetime of the structure, because of deterioration of the structure. Conditional failure rates therefore have a U-shaped (or bath tub curved) form (see also Langley, 1987).

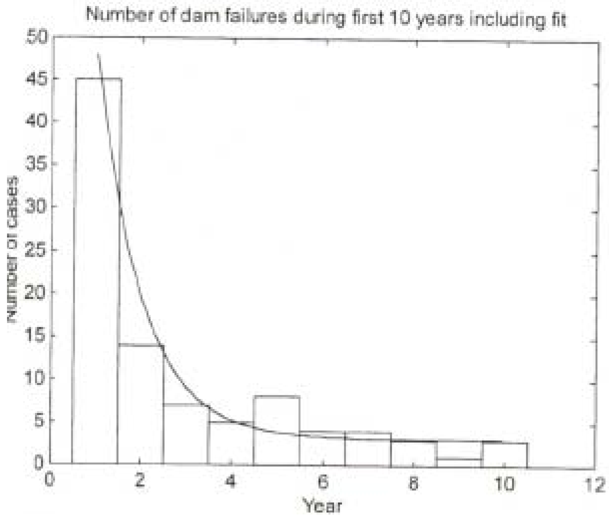

Figure 3.1: After Hoeg, 1996

Third observation

It was already noted that if the standard deviation of the yearly maximal wave load is large in relation to the standard deviation of the resistance, the dependence between failure in two subsequent years is low. Therefore, if the probability of failure is say p per year, then the failure probability is approximately N.p during the lifetime of N years. If however, keeping the standard deviation of the reliability function equal, the opposite is true, the failure in subsequent years are dependent and in that case the probability of failure in the first year is p and the probability of failure in the life time too. In the first case the conditional failure rate is constant over time and equal to p. In the second case however the conditional failure rate equals p in the first year and falls to zero afterwards.

In probabilistic calculations of structures a difference occurs between uncertainties that have many realizations during the structure's lifetime (e.g. wave heights) and uncertainties that have only one realization (e.g. soil parameters). This difference leads for example to the observation that every storm that was survived by the structure improves the knowledge of the p.d.f. of the resistance (pushing the lower tail to the right). Every year the knowledge of the owner of the structure grows and the probability of failure of the structure falls. In exactly the same way, the failure probability of the structure can be improved by investigating, e.g. the quality of the foundation assuming that this leaves the average unchanged and reduces the uncertainty. These ideas have been illustrated in this section by a simple hypothetical example of a flood defence structure. An analysis of the effect of inherent uncertainty in time and space on the model has been given.

4.0 REDUCING UNCERTAINTY BY STRESS TESTING

Apart from inherent (non-reducible) uncertainties, we can also distinguish epistemic (reducible) uncertainties. Uncertainties in decision and risk analysis can primarily be divided into these two categories: uncertainties that stem from variability in known (or observable) populations and therefore represent randomness in samples (inherent uncertainty), and uncertainties that come from basic lack of knowledge of fundamental phenomena (epistemic uncertainty).

Inherent uncertainties represent randomness or the variations in nature. For example, even with a long history of data, one cannot predict the maximum water level that will occur in, say, the coming year at the North Sea. It is not possible to reduce inherent uncertainties.

Epistemic uncertainties are caused by lack of knowledge of all the causes and effects in physical systems, or by lack of sufficient data. It might be possible to obtain the type of the distribution, or the exact model of a physical system, when enough research could and would be done. Epistemic uncertainties may change as knowledge increases.

The inherent uncertainty and epistemic uncertainty can be subdivided into the following five types of uncertainty (according to Paté-Cornell 1996): inherent uncertainty in time and in space, parameter uncertainty and distribution type uncertainty (together also known as statistical uncertainty) and finally model uncertainty, Figure 4.1).

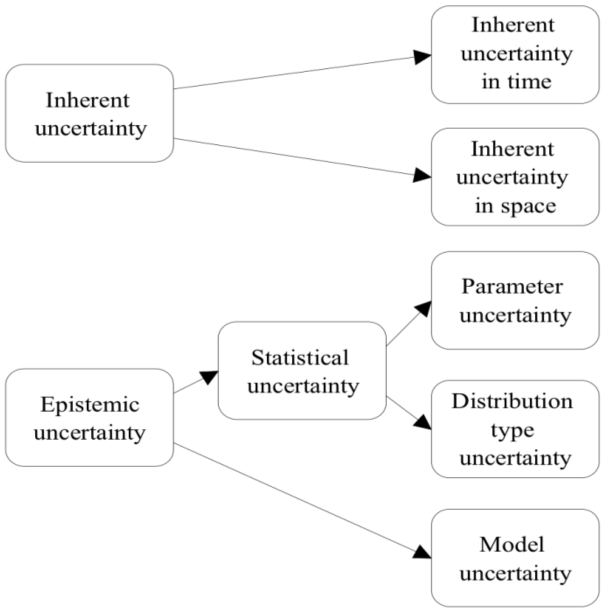

Figure 4.1: Types of uncertainty based on Pate-Cornell (1996)

i) Inherent uncertainty in time

Stochastic processes running in time such as the occurrence of water levels and wave heights are examples of the class of inherent uncertainty in time. Unlimited data will not reduce this uncertainty because the realizations of the process in the future stay uncertain.

ii) Inherent uncertainty in space

Random variables that represent the fluctuation in space, such as the dike height. Just as for inherent uncertainty in time, it holds that unlimited data (e.g. if the height would be known every centimetre) will not reduce this uncertainty. There will always still be a fluctuation in space.

iii) Parameter uncertainty

This uncertainty occurs when the parameters of a distribution are determined with a limited number of data. The smaller the number of data, the larger the parameter uncertainty.

iv) Distribution type uncertainty

This type represents the uncertainty of the distribution type of the variable. It is for example not clear whether the occurrence of the water level of the North Sea is exponentially or Gumbel distributed or whether it has a completely different distribution.

Note, a choice was made to divide statistical uncertainty into parameter- and distribution-type uncertainty although it is not always possible to draw the line; in the case of unknown parameters (because of lack of observations), the distribution type will be uncertain as well.

Later in this section we will try to distinguish what type of uncertainty a stochastic parameter represents. Since parameter uncertainty and distribution type uncertainty cannot be discerned, another practical – less scientific – division has been chosen. The statistical uncertainty is divided in two parts: 'statistical uncertainty of variations in time' and 'statistical uncertainty of variations in space'.

v) Statistical uncertainty of variations in time

When determining the probability distribution of a random variable that represents the variation in time of a process (like the occurrence of a water level), there essentially is a problem of information scarcity. Records are usually too short to ensure reliable estimates of low-exceedance probability quantiles in many practical problems. The uncertainty caused by this shortage of information is the statistical uncertainty of variations in time. This uncertainty can theoretically be reduced by keeping record of the process for the coming *centuries.*

vi) Statistical uncertainty of variations in space

When determining the probability distribution of a random variable that represents the variation in space of a process (like the fluctuation in the height of a dike), there essentially is a problem of shortage of measurements. It is usually too expensive to measure the height or width of a dike in great detail. This statistical uncertainty of variations in space can be reduced by taking more measurements.

vii) Model uncertainty

Many of the engineering models that describe the natural phenomena like wind and waves are imperfect. They may be imperfect because the physical phenomena are not known (for example when regression models without underlying theory are used), or they can be imperfect because some variables of lesser importance are omitted in the engineering model for reasons of efficiency.

It was mentioned before that inherent uncertainties represent randomness or the variations in nature. Inherent uncertainties cannot be reduced.

Epistemic uncertainties, on the other hand, are caused by lack of knowledge. Epistemic uncertainties may change as knowledge increases.

In general there are three ways to increase knowledge:

- Gathering data, for instance via stress testing

- Research

- Expert judgment

Data can be gathered by taking measurements, doing stress tests or by keeping record of a process in time. Research can, for example, be done into the physical model of a phenomenon or into the better use of existing data. By using expert opinions it is possible to acquire the probability distributions of variables that are too expensive or practically impossible to measure.

The goal of all this research obviously is to reduce the uncertainty in the model. Nevertheless it is also thinkable that uncertainty will increase. Research or a stress test might show that an originally flawless model actually contains a lot of uncertainties, or after taking some measurements the variations of the dike height can be a lot larger. It is also thinkable that the average value of the variable will change because of the research that has been done.

The consequence is that the calculated probability of failure will be influenced by future research. In order to guarantee a stable and convincing flood defence policy after the transition, it is important to understand the extent of this effect.

It will now be investigated and discussed what influence the future reduction of uncertainty can have on the probability of failure and how to present this influence.

4.1 The Effect of Uncertainty due to lack of Information

The problem of lack of information in engineering models is studied in detail in this section.

The information about a random variable can be updated with the help of expert judgements. Cooke (1991) describes various methods to do so. Expert judgements can also be used to reduce the uncertainty of the quantiles. In order to include the results of stress tests in the quantile estimation of a certain quantity, the following approach is proposed in this section. The approach is applied to a case study of quantile estimation of water levels.

Consider R the random variable describing the river level with exceedance probability p per year.

Consider H the random variable describing the height of the dike modelled with a normal distribution with mean μ_H and standard deviation Φ_H.

The effect of the value of information from the stress test on the random variable R may be modelled by correcting its original mean value μ_R to its new value $\mu_R + v\Phi_I$ in which v is the standard normal distribution and Φ_I is the standard deviation of the information following from the stress test. Furthermore the standard deviation of R will be reduced from $\sqrt{(\Phi_R^2 + \Phi_I^2)}$ to Φ_R under the new information resulting from the stress test. This is summarised in Table 4.1 below;

Without Information			With Information	
	μ	Φ	μ	Φ
R	μ_R	$\sqrt{(\Phi_R^2 + \Phi_I^2)}$	$\mu_R + v\Phi_I$	Φ_R
H	μ_H	Φ_H	μ_H	Φ_H
$\beta_{ni}=(\mu_R-\mu_H)/\sqrt{(\Phi_R^2 + \Phi_I^2 + \Phi_H^2)}$			$\beta_{wi}=(\mu_R+v\Phi_I-\mu_H)//\sqrt{(\Phi_R^2 + \Phi_H^2)}$	

Table 4.1: The effect of information of the stress test on the random variables R and H (wi = without information, wi = with information)

The exceedance probability or reliability index β_{wi} after including stress test information can be seen as a random variable with a normal distribution with the following mean and standard deviation:

$$\beta wi \sim N(\,(\mu_R - \mu_H)//({\Phi_R}^2 + {\Phi_H}^2), \Phi_I/\sqrt{(\Phi_R^2 + \Phi_H^2)}\,)$$

The implications of the changes in β and the related flooding frequency are an important question.

Some uncertainties might be reduced in the future. This means that the uncertainty in the future probability of failure will be smaller. One could say that those uncertainties add to the 'uncertainty' of the probability of failure.

The obvious problem is that it cannot be predicted which uncertainties will be reduced in the future. There are several philosophies about what uncertainty will remain part of the future probability of failure.

It is important to realize that, for every philosophy, the difference between the 'true' probability of failure and the 'uncertainty' of the probability of failure is an artificial difference. If all uncertainties are integrated this will result in the same, current probability of failure.

Let us consider four different philosophies (options) about the uncertainties that might be reduced and form the 'uncertainty' of the probability of flooding, and the uncertainties that cause the 'true' probability of flooding, as given in Table 4.2 below;

	Option 1	Option 2	Option 3	Option 4
Inherent uncertainty (in time)	P_f	P_f	P_f	P_f
Inherent uncertainty (in space)	P_f	P_f	P_f	-
Statistical uncertainty (of variations in time)	P_f	P_f	-	-
Statistical uncertainty (of variations in space)	P_f	-	-	-
Model uncertainty	P_f	-	-	-

Table 4.2: Four Different Options

Option 1

No uncertainty will be reduced. All uncertainties in the model will be integrated to determine the probability of failure. This probability of failure also represents the current, actual probability of failure.

Option 2

A practical division is made between the uncertainties that might be influenced by taking measurements or doing research and the remaining uncertainties. Statistical uncertainty of variations in time will be a part of the probability of failure. This uncertainty can only theoretically be reduced by keeping record of the underlying process for the coming centuries.

Option 3

A theoretical division is made between inherent uncertainty and epistemic uncertainty. The assumption is that through research, all epistemic uncertainty disappears. This would result in a probability of failure that is solely caused by inherent uncertainties.

Option 4

It is assumed that only inherent uncertainty in time causes the probability of failure. This is also the philosophy used in the present safety standards.

A simplified model which calculates the probability of failure of a dike, will be used to demonstrate the earlier philosophy. Let us assume a limit state function:

$$Z = m \cdot (h_d - h_w)$$

m is model_factor [-] : m = N(1; 0.5)

h_d is dike height [m +NAP] : h_d =N(5; 0.25)

h_w is water level [m +NAP] : h_w =N(3.5; σ_{hw})

σ_{hw} is distribution parameter of h_w : σ_{hw}=N(0.5;0.05)

It is possible to identify the type(s) of uncertainty that every variable represents, Table 4.3.

m	model uncertainty
h_d	inherent and statistical uncertainty (space)
h_w	inherent uncertainty (time)
σ_{hw}	statistical uncertainty (time)

Table 4.3: The Variables and their uncertainties

A FORM analysis of this reliability function results in a probability of failure of 2.7 x 10^{-4} and a reliability index of 3.46. The influence factors of the variables are presented in Table 4.4.

	α_i^2
h_w	0.77
σ_{hw}	0.06
h_d	0.16
m	0.01

Table 4.4: The Influence Factors

In the Table 4.5 is shown, per option of presentation and per parameter, whether or not the uncertainty is part of the probability of failure. It is assumed that the uncertainty in the dike height represents exactly half the inherent uncertainty and half the statistical uncertainty.

	Option 1	**Option 2**	**Option 3**	**Option 4**
h_w	1	1	1	1
σ_{hw}	1	1	-	-
h_d	1	0.5	0.5	-
m	1	-	-	-

Table 4.5: The Four Options of the Case Study

Let pi be the values in Table 4.6, then with *u2 = Σ pi. *i2 it is quite simple to obtain the results per option as displayed in the table below. The results are also presented in Figure 4.2.

	Option 1	**Option 2**	**Option 3**	**Option 4**
α_u^2	1	0.91	0.85	0.77
α_v^2	0	0.09	0.15	0.23
β_m	3.46	3.62	3.75	3.94
σ_β	0	0.31	0.42	0.55

Table 4.6: Results for the Four Options

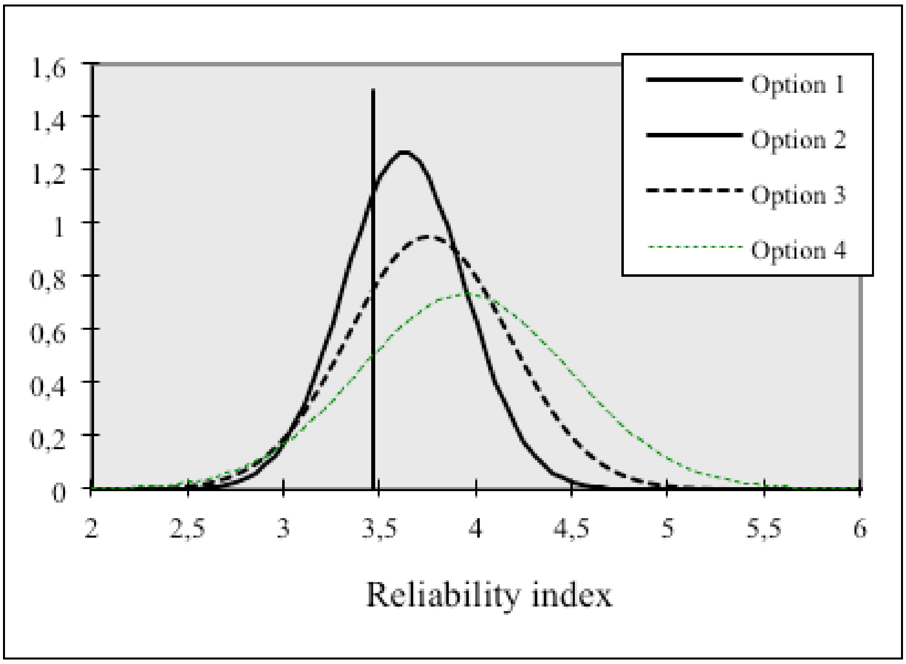

Figure 4.2: Probability distributions of the reliability index

In Figure 4.2 the current probability of failure is shown in four different ways. Option 1 gives that probability with no uncertainty. With each consecutive option, more uncertainty is removed from the 'true' probability of failure and added to the 'uncertainty' of the probability of failure. It is clear that the more uncertainty is removed from the resulting probability of failure; the higher the mean and the larger the standard deviation of the distribution of the reliability index will be.

In Table 4.4 it can be seen that the uncertainty in the occurrence of high water levels has the largest influence on the probability of failure. Since this uncertainty cannot be reduced, this will always be part of the probability of failure. Therefore the mean of the distribution of the reliability index has not shifted considerably.

It can be concluded that the method described in this section is a very practical and simple method to get insight into the problems concerned with the lack of information in engineering models and the use of stress test information.

It shows that the more uncertainty is expected to be reduced in the future, the higher the mean and the larger the standard deviation of the distribution of the reliability index will be. The method is demonstrated for a simple reliability function, but this methodology can also work for a very complex reliability function with a large number of variables.

Methodologies in stress testing are based on Monte Carlo simulation, and (non)-linear regression.

The Monte Carlo method uses the possibility of drawing random numbers from a uniform probability density function between zero and one. Practically all programming languages include a standard procedure for this. The non-exceedance probability of an arbitrary random variable is uniformly distributed between zero and one, regardless of the distribution of the variable. In formula:

$$F_X(X) = X_u$$

in which:

X_u is the uniformly distributed variable between zero and one;

$F_X(X)$ is the non-exceedance probability $P(X < X)$.

Thus, for the variable X:

$$X = F_X^{-1}(X_u)$$

in which:

$F_X^{-1}(X_u)$ is the inverse of the probability distribution function of X.

Using this formula a random number X can be generated from an arbitrary distribution $F_X(X)$ by drawing a number X_u from the uniform distribution between zero and one.

This way of drawing random numbers is generally applicable. However, for distributions for which the inverse probability distribution function $F_X^{-1}(X_u)$ is not known analytically, this method can lead to a lot of iterative calculations. Hence, other less maths-intensive methods for drawing from (for example) a normal distribution were sought.

More or less the same way, base variables of a statistical vector can be drawn from a known joint probability distribution function. However, the joint probability distribution function must then be formulated as the product of the conditional probability distributions of the base variables of the vector. In formula this is:

$$F_{\vec{X}}(\vec{X}) = F_{X_1}(X_1)\, F_{X_2|X_1}(X_2 \mid X_1) \cdots F_{X_m|X_1, X_2, \ldots, X_{m-1}}(X_m \mid X_1, X_2, \ldots, X_{m-1})$$

By taking m realisations of the uniform probability distribution between zero and one, a value can be determined for every X_i:

$$\begin{aligned} X_1 &= F_{X_1}{}^{-1}(X_1) \\ X_2 &= F_{X_2|X_1}{}^{-1}(X_2 \mid X_1) \\ &\ldots. \\ X_m &= F_{X_m|X_1, X_2, \ldots, X_{m-1}}{}^{-1}(X_m \mid X_1, X_2, \ldots, X_{m-1}) \end{aligned}$$

If the base variables are statistically independent, this can be simplified to:

$$X_i = F_{X_i}^{-1}\left(X_{u_i}\right)$$

By repeating this procedure a large number of times, the probability of failure can be estimated with:

$$P_f \approx \frac{n_f}{n}$$

in which:

n is the total number of simulations,

n_f is the number of simulations, for which $Z < 0$.

The simulation's relative error is:

$$\varepsilon = \frac{\frac{n_f}{n} - P_f}{P_f}$$

The expected value of the relative error is zero and the standard deviation is:

$$\sigma_\varepsilon = \sqrt{\frac{1 - P_f}{nP_f}}$$

Based on the central limit theorem, the error is normally distributed, provided n is sufficiently large. The probability that the relative error is smaller than the given value E is then:

$$P(\varepsilon < E) = \Phi\left(\frac{E}{\sigma_\varepsilon}\right)$$

Thus, for a reliability of $\Phi(k)$, the relative error is smaller than $E = k\sigma_\varepsilon$. For a wanted k and E the required number of simulations n can be determined with:

$$n > \frac{k^2}{E^2}\left(\frac{1}{P_f} - 1\right)$$

If, for instance, a reliability of 95 % is required for a maximum relative error E = 0.1, the required number of simulations amounts to:

$$n > 400\left(\frac{1}{P_f} - 1\right)$$

The number of simulations is therefore still dependent on the probability of failure. For small probabilities of failure, the required number of simulations is very large.

To reduce the large number of calculations, an approximation method is often used to estimate the probability distribution of the reliability function. With the simulated data the probability distribution of the reliability function is estimated. The found distribution can subsequently be used to estimate the probability of failure. This method does offer a quick way of gaining an insight into the shape of the probability distribution of the reliability function, but is generally too inaccurate for estimating the probability of failure. A method to reduce the number of simulations with sufficient accuracy, is applying Importance Sampling. Importance Sampling increases the failure space relative to the total integration space. If the random numbers are drawn from the Importance Sampling probability distribution function, instead of the original probability distribution, more points in the failure area are found.

When employing Importance Sampling, the probability of failure can be approximated with:

$$P_f = \frac{\sum_{i=1}^{n} \mathbf{1}(Z) \frac{f_{\bar{X}_i}(\bar{X}_i)}{h_{\bar{X}_i}(\bar{X}_i)}}{n}$$

in which:

$\bar{X}_i$ is the vector with random numbers, drawn from the Importance Sampling probability distribution;

$f_{\bar{X}_i}\left(\overline{X}_i\right)$ is the value of the original probability density function for the vector $\overline{X}_i$;

$h_{\bar{X}_i}\left(\overline{X}_i\right)$ is the value of the Importance Sampling probability density function for vector $\overline{X}_i$;

n is total number of realisations.

Figure 4.3 gives the flow chart of the Monte Carlo simulation with Importance Sampling. An increase of the failure space relative to the total space (from which the random variables are realised) with a factor 100, reduces the number of required simulations by a factor of 100.

For the choice of the Importance Sampling probability distribution, the same applies as for the application in numerical integration. For a one-dimensional problem Bucher determined the parameters of the optimal normal Importance Sampling probability distribution. These are:

$$\mu_{h_x} = \frac{\exp\left(-\frac{\beta^2}{2}\right)}{\sqrt{2\pi}\,\Phi(-\beta)} \quad \text{and} \quad \sigma_{h_x} = \sqrt{1+\beta\mu_{h_x} - \mu_{h_x}{}^2}$$

in which:

$$\beta = \Phi^{-1}\left(P_f\right).$$

The magnitude of the probability of failure P_f is not known in advance. The calculation of the probability of failure is iterative and involves determining the Importance Sampling probability distribution after each iteration.

Importance Sampling entails taking realisations from a distribution that has been moved to the design point. If the failure space is defined by more than one random variable, it is possible that the failure space is not concentrated around a design point. This is certainly the case if the definition of the failure space is discontinuous.

In such a case, Importance Sampling for one of the variables will lead to an underestimation of the probability of failure. In such cases a method known as Adaptive Sampling, can be applied. Adaptive Sampling first requires the calculation of the conditional probability of failure for each variable, to which end the other variables are set equal to their expected values. The probability of failure is determined by simulation of the variable concerned. The remaining variables are kept constant. The expected value of the conditional Importance Sampling function can be determined with the conditional probabilities of failure.

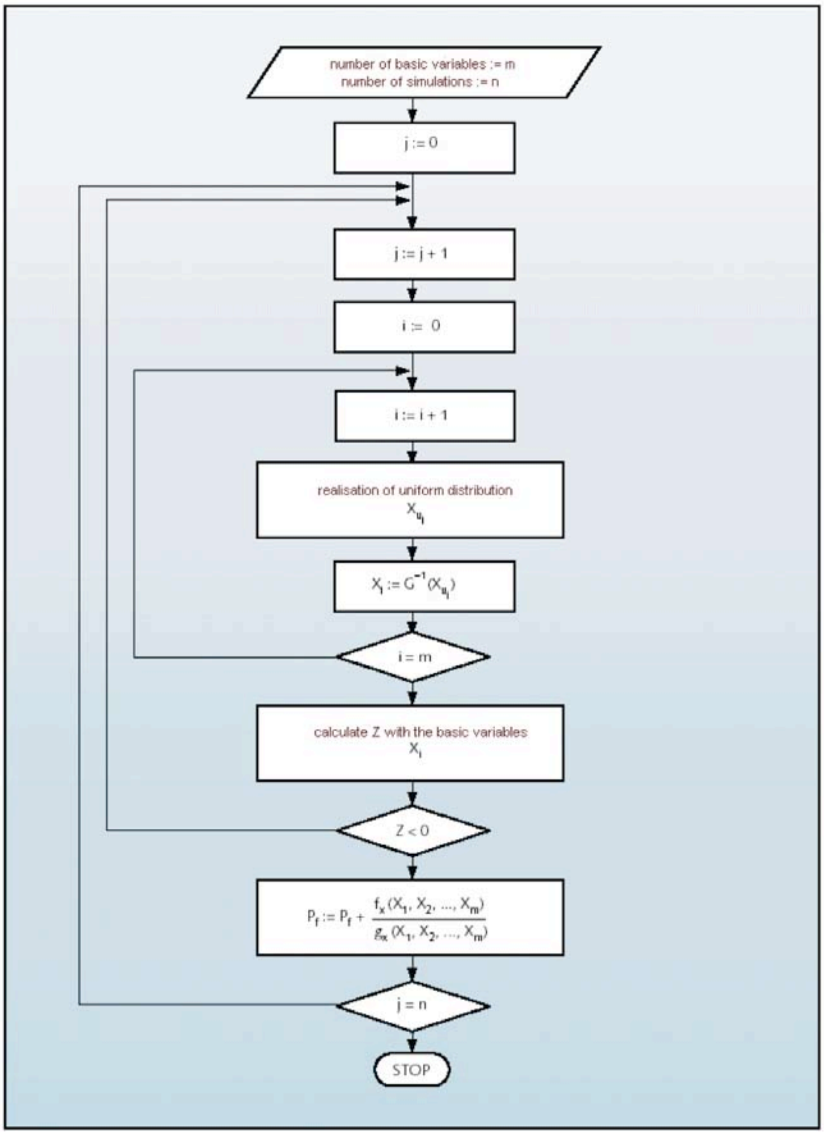

Figure 4.3: Monte Carlo simulation with Importance Sampling

The expected value of the joint Importance Sampling function is calculated with:

$$\mu_{g(X_i)} = \frac{\mu_{h(X_i)} P_{f_i}}{\sum_{i=1}^{n} P_{f_i}}$$

in which:

$\mu_{g(X_i)}$ is the expected value of the joint Importance Sampling function;

$\mu_{h(X_i)}$ is the expected value of the conditional Importance Sampling function with a probability of failure $P_f = P_{f_i}$;

P_{f_i} is the conditional probability of failure

Around the point $\mu_{g(\overline{X})}$ points are simulated in the failure space, as for the calculation of P_{f_i}. The variances and covariances of these points are determined with a known average $\mu_{g(X_i)}$. The covariance matrix is subsequently diagonalised and the variables are transformed. The transformation adjusts the eigenvalues of the covariance matrix according to:

$$\lambda_i^* = \sigma_\varepsilon \sum_{i=1}^{n} \lambda_i + (1 - \sigma_e \lambda_i)$$

in which:

λ_i is the calculated eigenvalue of the covariance matrix;

$$\sigma_\varepsilon = \sqrt{\frac{1-P_f}{nP_f}}\ .$$

The transformed variables are used to determine the probability of failure with the corresponding Importance Sampling function. After determining the probability of failure the Importance Sampling function can be adjusted. This method is suitable for processing in a computer program, but less appropriate for use in combination with general programs for Monte Carlo Simulations.

Regression analysis can be used in stress testing to extrapolate the outcomes of the non-destructive responses to the failure domain. The multiple regression analysis checks if there is, on the basis of the correlation of multiple independent variables with the dependent variables, a predictive dependency. The multiple regression analysis in stress testing uses continuous or ordinal data, but may also include one or more categorical variables as an independent variable (for instance manufacturer of the product, etc.).

Multiple regression analysis is used to test whether several independent variables influence a dependent variable and whether this is a positive or a negative effect. Also, it is possible to test interaction effects.

5.0 STRESS TESTS IN THE FINANCIAL INDUSTRY

For financial systems the stress test is a tool to determine the impact of shocks to a given financial instrument or financial institution (Committee on the Global Financial System, 2005). Two categories of methods could be identified: micro, conducted by individual institutions and macro, by central banks and supervisors, methods. If the stress test analyses the impact of one factor it can be considered as a sensitivity analysis; if it considers a few factors, then it is a scenario analyses, Figure 5.1.

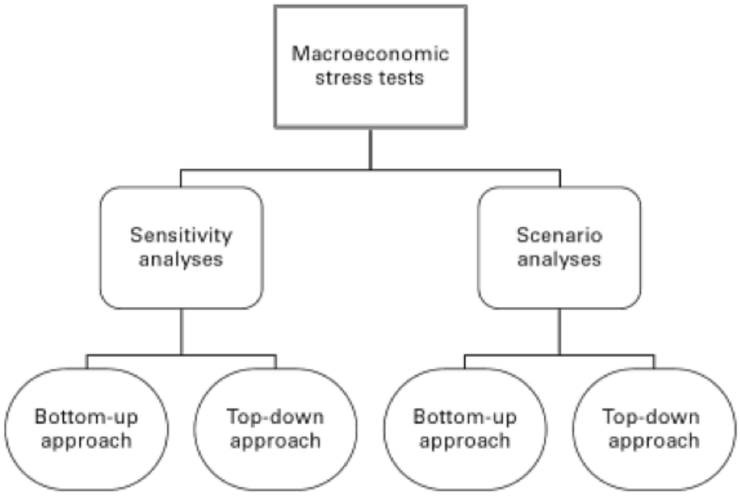

Figure 5.1: Macroeconomic stress tests (Stress-testing the Banking System: Methodologies and Applications, 2009)

The advantage of the bottom-up approach is that it is tailored to the individual institution and could be more realistic for a sensitivity analyses. The Top-down approach is more suitable for scenario analyses, more comparative, smaller loss of data during aggregation (Stress-testing the Banking System: Methodologies and Applications, 2009).

Scenario analyses could be based on historical scenarios (already happened before and may happen again), hypothetical (have not happened before) or reverse engineering (find scenarios looking at the failure of the system) (Licari J.M., Suárez-Lledó J, 2012). Three types of models are typically used: a structural econometric model, VAR (vector autoregressive), and statistical approaches. Foglia (2009) provides an overview of such models and their applications.

Financial stress tests examine what happens under certain extreme conditions. They answer questions such as the following:

- What happens if the market goes down by more than x%?
- What happens when interest rates rise or fall more than x%?
- What happens if the oil price rises 200%?
- What happens in a bankruptcy of a major customer, supplier or party in between?
- What happens if the currency rises or falls x%?

An example of a stress test report of a large Dutch bank (ING) is shown in the next Figure 5.2:

Name of the bank: ING Bank N.V.

Actual results at 31 December 2010	**million EUR, %**
Operating profit before impairments	7.999
Impairment losses on financial and non-financial assets in the banking book	-2.332
Risk weighted assets (4)	321.103
Core Tier 1 capital (4)	30.895
Core Tier 1 capital ratio, % (4)	9,6%
Additional capital needed to reach a 5 % Core Tier 1 capital benchmark	

Outcomes of the adverse scenario at 31 December 2012, excluding all mitigating actions taken in 2011	**%**
Core Tier 1 Capital ratio	8,7%

Outcomes of the adverse scenario at 31 December 2012, including recognised mitigating measures as of 30 April 2011	**million EUR, %**
2 yr cumulative operating profit before impairments	12.278
2 yr cumulative impairment losses on financial and non-financial assets in the banking book	-8.276
2 yr cumulative losses from the stress in the trading book	-1.052
of which valuation losses due to sovereign shock	*-237*
Risk weighted assets	391.282
Core Tier 1 Capital	33.860
Core Tier 1 Capital ratio (%)	**8,7%**
Additional capital needed to reach a 5 % Core Tier 1 capital benchmark	
Effects from the recognised mitigating measures put in place until 30 April 2011 (5)	
Equity raisings announced and fully committed between 31 December 2010 and 30 April 2011 (CT1 million EUR)	
Effect of government support publicly announced and fully committed in period from 31 December 2010 to 30 April 2011 on Core Tier 1 capital ratio (percentage points of CT1 ratio)	
Effect of mandatory restructuring plans, publicly announced and fully committed in period from 31 December 2010 to 30 April 2011 on Core Tier 1 capital ratio (percentage points of CT1 ratio)	

Additional taken or planned mitigating measures	percentage points contributing to capital ratio
Use of provisions and/or other reserves (including release of countercyclical provisions)	
Divestments and other management actions taken by 30 April 2011	
Other disinvestments and restructuring measures, including also future mandatory restructuring not yet approved with the EU Commission under the EU State Aid rules	0,7
Future planned issuances of common equity instruments (private issuances)	
Future planned government subscriptions of capital instruments (including hybrids)	-0,8
Other (existing and future) instruments recognised as appropriate back-stop measures by national supervisory authorities	
Supervisory recognised capital ratio after all current and future mitigating actions as of 31 December 2012, % (6)	8,6%

Notes

(1) The stress test was carried using the EBA common methodology, which includes a static balance sheet assumption and incorporates regulatory transitional floors, where binding (see http://www.eba.europa.eu/EU-wide-stress-testing/2011.aspx for the details on the EBA methodology).
(2) All capital elements and ratios are presented in accordance with the EBA definition of Core Tier 1 capital set up for the purposes of the EU-wide stress test, and therefore may differ from the definitions used by national supervisory authorities and/or reported by institutions in public disclosures.
(3) Neither baseline scenario nor the adverse scenario and results of the stress test should in any way be construed as a bank's forecast or directly compared to bank's other published information.
(4) Full static balance sheet assumption excluding any mitigating management actions, mandatory restructuring or capital raisings post 31 December 2010 (all government support measures and capital raisings fully paid in before 31 December 2010 are included).
(5) Effects of capital raisings, government support and mandatory restructuring plans publicly announced and fully committed in period from 31 December 2010 to 30 April 2011, which are incorporated in the Core Tier 1 capital ratio reported as the outcome of the stress test.
(6) The supervisory recognised capital ratio computed on the basis of additional mitigating measures presented in this section. The ratio is based primarily on the EBA definition, but may include other mitigating measures not recognised by the EBA methodology as having impacts in the Core Tier 1 capital, but which are considered by the national supervisory authorities as appropriate mitigating measures for the stressed conditions. Where applicable, such measures are explained in the additional announcements issued by banks/national supervisory authorities. Details of all mitigating measures are presented in the worksheet "3 - Mitigating measures).

Figure 5.2: Stress test result of ING bank

A summary of the 2011 EU-wide stress test conducted by the European Banking Authority (EBA), in cooperation with De Nederlandsche Bank (DNB), the European Central Bank (ECB), the European Commission (EC) and the European Systemic Risk Board (ESRB) is that under an adverse stress test scenario, the estimated consolidated Core Tier 1 capital ratio of ING would decline to 8.7% in 2012 compared to 9.6% as of end of 2010. ING would remain well above the hurdle rate of 5% Core Tier 1 ratio with surplus Core Tier 1 capital in 2012.

Financial stress tests have been made mandatory by governments (such as the UK FSA) for financial institutions, so that there is sufficient capital to cope during extreme conditions with possibly large losses. Such emphasis on adequate risk-adjusted capital requirements is demanded by Basel II. In stress testing, combinations of factors may be tested. There is also the possibility to use test cases such as the stock market crash of 1929, the Russian financial crisis in 1998, or other historical events such as the attacks on September 11, 2001, the collapse of Lehman Brothers, the Greek government in 2010 etc.

Stress tests reveal what happens when the scenarios come true. They provide insight into the vulnerability of a portfolio or a financial institution. Although extreme events do not need to actually happen, it does help to identify potential problems.

The stress tests in the financial sector can be seen, more or less, as sensitivity tests (showing the slopes of the outcome caused by a change in one of the input variables).

A proper stress test should show the reduction of the uncertainty in the probability of failure, as argued in Section 2 and 3 of this report. For this purpose, failure of a bank should be defined.

A bank is exposed to several types of risk, Figure 5.3, such as:

- Interest rate risk, Equity risk, Currency risk and Commodity risk

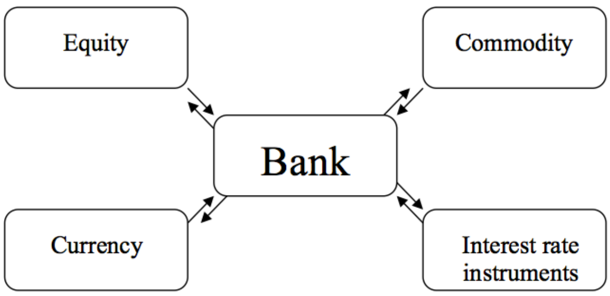

Figure 5.3: Risk factors of a bank

The value at risk (VAR) method calculates the PDF of the monetary loss of a bank within a certain period (1 day, 1 month, etc), as in Figure 5.4.

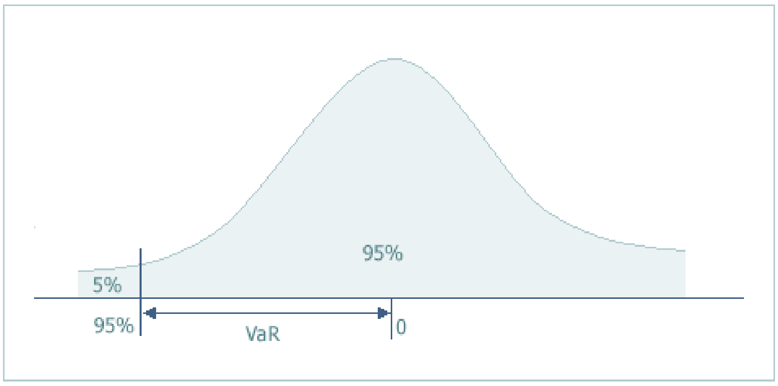

Figure 5.4: Value at Risk

A bank can be modelled as in Figure 5.5

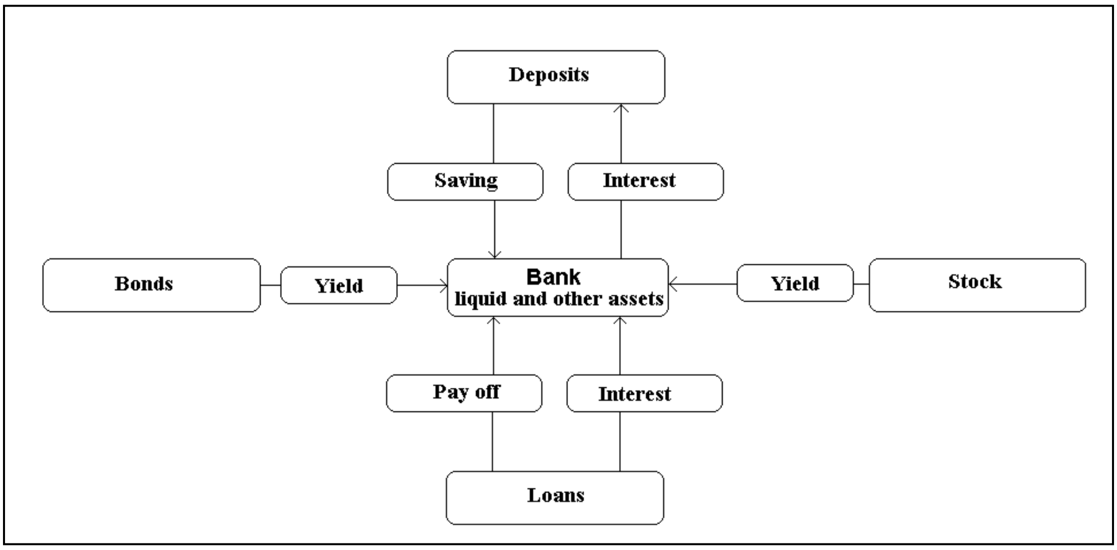

Figure 5.5: A model for a bank with inflow and outflow

Based on the above schematic model of a bank, the limit state function of a bank can be given by:

$$Z = C + L * (1 + r_l) + S * (1 + r_s) + B * (1 + r_b) + O * (1 + r_o) - D - (D + Sa) * r_d - Co$$

in which Z the bank's equity and the other variables are given by:

- Liquid funds *C*
- Deposits by clients *D*
  Interest paid on the deposits in one year r_d
  Savings by clients during one year *Sa*
- Amount invested in stocks *S*
  Rate of return on stocks in one year r_s
- Amount invested in bonds *B*
  Rate of return on bonds in one year r_b
- Value of outstanding loans *L*
  Interest collected on loans in one year r_l
  Pay off on loans in one year *Po*
- Other assets *O*

Rate of return on other assets in one year $\boldsymbol{r_o}$
Cost of operations in one year *Co*

P(Z<0) can now be calculated given the probability distributions and parameters of the above variables. The uncertainty in P(Z<0) can be reduced after a proper stress test has been applied.

6.0 STRESS TESTS IN MECHANICAL AND CIVIL ENGINEERING

Possible stresses that are being applied in the mechanical sector are:

- random vibration
- rapid temperature transitions
- voltage margining
- frequency margining
- proof loading of a structure

The product is stressed far beyond its specifications. The stress test can be set up to find the destruct limits. For instance Motorola's Process for stress testing of their mobile telephones involves:

- Drop Testing (samples of the phone are dropped on different surfaces multiple times);
- Hot and Cold Temperatures (the phones receive extreme temperatures of hot and cold; testing simultaneously with audio quality and simulated calls);
- Electro-Static Discharge (zap products with an electric pulse at "discharge" points);
- Dust Exposure (products are placed in a dust chamber and dust is shaken all over the product and then left to settle).
- Vibration Testing (horizontal and vertical vibration testing; placed into vibration table and controlled by tester; the tests vary in the extremes of vibrations given; during the test it places a call and must be stable throughout the process).
- Environmental Ageing Tests (temperature cycling: moving from low to high heat and cold, as well as temperature shocks).
- Certification Testing (Sudden Impact Testing, Endurance Testing, Package Drop Testing, Transport Simulation).

Stress tests in the field of (electro-) mechanical engineering usually include testing of hardware (Central processing unit and memory) and software. For hardware that means to determine the "extremes of workload, type of task, memory use, thermal load (heat), clock speed, or voltages". In software, stress test focuses on "robustness, availability, and reliability under extreme conditions" (Meier et al., 2007).

In engineering a wide spectrum of experimental methods (destructive or non-destructive) are used to determine the condition of structures. Destructive methods (often known in the mechanical- and car industry as crash tests) are carried out until failure of element/structure by given different loads on the element/structure and recording the responses on cameras or with other equipment. Non-destructive methods are used to find the breaking point without destroying. Such methods include ultrasonic, acoustic (emission and resonance), electromagnetic, magnetic-particle, liquid penetrant, infrared, radiographic, remote visual inspection eddy-current testing, low coherence interferometry. (See http://www.engineeringtoolbox.com/ndt-non-destructive-testing-d_314.html).

Mathematical approaches can also be used to generate stress test results. Stress testing in the electromechnical sector cuts costs. There is less return of products that fail for the customer. However, there is no clear guidance on the stress levels. A probabilistic approach as shown in Section 2 and 3 of this report, is strongly suggested by WP6 of INFRARISK.

Within the civil engineering sector and in particular bridge engineering, a nice overview of the different types of testing has been published in the ARCHES project, Figure 6.1.

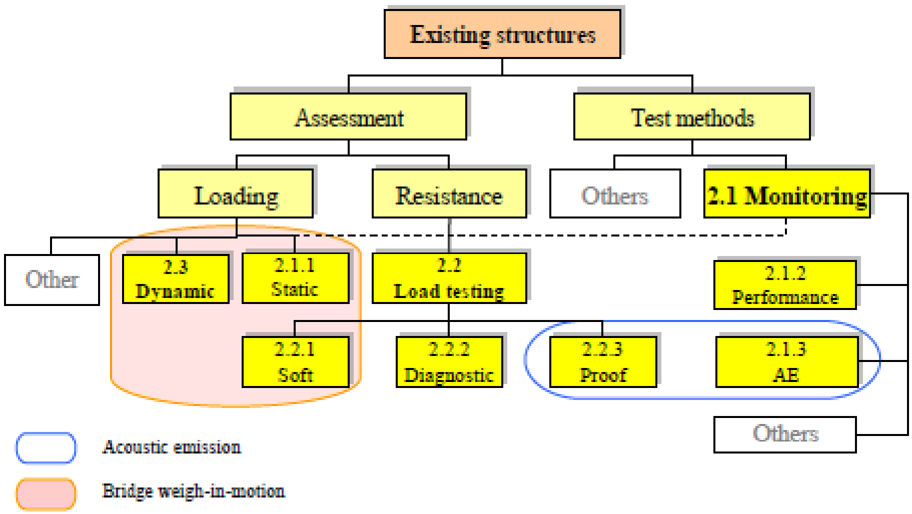

Figure 6.1: Different techniques for assessment of civil engineering structures, from ARCHES, 2009 (AE=Acoustic emission)

Soft load testing of bridges is a sub-type of diagnostic load testing that uses responses from the normal traffic to fine-tune the structural model of the structure.

In diagnostic testing of bridges, the selected load is placed at designated locations on the bridge and the effects of this load on individual members of the bridge are measured by the instrumentation attached to these members. The resulting field measured effects are then compared to effects computed based on the applied loading and standard engineering analysis principles and practices.

Proof load testing is a full-scale and non-destructive examination of high load-carrying capacity.

Acoustic emission (AE) is a useful technique in proof load tests in order to stop the load increase before any damage can be inflicted to the bridge. For instance, AE signals can evaluate the cracking limits without introducing any significant damage to the girders.

The necessity of the assessment of the safety of structures (such as bridges) is due to the fact that all structures face Ageing, have damage and sometimes are not well enough maintained. Other reasons could be changes in the use or increasing traffic loads. Analytical or experimental methods can be used for their safety evaluation (Wí sniewski, 2007). Analytical methods include probabilistic and deterministic approaches. Experimental method may consist of diagnostic tests and stress- or proof tests.

The experimental method consists of diagnostic tests and proof tests. Diagnostic methods are non-destructive methods such as:

- Ultrasonic Pulse Velocity
- Rebound Hammer Test
- Profometer (for detection of rebar locations and measurement of the concrete cover and rebar diameters)

- Electrochemical Half-cell Potentiometer test
- Carbonation Test
- Concrete Core extraction and Testing
- GPR - Ground Penetrating Radar (Geophysical Services)

The proof load test is useful when the documentation about design and properties of the bridge is absent or inadequate, detailed inspection of the bridge difficult and the analytical methods do not give an accurate result (Casas and Gómez, 2013). During this test the actual loads are applied to the bridge in a few stages. In the first stage these authors propose that the load should not be higher than 25% of the total target proof load. The response of the bridge should be carefully monitored in order to avoid permanent damage or failure.

The paper Proof Testing of Reinforced Concrete Bridges; A Useful Way to Provide Load Rating by Bernhardt et al. (in TECHBriefs 2004 No. 3), argues that proof testing can be a safe, scientific method of determining bridge capacity when traditional methods of determining that capacity are not available.

Figure 6.2 and Figure 6.3 illustrates proof load testing of bridges.

Figure 6.2: Bulldozer on low-boy trailer during proof testing (from Bernhardt et al., 2004)

Figure 6.3: Heavily overloaded trucks crossing a damaged bridge (ARCHES, 2009)

Task 6.2 of INFRARISK will further examine and develop a framework for design and optimization of stress tests in infrastructural systems.

Recently, stress testing under climate changing conditions has become an important issue in different fields. The summit on Climate Risk at the UN Headquarters at New York asked oil and gas companies for stress tests of their input into the greenhouse gas emissions.

A stress test under climate change could be defined as "an exercise to assess the likely effects on particular countries, populations, or systems of potentially disruptive climate events to which they have some likelihood of exposure in the coming decade" (Steinbruner et al., 2013). These authors suggest to consider two kinds of events: 1) the event which is evidential and probability of occurrence or scale of the event is increasing; and 2) potential events which are not evidential but possible according to the theoretical knowledge.

Swart et al. (2013) suggest a broader definition of the stress test as "the evaluation of the level of climate resilience offered by current strategies and policies, and the design of effective improvements and new transformative solutions", Figure 6.4.

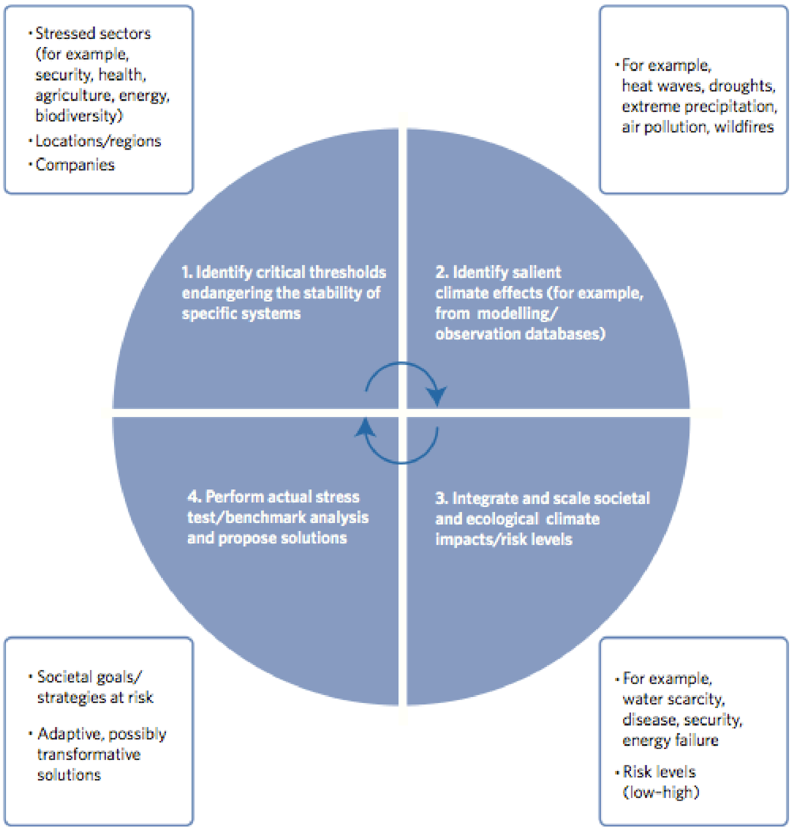

Figure 6.4: Model of stress test (from Swart et al. (2013)

Stress tests of roadway/railway could consider the behaviour of material and structure under forces and loads (conditions of road profile, capacity of network, and other technical factors) but also the human factor (behaviour of drivers). The work of Li et al. (2009) applies a stress test to determine the limit of the roadway network capacity. The average saturation level and average speed of a roadway network were chosen as an evaluation index of roadway operations. The modelling of the transportation system was done in the software package Cube (Base and System). To study the performance of the stress test, the traffic loads were increased by using a regression model (network flow as function of time). The Huairo district (Beijing, China), divided into 12 traffic zones, was chosen as a case study. The operational model consists of 5 modules: network analysis; trip production; traffic distribution; traffic assignment and analysis module (for graphical display of results). Traffic flow was increased by using a linear regression equation until the saturation level of roadway network became 0.6 and the average speed of roadway network was 27.4 km/h. At this level the roadway network was deemed to have reached the capacity at which it was still operated normally. The total capacity therefore was 104 923 pcu/h according to the Figure 6.5.

Time of test	Traffic flow of road network (pcu/h)	Average saturation of road network	Average speed of road network (km/h)
1	46297	0.28	49.58
5	52020	0.31	47.06
10	60845	0.34	45.29
15	69670	0.41	45.97
20	78495	0.46	39.05
25	87320	0.50	35.67
30	96145	0.55	30.08
35	104923	0.60	27.41

Figure 6.5: Result of the stress test of Huairo roadway network (from Li et al. (2009))

The authors concluded that their method can be successfully used to determine the limitation of network capacity.

Mercier et al. (2013) investigate mobility behaviour in response to different mobility shocks. Different householder's groups in France (Lyon) and Germany (Munich) were analysed. For Lyon these were: family with two children and family without children from a suburban area, single student and retired couple from the city centre. For Munich these were: single man from the city centre, middle aged couple from outskirts, family with two children from suburban area. Regional databases helped to determine the housing, activity locations and costs.

The stress test includes three scenarios: Crude oil price at a level of $200/barrel; Fuel prices at the petrol station triples; Oil shortage and rationing of fossil energy resources. It was found out that the first scenario did not affect much the activity and mobility behaviour (Mercier et al., 2013). The second scenario had more impact: for the four-person family living on a suburban area in Lyon, the housing and mobility costs increased by 51% (67% of family's income); for a family living in the Munich metropolitan region this increase is 17% and costs take 78% of the family's income. However, the changes in the mobility behaviour (public transport, car-pooling, etc.) can sufficiently reduce the costs. The third scenario (oil shortage and rationing of fossil energy resources) has the most dramatic effect. For example, the four-person family living on a suburb in Lyon should limit their travel distance to 161 km per month (from 2566 km in the base situation) and apply other different strategies to reduce costs. Therefore, stress test is a useful tool to analyse the resilience of householders on mobility shocks.

For the domain of flood defences, a stress test could consist of a test evaluating whether measures on flood defences are actually required by loading it "extremely". If there are doubts about the strength (caused by uncertainty in data and/or models for example) a real stress test can be performed. The intended embankment is dammed and screened with (above) normative hydraulic loads. On the basis of these data, the current strength and behaviour of the flood defence is analysed, and it is determined whether reinforcements are really needed, or at lower costs and efficiency.

Also stress tests on the flood prone area can be designed, in which the following issues can be investigated:

- Looking at the environment in a dike reinforcement
- Combining space defences and disaster management in achieving safety
- The functioning of sensitive and critical infrastructure in the event of a flood

Examples include:

- Wave Overtopping Simulator, an instrument to test the strength of real dikes for overtopping waves.
- Wave Impact Generator, a machine which enables real testing of river dikes against wave impacts.
- Wave Run-up Simulator, to test the strength of grassed slopes of sea dikes in the wave run-up zone.

For and on site test see: http://www.youtube.com/watch?v=Ldf2MS2RbXY

7.0 STRESS TESTS IN NUCLEAR ENGINEERING

Stress tests in this domain include 2 tracks:

1. Safety track:

The safety track covers such extreme natural events as earthquakes, floods, extreme cold, extreme heat, snow, ice, storms, tornados, heavy rain; consequences of prolonged loss of electric power and/or loss of the ultimate heat sink; and severe accident management as reported in the technical summary on the implementation of comprehensive risk and safety assessments of nuclear power plants in the European Union (EU, 2012). Earthquakes and floods were particularly in focus after the March 2011 accident at Fukushima.

2. Security track:

The security track covers security treats (such events as terror attacks), but the publications on this track are classified and not open for public.

Seismic hazard corresponds to maximum horizontal peak ground acceleration in 10 000 years and minimum of 0.1g according to International Atomic Energy Agency guidelines. The result of an EU stress test showed that the design of nuclear power plants in some countries did not consider this type of hazard (Technical summary on the implementation of comprehensive risk and safety assessments of nuclear power plants in the European Union, 2012):

"In France, no probabilistic seismic hazard assessment (PSHA) is used except for 3 plants…"

If they do consider the seismic hazard, the exceedance probability could be different: In Romania it is 10^{-3} / year; in Germany it is 10^{-5}/year.

The size of the region around nuclear power plants which is deemed to be affected by the earthquake is also different. According to IAEA Safety Standards Series (2010), it is typically 300 km.

The methodologies for the seismic hazard assessment of a nuclear power plants includes geological, geotechnical and hydrogeological investigations, creation of prehistoric, historical and instrumental earthquake databases, development of seismic models, probabilistic and deterministic analyses. In Greenpeace (2012) it was pointed out that the most attention was focused on the impact on structure and components due to vibration and less to the indirect impacts as "damage of non-seismically qualified buildings, fires or flooding of corridors".

The floods hazard are associated with rain and snowmelt flooding, flooding due to high tide, storm surge, seiche (standing wave in an enclosed or partially enclosed body of water), wind waves, sudden release of impounded water, bores and mechanically induced waves, high ground water levels or combinations of these factors. The corresponding exceedance probability of 10^{-4} / year as a minimum was recommended (EU, 2012). According to the report, some countries did not fulfil the requirement: "In Belgium, the Tihange site is currently protected by its design against a reference flood with a statistical return period up to 400 years. In the Netherlands, the Borssele site is protected against flooding by the network of dikes in Zeeland. This network will be improved to comply with the legal requirements of 4000 year return period" (EU, 2012).

The impact of flooding which should be considered in the EU stress test is: failure of power supply system, communication and transport network, damage to the structure and foundation, dispersion of radioactive material, etc. Probabilistic, deterministic methods or both are used to analyse this hazard.

The assessment of consequences of loss of safety functions is relevant also if the situation is provoked by events such as malevolent acts (even if these initiating factors are not studied as such in the course of a complementary reassessment). The review of the severe accident management issues focuses on the operator’s provisions but it may also comprise relevant planned off-site support for maintaining the safety functions of the plant. Although the experience feedback from the Fukushima accident may include the emergency preparedness measures managed by the relevant off-site services for public protection (fire-fighters, police, health services….), this topic is outside the scope of these stress tests.

The risk of airplane crashes on nuclear power plants is not considered in the EU stress tests. Also degradation effects because of Ageing and/or material fatigue and human failure is not considered in the EU stress tests, for unknown reasons.

Furthermore, the following aspects need to be reported in the stress reports:

Provisions taken in the design basis of the plant and plant conformance to its design requirements;

Robustness of the plant beyond its design basis - For this purpose, the robustness (available design margins, diversity, redundancy, structural protection, physical separation, etc) of the safety-relevant systems, structures and components and the effectiveness of the defence-in-depth concept have to be assessed. Regarding the robustness of the installations and measures, one focus of the review is on identification of a step change in the event sequence (cliff edge effect) and, if necessary, consideration of measures for its avoidance. Defence in depth denotes the practice of having multiple, redundant, and independent layers of safety systems for the single, critical point of failure: the reactor core. The aim is to reduce the risk that a single failure of a critical system could cause a core meltdown or a catastrophic failure of reactor containment. With the 2011 Japanese Fukushima nuclear disaster, all of those lines of defence collapsed.

Any potential for modifications likely to improve the considered level of defence-in-depth, in terms of improving the resistance of components or of strengthening the independence with other levels of defence.

In addition, the operator may wish to describe protective measures aimed at avoiding the extreme scenarios that are envisaged in the stress tests in order to provide context for the stress tests. For this aim, the operator shall identify:

The means to maintain the three fundamental safety functions (control of reactivity or preventing the risk of criticality, fuel cooling and heat removal, confinement of radioactivity) and support functions (power supply, cooling through ultimate heat sink), taking into account the probable damage done by the initiating event and any means not credited in the safety demonstration;

Possibility of mobile external means and the conditions of their use;

Any existing rescue procedure for the installation using means from another installation.

As for severe accident management, the operator shall identify, where relevant:

The time before damage to the fuel. If the core is in the reactor vessel, indicate the time before the water level reaches the top of the core, and the time before the fuel degradation occurs (for instance, fast cladding oxidation with hydrogen production);

For research reactors, specify the risk of critical excursion and the means to control this;

If the fuel is stored in the spent fuel pool, the time before pool boiling, the time for which adequate shielding against radiation is maintained, the time before the water level reaches the top of the fuel elements and the time before fuel degradation starts;

For research reactors in which the fuel is in dry storage, the time for which adequate shielding against radiation is maintained and the time before fuel degradation starts.

8.0 SURVIVAL ANALYSIS AND STRESS TESTS

A 'free of charge' stress test is one based on survival. If a structure has survived a certain load in the field, it has shown its strength. This principle can be very well illustrated based on demographic data. Figure 8.1 shows the probability of dying before a certain age for white people in the USA in 2003. The data is retrieved from a CDC/NCHS report at :

http://www.cdc.gov/nchs/data/nvsr/nvsr54/nvsr54_14.pdf

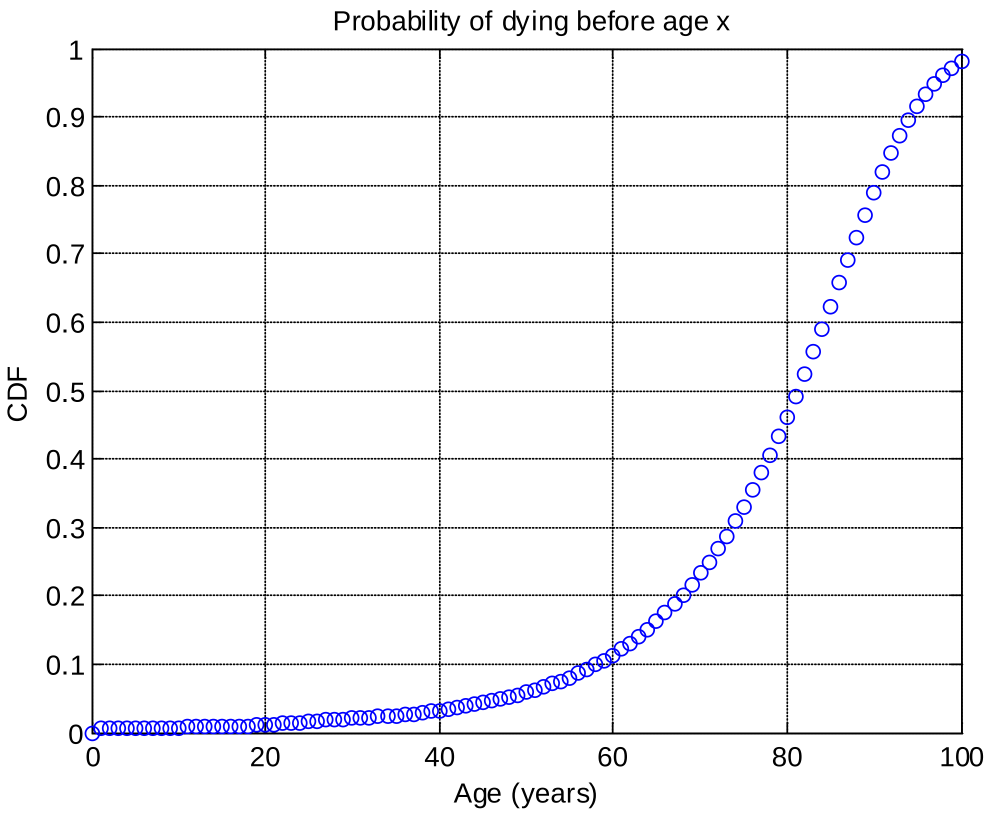

Figure 8.1: Cumulative distribution function based on life tables USA (2003)

From the cumulative distribution, the hazard rate can be immediately derived from the equations of Section 3.3 (dividing the probability density by the complementary cumulative distribution). This results in Figure 8.2, which shows the probability of surviving the next year. This figure can also be plotted on semi-logarithmic paper, Figure 8.3, in which the minimum rate is obtained at the age of 10. From the age of 30 the hazard rate increases almost linearly over time

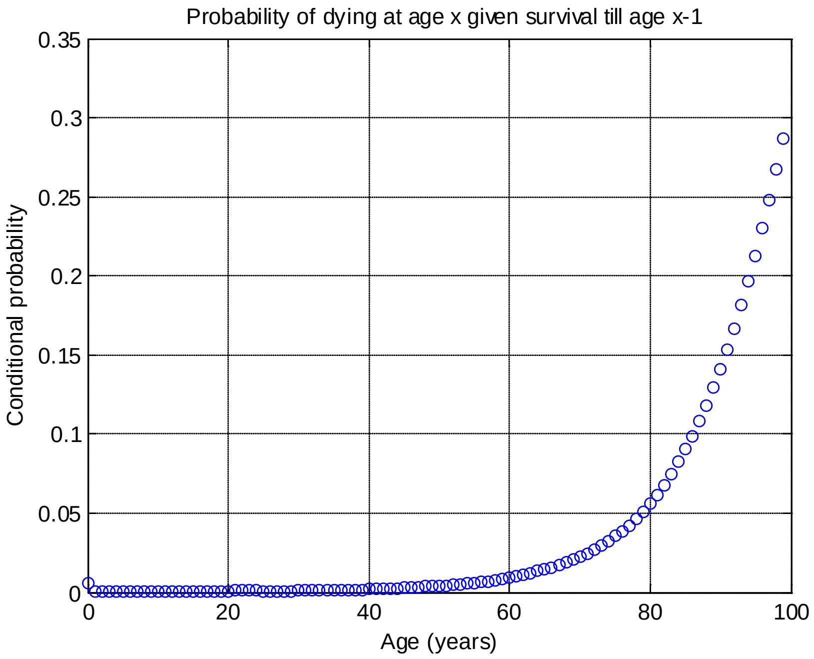

Figure 8.2: Conditional probability distribution

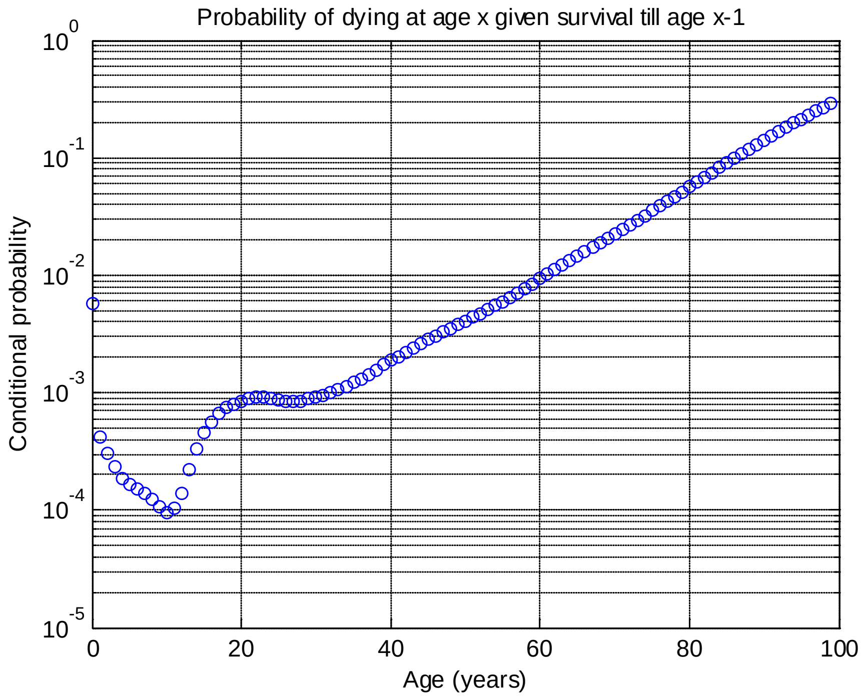

Figure 8.3: Conditional probability on semilogarithmic paper

The probability density function is shown in Figure 8.4. The highest probability is dying at the age of 85 years old (the mode of the distribution).

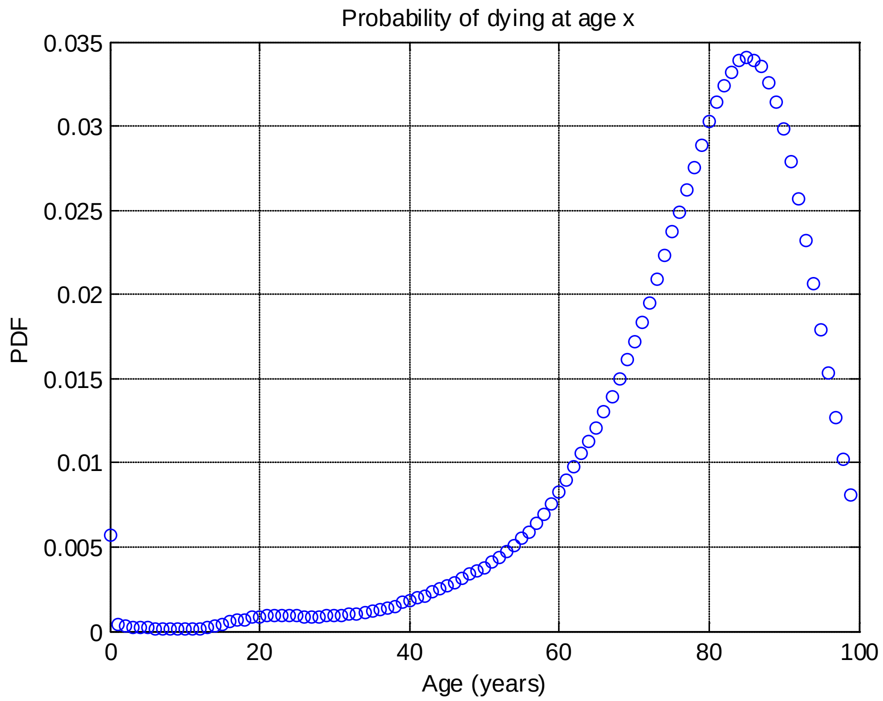

Figure 8.4: Probability density function

The last figure, Figure 8.5, shows the ‘proven strength’ principle. The probability of dying before the age of 80 is 47% (the blue circles). This is an unconditional probability ‘at birth’. Given the information that somebody has survived till the age of 79 years old, the probability of dying before the age of 80 can be updated to 5% (the green crosses). The ‘information advantage’ decreases over time and is negligible about 20 years later, as can be seen by the convergence of the green and blue lines.

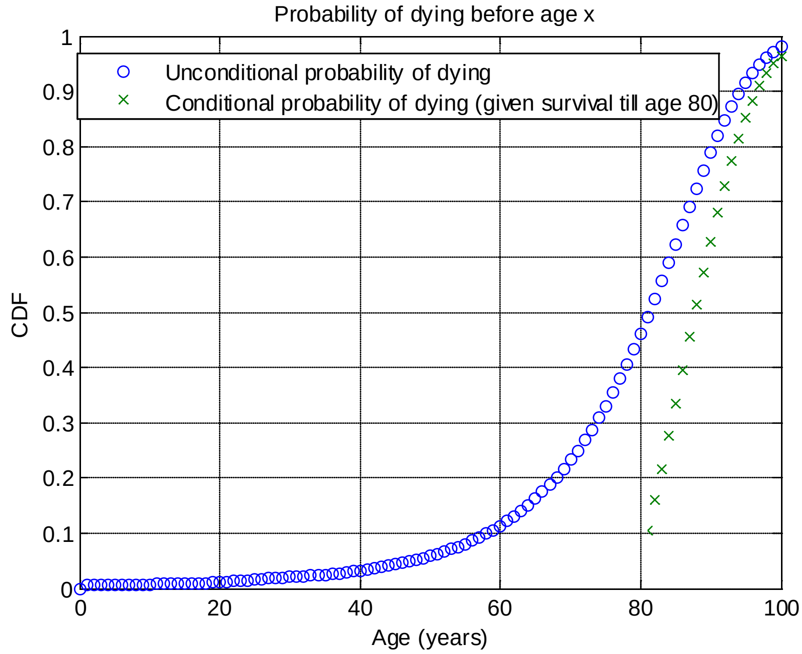

Figure 8.5: Probability density function

9.0 CONCLUSION: A VIEW TOWARDS STRESS TESTING IN INFRASTRUCTURAL SYSTEMS

This report has reviewed some stress test methodologies in other domains (financial industry, mechanical industry and nuclear power industry) and has proposed a general methodology to model stress tests. The essential part in this methodology is the uncertainty reduction in the failure probability of the system. This uncertainty can be reduced by testing (ranging from soft load testing, diagnostic testing to proof load testing) or by simply observing the survival of the system over time

The modelling approach of sections 3 and 4 in this report is recommended to be adopted for infrastructures. In task 6.2, the framework will be further developed in order to design and optimize stress tests in infrastructural systems, taking account for multi-risk scenario's and temporal and spatial uncertainties. Infrastructural systems will be modelled as fault tree systems with a large number of (possibly dependent) components. Elevated stress levels can be applied to individual components and/or a combination of components. Different temporal and spatial correlation distances can lead to different stress levels needed for the components within the system, for which a quantification model is needed.

Decision-making after stress tests observations become available can be based on a cost benefit analysis. A small uncertainty in the failure probability represents a value-of-information, which can be monetised. It depends on the possible consequences of failure. Task 6.3 will develop methods and techniques how to proceed after a stress test would result in a shift of the failure probability to higher values. It will investigate if immediate safety measures need to be taken to improve the situation (and to which extent) or if actions can be delayed. If the stress tests results are slightly negative or just in the 'green zone', alternative actions can be recommended. Actions can be categorized according to an optimal ranking strategy.

10.0 REFERENCES

Basu, A.P., and Ebrahimi, N. (1991). Bayesian approach to life testing and reliability estimation using asymmetric loss function. *Journal of Statistical Planning and Inference*, 29, 21-31.

Basu, A.P., and Thompson, R.D. (1992). Life testing and reliability under asymmetric loss. In J.P. Klein, editor, *Survival Analysis - State of the Art*, pages 1-7, Dordrecht: Kluwer Academic Publishers.

Casas J.R., and Juan D. Gómez. Load Rating of Highway Bridges by Proof-loading. KSCE *Journal of Civil Engineering*, 2013, 17(3):556-567

Casas, Joan R. , Piotr Olaszek, Aljońa Ńajna, Aleń Ņnidarič, Igor Lavrič, Recommendations on the use of soft, diagnostic and proof load testing, Report *ARCHES Assessment and Rehabilitation of Central European Highway Structures*, August 2009

CFGS, Stress Testing by Large Financial Institutions: Current Practice and Aggregation Issues, *CGFS Publications No 14*, April 2000

Committee on the Global Financial System. Stress testing at major financial institutions: *survey results and practice*, 2005.

Cooke, Roger M.. Experts in uncertainty; opinion and subjective probability in science. Publisher *New York : Ox-ford University Press*, 1991 ISBN 0-19-506465-8.

EU, 2012. Technical summary on the implementation of comprehensive risk and safety assessments of nuclear power plants in the European Union, Oct. 4th, 2012.

Foglia A. Stress Testing Credit Risk: A Survey of Authorities' Approaches. I*nternational Journal of Central Banking*. September Issue. 2009.

Greenpeace, 2012. The Critical review of the EU Stress Test performed on Nuclear Power Plants. Report Greenpeace.

Hoeg, K., 1996, Performance evaluation, safety assessment and risk analysis for dams, *International journal on hydropower and dams*, Issue 6, Vol. 3, November 1996.

IAEA 2010, INTERNATIONAL ATOMIC ENERGY AGENCY, Seismic Hazards in Site Evaluation for Nuclear Installations, *IAEA Safety Standards Series No. SSG-9*, IAEA, Vienna (2010).

Langley, R.S., 1987, Techniques for assessing the lifetime reliability of engineering structures subjected to stochastic loads, *Engineering structures*, Vol. 9, April 1997.

Licari J.M., Suárez-Lledó J. Reverse Stress Testing from a Macroecomic Viewpoint: Quantitative Challenges & Solutions for its Practical Implementation. *Moody's Analytic: Economic and Consume Credit analytics*. 2012, www.economy.com

Li L., Quan Y., Yang B. Stress test of roadway network capacity, *World Congress on Computer Science and Information Engeneering*, 2009, pp.467-471.

Mercier, A.; Crozet, Y.; Ovtracht, N.; Buettner, B.; Wulfhorst, G. Stress tests on urban mobility: lessons for public policies. *13th WCTR*, July 15-18, Rio de Janeiro, Brazil, 2013, 26 p.

Meier J.D., Farre C., Bansode P., Barber S., Rea D. Performance Testing Guidance for Web Applications.. Microsoft Corporation, 2007

Paté-Cornell, M.E. 1996. Uncertainties in risk analysis; Six levels of treatment. *Reliability Engineering and System Safety. Nr. 54*. Northern Ireland: Elsevier Science Limited.

Quagliariello, M., Stress-testing the Banking System: Methodologies and Applications. Edited by Mario Quagliariello, Cambridge University Press, 2009.

Steinbruner J.D., Stern P.S., and Husbandset J.L. Climate and Social Stress: Implications for Security Analysis, *The National Academies Press*, Washington, 2013

Swart R., Fuss S., Obersteiner M., Ruti P., Teichmann C., Vautard R. Beyond vulnerability assessment. *Nature Climate Change 3*, 2013, pp. 942–943.

Wenisch A., Becker O. Critical Review of the EU Stress Test performed on Nuclear Power Plants. Study commissioned by Greenpeace, 2012.

Wí sniewski, Dawid F., Safety Formats for the Assessment of Concrete Bridges with special focus on precast concrete. Doctoral thesis, Department of Civil Engineering School of Engineering - University of Minho Guimar˜aes, Portugal, 2007

APPENDIX A: Simple example of destructive stress testing

The concept of stress testing can be very well explained based on the following paper clip example, Figure A.1

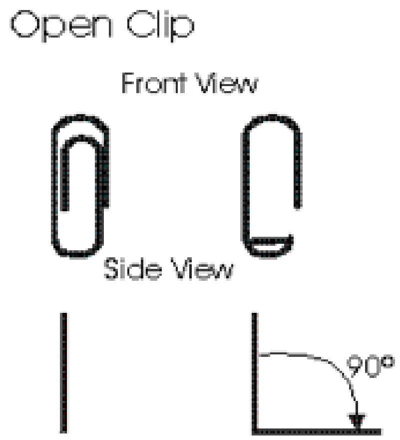

Figure A.1: Example to illustrate the methodology: Paper Clip Life Testing

Normally, a paper clip is bent over a very small angle (let's say 1 degree). We are interested in the lifetime of a paper clip (the number of cycles until failure). A (destructive) stress test would count the number of cycles over a large angle until failure, over many runs, Figure A.2

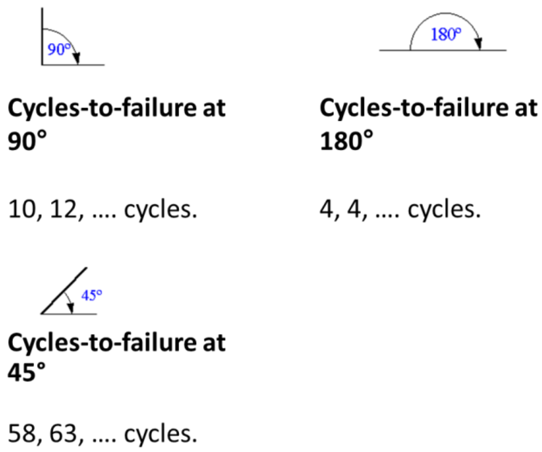

Figure A.2: Number of cycles until failure of paper clip under extreme loading

The observed number of cycles during extreme loading can be modelled by so-called Weibull distributions in which one of the Weibull parameters is modelled as a function of the applied angle (according to an inverse power law). In this way, the number of cycles during normal loading (bending over small angles) can be predicted by extrapolating the observations at high stress levels to low stress levels. Statistical fitting and extrapolation are therefore the main tools in a stress test analysis.