Processing Risk
in
Asset Management

Exploring The Boundaries Of Risk Based Optimization Under Uncertainty For An Energy Infrastructure Asset Manager

Ype Wijnia

PROCESSING RISK IN ASSET MANAGEMENT

EXPLORING THE BOUNDARIES OF RISK BASED OPTIMIZATION UNDER UNCERTAINTY FOR AN ENERGY INFRASTRUCTURE ASSET MANAGER

Ype Wijnia

PROCESSING RISK IN ASSET MANAGEMENT

EXPLORING THE BOUNDARIES OF RISK BASED OPTIMIZATION UNDER UNCERTAINTY FOR AN ENERGY INFRASTRUCTURE ASSET MANAGER

PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Technische Universiteit Delft,

op gezag van de Rector Magnificus prof.ir. K.C.A.M. Luyben;

voorzitter van het College voor Promoties,

in het openbaar te verdedigen op

19 april 2016 om 15.00 uur

door

Ype Cornelis WIJNIA

Bestuurskundig Ingenieur,

Technische Universiteit Delft, Nederland

geboren te Alphen aan den Rijn, Nederland

Dit proefschrift is goedgekeurd door:

promotor: Prof.dr.ir. P.M. Herder

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof.dr.ir. P.M. Herder	Technische Universiteit Delft

Onafhankelijke leden:

Prof.dr. J.P. Liyanage	University of Stavanger, Norway
Prof.dr.ir. P.H.A.J.M. van Gelder	Technische Universiteit Delft
Prof.dr.ir. M.P.C. Weijnen	Technische Universiteit Delft
Prof.dr. J.J. Smit	Technische Universiteit Delft
Dr. K. Komonen	Finnish Maintenance Society Promaint, Finland
Ir. J.C.F.M. Peters	Enexis B.V., Nederland

ISBN 978-94-028-0149-1

Keywords: Asset Management, Risk Management, Optimization, Uncertainty, Energy distribution

Cover design by Ype Wijnia

Photographs on front page by Enexis B.V.

Printed in the Netherlands by Ipskamp Printing, Enschede

Contents

List of figures

List of tables

Summary

In 1999, the Dutch energy market was liberalized. The operation of the distribution grids, which were regarded as a natural monopoly, had to be unbundled from production and sales. To safeguard a fair market, a regulator was established. One of the tasks of the regulator was to determine the income of the Distribution Network Operators (DNO). In the United Kingdom, which preceded the Netherlands in liberalization, this resulted in a significant reduction of income. The “solution” many UK companies implemented to deal with this challenge was asset management. In those early days asset management was by no means a coherent concept. However, the idea of balancing cost, risk and performance seemed to be characteristic for asset management. Coherence among companies and practices was only achieved with the formalization in a normalized management system: the PAS 55 specification. Although asset management revolves around improved (optimal) decisions, the actual decision making received very little explicit attention in PAS 55 (and neither in its successor the ISO 55000 series). Within regulation of the Dutch network operators, there was also little standardization of decision making, despite the adoption and modification of PAS 55 into NTA 8120, the Dutch norm for asset management for network operators. Lack of standardization would be understandable for a norm with an intended wide application, but was remarkable within this limited context of the management of a distribution infrastructure. This raised the question whether there is a more fundamental reason for this lack of standardization. And if so, what level of standardization is possible? In other words, what are the boundaries for standardized optimization in asset management?

When the concept of asset management is applied to the infrastructure for energy distribution, most of the management attention is on managing risk. Most risks are regarded as normal risks that can be objectified. This raised the question whether the management of infrastructures for energy distribution could be regarded as cost-benefit-consideration with regard to risk for the whole-system (Risk Based Optimization, RBO). Managing risk, however, is a problematic concept. In literature there is no precise agreement on how to approach this, though the available views (like COSO and ISO 31000) seem to align. Unfortunately, behind the superficial differences a more fundamental conflict is hidden, on what a risk precisely is, and how a good decision should be made about risk. Many different definitions can be used for risk, with most of these recognizing the importance of uncertainty. The most important difference between the definitions is that of risk as a concept and the way this concept is measured, though some limit the concept of risk to the negative domain whereas others also consider positive deviations as (upward) risk.

We have selected Cost Benefit Analysis (CBA) as the appropriate approach for decision making in asset management, after a thorough literature study. However, CBA is not sufficient to achieve a full risk based optimization of the asset base: theoretically it is possible to account for non-financial effects in CBA, but there is no generally accepted scheme for incorporating them into such analysis. Besides, CBA does not provide a means for evaluating risks on their importance as such. If risk evaluation criteria were available, the most important problems could be selected from an overview of problems. There was, however, no agreed method for generating such an overview. It was also not clear whether using CBA for every decision would be accepted. And finally, it was not clear what the added value of RBO would be in an energy distribution infrastructure. A pivotal point in understanding these knowledge gaps was that the gaps are not independent in practice. For

example, a very sophisticated value system may be generally accepted but may also be very difficult to apply. On the other hand, a very simple value system (e.g. only financial) may be easy to apply, but results may not be accepted.

Based on these knowledge gaps the central research question was:

To what extent is formal Risk Based Optimization of the whole system feasible in managing assets of the infrastructure for energy distribution?

This central question was divided into 4 sub questions:

1. What is an adequate representation of the value system that facilitates both CBA as the selection of most important risks?
2. What is an adequate representation of the risk position?
3. What is the effectiveness of applying Risk Based Optimization by means of these adequate representations?
4. How robust is this effectiveness of Risk Based Optimization?

A significant part of the experimental research on these questions was conducted within Enexis, one of the three large Distribution Network Operators in the Netherlands.

Single Asset Optimization: The first experiment concerned the value of RBO like approaches for an individual problem of network capacity expansion. Evaluating several ways to include more risk based considerations into decision making revealed that a drastic change of the decision could be reached by means of postponing the investment with many (10 or more) years. In terms of the Total Cost of Ownership (TCO) the improvement was about 20% compared with the conventional approach. Additionally, it was shown that a very robust decision could be made without fully understanding the precise failure mechanism behind the risk.

Portfolio Optimization: Based on this result, the scope of the research was expanded to the value of RBO in determining a portfolio of interventions. The most important question in this experiment was whether mixing several separate portfolios and prioritizing them on a uniform value system would be accepted, both by the contributing teams as by the ultimate decision maker. We developed an approach that led to a fully accepted portfolio decision. Key in our approach was that the value system was used as a flexible aid in ranking the interventions and not as an ultimate truth by which to measure and judge every individual decision. Also, our approach combined CBA with a well-designed (social) decision process. It resulted in a performance improvement of about 20% compared with an unranked budget allocation.

Risk Position: In order to move from a procedural optimization of the portfolio of interventions to a formal optimization of the whole system, an adequate representation is needed of all the risk in the system, the risk position. Our method to obtain such adequate representation has two key elements: the risk process and its, rather pragmatic, application. The method selected for structuring risk was the use of a risk process, from cause to consequence, where every phase of the risk process could be used as a starting point for identification and clustering. The risk process was then applied to establish the total incident risk of the gas distribution grid. We found that the risk process helped in understanding that only few precursors (=combination of cause and asset) preceded the bulk of the incidents. The risk in the whole system could then be approximated by first establishing the risk per precursor (= average expected impact per occurrence) and then by simply counting the occurrence of precursors.

System Optimization: The risk process was then applied in a formal system optimization of the replacement of ageing assets. Using the risk process, the asset base (consisting of thousands of varieties of assets) was condensed into several tens of asset types. Per asset type a risk profile was established, consisting of an age dependent failure rate, age profile and the failure consequences. Summed over all asset types this gave a prognosis of the performance of the asset base. Then, the optimal replacement age was determined in a cost marginal approach (comparing the cost of advancing replacement one year with the risk reduction in that year). The added value of risk based optimization of the system turned out to be in the 20% region.

Risk register: In a further experiment, all risks identified were combined into a single register, and the total risk was calculated by running a Monte Carlo simulation over the collected risks. The resulting distribution of the predicted performance differed from the measured, empirical performance distribution. This was due to overlap between risks resulting in double counting of effects. A risk register usually contains only a small number of high risks (with a large average impact) and a large number of low risks that show the most overlap (small average impact, orders of magnitude below the high risks). We found that when the most relevant risks are modeled adequately, the less important risks do not matter that much, even if they would have significant overlap. This meant that some overlap does not have to be resolved (at high cost) in a risk register. Another key finding was that the specific definition of the concept of risk was not needed to adequately model the risk position.

Value system: An adequate representation of the value system proved to be rather straightforward. For the majority of risks under consideration it proved to be sufficient to establish the impact on three values: costs, safety and reliability. The expected amount of misery (probability times impact, the exposure) proved to be an appropriate measure for the importance of the risk, especially if the risk level was measured on a logarithmic scale. If the values are aligned (the decision maker is indifferent between impacts of the same severity on different values) then non-financial effects could be substituted by their financial counterpart. Another key finding was that decisions were not very sensitive to the exact monetization factor of non-financial values in the risk matrix.

Our main conclusions from the theoretical study and our experiments are:

1. A value system for normal risks can be adequately represented by a properly designed risk matrix. It can be used both for prioritizing risks as well as for cost benefit analysis by means of a monetization factor.
2. The risk process helps in structuring the risk position into a limited number of risks that provide an adequate representation of the risk position, i.e. the total value at risk.
3. The effectiveness of applying Risk Based Optimization, in a well-designed social decision making process, is high. It reduces the total cost of ownership of the assets by about 20%.
4. The effectiveness of RBO is robust under variable representations of value systems and risk registers. In practice, only a fraction of the RBO outcomes were rejected by the decision makers. This justifies our relatively pragmatic RBO as the cost of detailing and improving it even further would not be compensated by overall improved decision making and outcomes.

A formal risk based optimization of the whole system is feasible to a very large extent in managing assets of infrastructure for energy distribution.

The acceptance of RBO was reached in a system with largely normal risks. The results can therefore not be generalized without further consideration to other infrastructures or to different industries, which possibly have more non-normal risks. A relevant question for those systems is whether RBO could be applied for the normal part of the risks only, or that it would be better to switch to a completely different approach. Another relevant issue for further research is the independence of risks. In more tightly coupled systems than the ones studied in this thesis, risks may have to be modelled in an integrated approach.

Samenvatting

In 1999 werd de Nederlandse energiemarkt geliberaliseerd. Het beheer van het netwerk, dat als een natuurlijk monopolie werd beschouwd, moest hierbij onafhankelijk worden van productie en verkoop. Om misbruik van de machtspositie te voorkomen werd een toezichthouder ingesteld. Een van de taken was het vaststellen van het inkomen voor de netwerkbedrijven. In het Verenigd Koninkrijk, dat Nederland was voorgegaan in de liberalisering, had dit geleid tot een sterke daling van de inkomsten. De “oplossing” die veel Engelse bedrijven hadden geïmplementeerd om met de inkomstendaling om te gaan was asset management. In deze tijd was asset management nog zeker geen samenhangend concept, maar het idee van het balanceren van kosten, risico en prestatie leek karakteristiek voor asset management. Samenhang tussen bedrijven en praktijken kwam pas met de formalisatie in een genormaliseerd management systeem, PAS 55. Alhoewel asset management draait om het nemen van de betere (optimale) beslissingen, krijgt besluitvorming nauwelijks expliciete aandacht in PAS 55 (en ook niet in de opvolger de ISO 55000 serie). Ook binnen de regulering van de netwerken in Nederland heeft er nauwelijks standaardisatie van besluitvorming plaatsgevonden, ondanks dat PAS 55 grotendeels is overgenomen in NTA 8120, de Nederlandse norm voor asset management voor netbeheerders. Dit gebrek aan standaardisatie is begrijpelijk voor een norm met een brede beoogde toepasbaarheid, maar opvallend binnen deze beperkte context. Dit riep de vraag op of er wellicht fundamentele redenen zijn waarom dit niet plaats vindt. En zo ja, tot welk niveau kan er dan wel gestandaardiseerd worden? Met andere woorden, wat zijn de grenzen voor gestandaardiseerde optimalisatie binnen asset management?

Wanneer het concept van asset management wordt toegepast op de infrastructuren voor energiedistributie dan gaat de meeste aandacht uit naar het beheersen van risico. De meeste risico’s worden beschouwd als normale risico’s die geobjectiveerd kunnen worden. Dit riep de vraag op of het management van infrastructuren voor energiedistributie beschouwd kan worden als een “whole system” kosten-baten afweging met betrekking tot risico (Risico geBaseerde Optimalisatie, RBO).

Echter, het beheersen van risico is een problematisch concept. In de literatuur is geen precieze overeenstemming te vinden over hoe je dit aanpakt, al zijn de beschikbare visies (zoals bijv COSO of ISO 31000) op hoofdlijnen zeer sterk vergelijkbaar. Helaas ligt achter de oppervlakkige verschillen een fundamenteler conflict verborgen, namelijk de vraag van wat risico precies is en hoe je een goed besluit neemt over risico. Er zijn vele risico definities in omloop, waarbij de meeste definities het belang van onzekerheid erkennen. Het belangrijkste onderscheid tussen de definities betreft risico als concept en hoe het wordt gemeten, al beperken sommigen het tot het negatieve terwijl anderen ook positieve afwijkingen als (opwaarts) risico beschouwen.

Wij hebben Kosten Baten Analyse als de geschikte methode voor besluitvorming geselecteerd op basis van een uitgebreide literatuurstudie. Echter, KBA alleen is niet voldoende om tot een volledige risico gebaseerde optimalisatie te komen: theoretisch is het mogelijk ook niet-financiële effecten mee te nemen in de afweging, maar hiervoor bestond geen algemeen geaccepteerd schema. Ook geeft KBA geen antwoord op de vraag voor welke problemen een beslissingen genomen moet worden. Middels risico-evaluatiecriteria kunnen de belangrijkste problemen uit een volledig overzicht geselecteerd worden, maar helaas was er geen methode beschikbaar om zo’n volledig overzicht te maken. Ook was nog niet duidelijk of het onverkort toepassen van KBA niet tot

acceptatieproblemen leidt. Tot slot was ook niet helder welke toegevoegde waarde RBO zou hebben indien het volledig zou worden toegepast in de energiedistributie infrastructuur. Een kantelpunt in het begrip van deze kennislacunes is dat de lacunes niet onafhankelijk zijn in de praktijk. Een zeer geavanceerde weergave van het waardesysteem kan bijvoorbeeld algemeen geaccepteerd worden maar ook zeer moeilijk toe te passen zijn. Aan de andere kant, een zeer eenvoudig waardesysteem (met bijvoorbeeld alleen maar financiële aspecten) is wellicht makkelijk toe te passen, maar wordt mogelijk niet door iedereen geaccepteerd.

Op basis van deze 4 kennislacunes was de centrale onderzoeksvraag:

In hoeverre is formele risico gebaseerde optimalisatie van het gehele systeem mogelijk in het managen van de assets van de infrastructuren voor energiedistributie?

Deze onderzoeksvraag is vertaald in 4 deelvragen:

1. Wat is een adequate manier om het waarde systeem te representeren, waarmee zowel KBA als selectie van belangrijkste risico's mogelijk is?
2. Wat is een adequate representatie van de risicopositie?
3. Wat is de effectiviteit van het toepassen van Risico geBaseerde Optimalisatie middels deze adequate representaties?
4. Hoe robuust is deze effectiviteit van Risico geBaseerde Optimalisatie?

Een belangrijk deel van het experimentele onderzoek naar deze vragen is uitgevoerd binnen Enexis, één van de drie grote Nederlandse netbeheerders.

Optimalisatie enkel vraagstuk: Het eerste experiment betrof de waarde van RBO-achtige benaderingen voor een individueel vraagstuk van capaciteitsuitbreiding van het netwerk. Uit de evaluatie van een aantal manieren om meer risico gebaseerde afwegingen te maken bleek dat dit kon resulteren in een drastische verandering van het besluit in de vorm van uitstel van investeringen met vele (10 of meer) jaren. De toegevoegde waarde van een dergelijke optimalisatie ligt in de orde van 20% van de Total Cost Of Ownership (TCO) ten opzichte van een conventionele benadering. Aanvullend werd getoond dat een zeer robuuste beslissing genomen kon worden zonder dat er een volledig begrip van het precieze faalmechanisme achter het risico was.

Optimalisatie van de portfolio: Op basis van dit resultaat is de scope van de vraag uitgebreid tot het bepalen van de waarde van RBO in de vaststelling van een portfolio van interventies. De belangrijkste vraag bij dit experiment was of het mengen van een aantal deelportfolio's met een uniform waardesysteem tot een geaccepteerd eindresultaat zou leiden, zowel bij de samenstellende teams als bij de uiteindelijke beslissers. We ontwikkelden een aanpak die tot een volledige acceptatie van de portfolio beslissing leidde. De sleutel in onze aanpak was dat het waardesysteem werd gebruikt als een flexibel hulpmiddel voor het ranken van de maatregelen en niet als de ultieme waarheid waarmee individuele besluiten genomen moesten worden. Daarnaast combineerde onze aanpak CBA met een goed ontworpen besluitvormingsproces. Het resulteerde in een prestatieverbetering van grofweg 20% vergeleken met een ongeordende budgettoewijzing.

Risicopositie: Om van een procedurele optimalisatie van de portfolio van interventies naar een formele systeemoptimalisatie te komen is een methode voor adequate weergave van het totaal aan risico (de risicopositie) nodig. Onze methode om zo'n adequate weergave te verkrijgen bestond uit twee sleutelelementen: het risicoproces en de pragmatische toepassing daarvan. De gekozen methode voor het structureren van de risico's was het gebruik van het risicoproces, van oorzaak tot gevolg, waarbij elke fase gebruikt kon worden als ankerpunt voor identificatie en clustering. Dit risicoproces is vervolgens toegepast in het vaststellen van het totale incidentrisico van het gasnet. We vonden dat het risicoproces hielp in het inzichtelijk maken dat slechts een klein aantal precursors (=combinatie van oorzaak en asset) vooraf gingen aan de bulk van de incidenten. Het risico in het gehele systeem kon worden benaderd door eerst per precursor het risico (= gemiddeld verwacht effect per optreden) te bepalen en vervolgens het voorkomen van de precursoren te tellen.

Systeemoptimalisatie: Het risicoproces is vervolgens toegepast in een systeemoptimalisatie van de vervanging van verouderende assets. Met het risicoproces kon de gehele assetbase (bestaande uit duizenden verschillende soorten assets) worden ingedikt tot enige tientallen verschillende asset typen. Per type werd een risicoprofiel vastgesteld, bestaande uit een leeftijdsafhankelijke faalkans, een leeftijdsprofiel en de faaleffecten. Opgeteld over alle typen leverde dit een prognose van de prestatie van de asset base. Vervolgens is per type een optimale vervangingsleeftijd bepaald in een marginale benadering (het vergelijken van de kosten van het een jaar vervroegen van de vervanging met de risicoreductie in dat jaar). De toegevoegde waarde van de systeemoptimalisatie was een prestatieverbetering van ongeveer 20%.

Risico register: In een vervolgexperiment werden alle tot dusver geïdentificeerde risico's ondergebracht in één register, en het totale risico werd berekend middels een Monte Carlo simulatie over de verzamelde risico's. De resulterende distributie van de voorspelde prestatie bleek af te wijken van de gemeten, empirische distributie van de prestatie. Dit kwam door overlap tussen de risico's waardoor sommige gevolgen dubbel geteld werden. Een risicoregister bevat normaal gesproken slechts een paar hoge risico's (met een grote gemiddelde impact per jaar) en vele lage risico's (kleine gemiddelde impact per jaar, ordegroottes kleiner dan de hoge risico's) die de meeste overlap geven. We vonden dat wanneer de meest relevante risico's adequaat gemodelleerd zijn, de minder relevante risico's er niet meer zoveel toe doen, ook al hebben ze grote overlap. Dit betekent dat niet alle overlap in het risicoregister opgelost hoeft te worden tegen hoge kosten. Een andere belangrijke constatering was dat een specifieke definitie van het risicoconcept niet nodig was om een adequaat beeld van de risicopositie te verkrijgen.

Waardesysteem: Een adequate weergave van het waardesysteem bleek behoorlijk eenvoudig te zijn. Voor de overgrote meerderheid van risico's bleek het voldoende effecten vast te stellen op drie waarden: kosten, veiligheid en betrouwbaarheid. De verwachte hoeveelheid ellende (kans maal effect) bleek bovendien een geschikte maat voor het risiconiveau, zeker bij gebruik van een op ordegroottes gebaseerde logaritmische schaalverdeling van de risicomatrix. Als bovendien voor uitlijning van de impacts wordt gezorgd (de beslisser is indifferent tussen impacts met een zelfde ernstgraad op verschillende waarden) dan konden niet-financiële effecten gesubstitueerd worden door hun financiële evenknie. Een belangrijke bevinding was de beslissingen niet erg gevoelig waren voor de exacte monetarisering van de niet financiële effecten in de matrix.

Onze belangrijkste conclusies uit de theoretische studie en de experimenten zijn:

1. Het waardesysteem kan adequaat gerepresenteerd worden met een juist ontworpen risicomatrix. Deze kan zowel gebruikt worden voor het prioriteren van risico als voor het maken van een kosten-baten afweging via de monetariseringsfactoren.

2. Het risicoproces helpt in het structureren van de risicopositie in een beperkt aantal risico’s dat een adequate weergave van het totale risico vormen.
3. De effectiviteit van het toepassen van Risico gebaseerde Optimalisatie in een goed ontworpen sociaal besluitvormingsproces is hoog. Er wordt ca 20% op de Total Cost of Ownership bespaard.
4. De effectiviteit is robuust onder variabele representaties van het waardesysteem en de risicopositie. In de praktijk bleek slechts een klein deel van de RBO resultaten verworpen te worden door de beslissers. Dit rechtvaardigt onze relatief pragmatische RBO omdat de kosten van meer details en verdere verbetering niet gecompenseerd zouden worden door een algehele verbetering van besluitvorming en resultaten.

Een formele optimalisatie van het hele systeem is in zeer hoge mate mogelijk in het beheer van de assets van de infrastructuren voor distributie van energie.

De acceptatie van RBO werd bereikt in een systeem met grotendeels normale risico’s. De resultaten kunnen daarom niet zomaar gegeneraliseerd worden naar andere infrastructuren of andere industrieën, die mogelijk meer niet-normale risico’s kennen. Een belangrijke vraag voor dergelijke systemen is of voor het normale deel van de risico’s RBO nog steeds gebruikt kan worden, of dat beter is om voor het geheel op een andere methode over te stappen. Een ander belangrijk onderwerp voor vervolgonderzoek is de onafhankelijkheid van risico’s. In sterker gekoppelde systemen dan bestudeerd voor deze thesis moeten risico’s mogelijk gemodelleerd worden in een geïntegreerde benadering.

Acknowledgements

The research project culminating in this thesis started well over a decade ago, as is witnessed by the experiments. It connected with science in September 2004 when Enexis and Next Generation Infrastructures agreed on a four year, one-day-a-week PhD project. I thank these organizations for the opportunity they provided me for becoming a scientist. A well-known fact to all who have preceded me in proving they can do science, becoming a scientist takes much more than such a small timeframe of effectively 0,8 year. Yet it was enough to acquire a taste for science, at least enough to continue the part time one-day-a-week schedule after the end of the contract on a self-funded basis. I guess completion took a little bit longer than many (including myself) expected. Whether that is a bad thing can be debated. Science is about venturing into unchartered territory, and sometimes it just takes a while for the new discoveries to settle and come to terms with existing knowledge. Science in a way is an unconditional commitment for adding knowledge to the world, independent of the time it takes to do so. Being self-funded then becomes an enormous asset as it removes all deadlines and the associated urges to cut corners.

One of the benefits of being a part time scientist is that it makes one continuously aware that theory and practice are two conflicting views on reality. To name a few obvious differences, science is about asking questions whereas practice is about giving answers, and science (at least in theory) has zero tolerance for imprecisions and errors whereas practice is more about being good enough. However, there is also a much more subtle difference. Science can limit its view on the world to a single consistent perspective in order to create meaning. Practice cannot afford such a luxury, as dealing with different viewpoints is almost unavoidable in order to get results. Unfortunately, being aware of different views does not mean that both views can be assumed at the same time. I guess a blend of views is more likely, resulting in a part time scientist being more precise and analytical than a practitioner and more outlining and solution oriented than a full time scientist. The consequence is that one is always blamed (or complimented?) for belonging to the other group: being regarded as a practitioner in academia, but as a theoretician in practice. This may seem a lonely position, never really belonging to the group in which one operates. However, it can also be considered as a very connected position. Being in between allows challenging results from the other perspective and thus for bridging the gap. How does a scientist know that the results have any relevance in practice if they do not account for many other imprecisions practitioners face, and how do practitioners know that it is good enough if they do not know what is achievable according to theory? I prefer to think I bridged the gap, evidenced by this thesis that proves I am a scientist whilst still working in practice. Yet, this just may be a fantasy. It is also possible that I just crossed the bridge into academia and that I am lost forever for any practical purpose. Only time will resolve this uncertainty.

What is not uncertain though is that this bridging or crossing was only possible because of an exchange of ideas with many people. I cannot thank them enough for providing an opportunity to ask challenging questions and to challenge my results in return, in both views. This includes all colleagues at E&I, WCEAM, EURENSEAM, Enexis and of course my business associates. A special thanks goes to the power rangers, which provided an opportunity for discussion on a very regular basis. A very special thanks goes to the two colleagues at E&I who never got tired of trying to change my perspective into something completely different: Igor Nikolic and Rob Stikkelman. Thank you very much for the valuable discussions we had inside and outside the office.

I would also like to express my gratitude to all people that directly contributed to the experiments described in this thesis. In order of appearance these are my promotor and supervisor Paulien Herder, Joost Warners, Rene Hermkens, Jan Flonk, Martijn Korn and Saskia de Jager. The original publication of the experiments was made possible by Next Generation Infrastructures.

With regard to preparing this thesis, I would like to thank the committee for their useful comments on the draft they received. A very very special thanks goes to my promotor. Paulien, being a part-time scientist inevitably results in a decrease in the rate of progress, but without your prompt and enthusiastic response to any draft thesis I turned in and your support in preparing the next step it would have taken many many more years. I would also like to thank John de Croon for his help in the final edit of this thesis. Without a pair of fresh eyes it is simply impossible to remove the small inconsistencies that creep into a research project that spans more than a decade.

Last but not least I thank my family. Besides their occasional provision of a meal, drink or a listening ear, they taught me the importance of unconditional commitments. Without that, this thesis never would have been finished.

Ype Wijnia

Zwolle, March 2016

PART 1 RESEARCH FRAMING

In this part, the research problem is introduced, followed by a description of the research context, the theoretical framing and the research design.

1 Introduction

In 1999, the Dutch energy market was liberalized, following the European directive. The general idea was that consumers should have a choice in the supplier of their energy[1]. An anticipated side effect was that due to competition the prices would drop and efficiency would increase. The energy suppliers would thus be exposed to true market forces. The distribution grids are a natural monopoly, as it would be very costly (and thus inefficient) to install and operate multiple grids alongside each other. Yet, allowing a market party to operate a natural monopoly leaves many options to abuse this monopoly, either by direct blocking of new entrants or by cross subsidizing the market operations by distribution fees. Hence, the existing energy companies were required to unbundle into a (commercial) energy supplier and an independent Distribution Network Operator (DNO). To safeguard a fair market, a regulator was established, the DTe[2].

One of the tasks of the regulator was to determine the income of the network operators. To prevent any abuse of their monopolistic powers, network operators were only allowed to charge “fair market prices” to the customer. In 1999, the DTe started a consultation on price cap regulation. Under a price cap regime, the allowed income is based on the costs the grid operator could have (the efficient costs), not on the cost the grid operator actually has. This latter form is cost plus regulation, essentially[3] in place before liberalization. The efficient costs would be determined by benchmarking the grid operators and using the lowest cost per unit as a reference. If operators could do better they could keep the difference, but if they performed worse they would have to pay themselves (i.e. pay out less dividend to the shareholders). The benchmark would be held periodically, so that if companies improved their performance, the efficient cost level (the frontier) would move[4]. In order to remain profitable, the companies would have to improve again and so on. In essence, a cycle of continuous improvement on a very high level.

The Netherlands was not the first country in the EU to liberalize the markets. The UK had done so in the early 90s. Price cap regulation was also practiced in the UK. When liberalization happened in the Netherlands, it was quite obvious for the network operators to look overseas to get an impression of what to expect. Unfortunately, the outlook was not very pleasant. Precisely in the period the Dutch sector became aware of the changing reality, the regulator in the UK imposed some very stiff income

---

[1] The freedom of choice was introduced gradually. In 1999, only the very large consumers would have a choice, followed in 2001 by small and medium enterprises and in 2004 (originally planned for 2007) all consumers would be liberalized.

[2] The regulator had several changes of name. The original name was DTe (Dienst uitvoering en Toezicht Energie). In 2005 it became part of NMA (Nederlandse Mededingings Autoriteit), the Dutch regulator for (general) competition. The name changed to Directie Toezicht Energie, the abbreviation remained DTe. In 2013, the NMA itself was transformed into the Authority Consumer and Market (ACM).

[3] Though it was not the regulator but the local government which owned the energy company that allowed the proposed income by the energy company.

[4] It would be corrected for the inflation by the formula CPI-X, with CPI for consumer price index and X for the general efficiency improvement.

reductions, up to 30% per year (Krol, 2000). Such drastic income reductions were not foreseen immediately, but a reduction rate of a few % per year was within expectations. A single small reduction of income is not very troublesome. But if the income is reduced by a certain rate year after year, even small reductions would cumulate into a significant total. Whether a 30% income reduction is the effect of a single measure or the cumulative effect of several measures, it is a serious challenge by all means.

Together with the awareness that something was going to change drastically, the sector became aware of a potential solution. This was asset management. Companies that "suffered" from income reductions were proudly presenting[5] the form of asset management they implemented to maintain profitability whilst at the same time improving their performance. Asset management was by no means a coherent set of concepts in those days, demonstrated by the wild divergence of the presentations held. Yet, a concept from those early days that has survived the test of time is the balance between costs, risk and performance, as shown in the Figure 1[6].

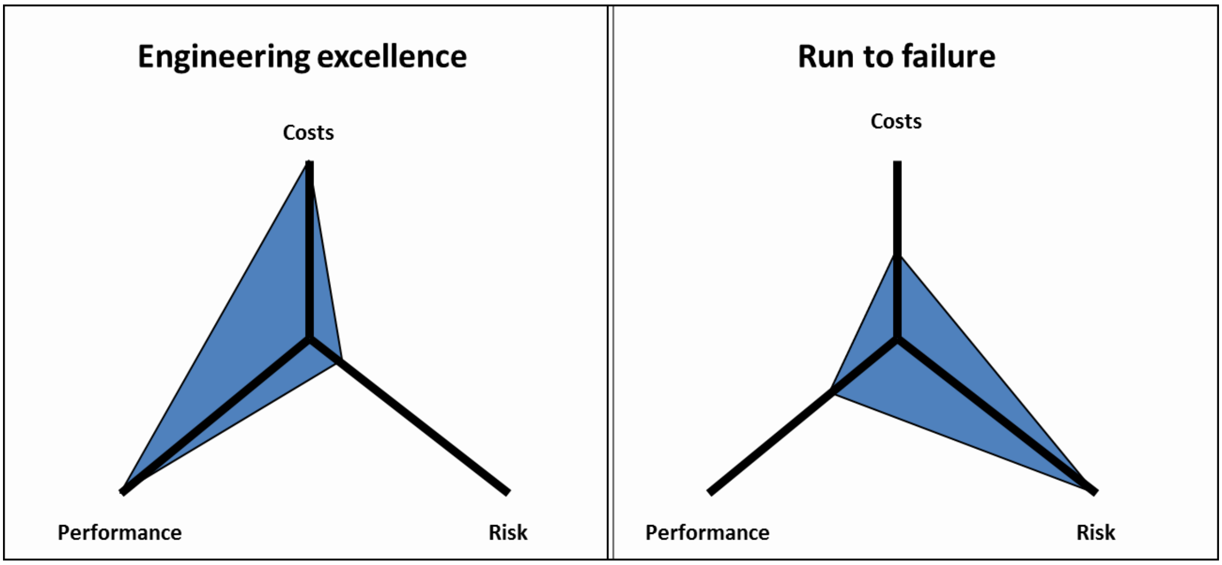


Figure 1: Extremes in the cost/risk/performance balance, adapted from Yorkshire Electricity (Wijnia and Huisma, 2007).

The diagram relates the business relevant attributes of an asset (costs, risk and performance) to the asset strategy that is applied. As an example, two extremes[7] in maintenance strategies are displayed. One extreme (on the right) is "Run to Failure". The asset is operated until something goes wrong and a corrective action is needed. The other extreme is that of engineering excellence. In this strategy,

---

[5] The first Asset Management conference attended by Enexis was organized by IIR, 18-19 january 1999 Londen. Among the presenting utility companies were Yorkshire Electricity, Railtrack, Severn Trent Water, National Power, Hyder Utilities, London Electricity, South West Water and Transco. Other organizations present were consultancy firms like the Woodhouse Partnership, Logica, Andersen Consulting.

[6] The precise origin of this diagram is unknown. Earliest records in possession documenting the use by Yorkshire electricity date from April 11, 2001.

[7] Extreme in the attention that is given to the asset, not necessarily in any of the attributes.

preventive maintenance is frequently applied to ensure risk free operation. It is the role of the asset manager to understand what options are available and what their impact on cost, risk and performance is. This requires a deeper understanding of the asset, for example by means of a fault tree analysis. The novelty of the diagram was in the explicit consideration of risk, whereas it used to be implicitly covered by applying technical standards[8].

Some of the organizations promoting asset management established the Institute of Asset Management[9] (IAM). In a collaboration between IAM and the British Standards Institution (BSI), chaired by John Woodhouse, a Publicly Available Specification 55 (PAS 55) on the optimal management of physical infrastructure assets (BSI, 2004a) was developed. The idea of optimizing the balance between costs[10], risks and performance over the whole lifecycle is reflected in the definition of asset management in PAS 55:

> *[Asset Management is the] systematic and coordinated activities and practices through which an organization optimally manages its physical assets, and their associated performances, risks and expenditures over their lifecycle for the purpose of achieving its organisational strategic plan.*

PAS 55-1 describes[11] the requirements for an asset management system. Its structure is aligned with other standards on management systems like ISO[12] 9001. Requirements address for example documents that need to be in place, processes that need to be established, and roles and responsibilities that have to be defined. With regard to risk identification, a list of topics that should be included in the assessment is put forward as a requirement.

Like other standards for management systems, the focus of the requirements is on what needs to be in place and not on how to realize that. That "how" is addressed in part two, the guidelines for the

---

[8] In hindsight, explicit risk management was one of the three pillars of asset management (Woodhouse, 2014)

[9] According to the term of reference for patrons (IAM, 2006), the seven Founding Members are Anglian Water (AWG), London Electricity/LPN, National Grid, Northumbrian Water, Railtrack, Severn Trent Water and Yorkshire Electricity and their legal successors. IAM was founded in 1994 (IAM, 2013). John Woodhouse holds member certificate number 001 (Burns, 2010)

[10] The terms cost and expenditure are often interchanged. Technically, the expenditure is the (observable) cash flow, whereas costs are the way the expenditures are accounted for in the income statements. By agreement, maintenance expenditures are generally booked as onetime costs (hence the interchangeability), but investment expenditures are translated into costs by depreciation. As asset management is not limited to maintenance, the correct term should be expenditure, but cost is used more often. PAS 55 itself uses expenditure in the definition of asset management, but costs in the definition of optimal in the guidelines (PAS 55-2, section 0.1, 5th bullet).

[11] The comments here are made on PAS 55 as published in 2004. As it is no longer valid, past tense should have been used. To a large extent, however, the comments still are valid for its successors, PAS 55:2008 and the ISO 55000 series. Therefore present tense is used and not past tense, as that would suggest the mentioned problems were solved.

[12] International Organization for Standardization. Throughout this thesis the abbreviation ISO will be used.

application. Given the wide range of organizations in which the standard should be applicable, it is clear the guidelines cannot be very specific. It is more a list of items that can be considered in implementing asset management than an implementation plan.

A paradoxical item in PAS 55 is (risk) decision making. The whole idea of asset management is to make the organization think about the value the assets deliver to the stakeholders as a primary concern and derive the technical requirements from that consideration, instead of considering what is technically achievable given the budget and other constraints. Considerations only become effective in decisions. Asset management in this view thus is centered around decision making. This is reflected in both the specification itself (one of the recognized benefits is to provide evidence that the right decisions are made)[13], and even more so in the guideline on the application (a systematic approach for consistent decision making)[14]. Yet, PAS 55 does not specify how decisions in asset management should be made. The term decision or decision making is not even mentioned in the requirements at all. In the requirements, decision is only mentioned in a note to the asset management policy (that other policies may exist to provide guidance and a clear framework for decision making)[15].

From a certain perspective this is understandable. PAS 55 was drafted by a diversity of infrastructure managers and should be applicable by all of them. Too much specification then can become an obstacle. But asset management is also making organizations think about the value their assets provide. This is a change of paradigm compared with the “old” organizations dominated by technical regulations, technical considerations and hidden value judgments. Replacing a prescription on how to build things by a prescription on how to decide how to build things would most likely not bring the needed cultural change.

From a more distant scientific perspective however it is quite strange. Even though the technologies between the public infrastructures are different, the stakeholders for those infrastructures are comparable, if not precisely the same. Roads, railways, electricity and gas grids, water and sewage networks all have the same users to a large extent. As the infrastructures are within the public domain, failing assets may impact (outside the users) the people living nearby. Again, impacts may differ between the infrastructures, but the impacted stakeholders are very much alike. Therefore, if the interests of the stakeholders are considered in decision making on the risks the assets present to them, it seems very reasonable to assume that those considerations are quite comparable between infrastructures. That implies there is an opportunity for standardizing the considerations, but that did not happen. There is only a small hint on the interests, requirements or values the stakeholders could have by stating that stakeholder requirements should include health, safety, sustainability and environmental requirements. In the guideline (PAS 55-2) some more hints with regard to decision making are made, like putting a monetary value on non-financial aspects, the use of cost benefit

---

[13] PAS 55-1, list of benefits of asset management, top of page vii (BSI, 2004a).

[14] PAS 55-2, section 0.1 General, page V second bullet, on a successful implementation of asset management (BSI, 2004b).

[15] Note 2 on the asset management policy (4.2.1), page 4 (BSI, 2004a).

analysis, net present value calculations and so on, but these are not prescriptions. The idea of PAS 55 really seems to be that organizations determine that for themselves.

The paradox with regard to decision making is not unique to PAS 55 and its successors. It is also observable in regulation, for example regarding DNOs for electricity and gas in the Netherlands. As mentioned, the regulator determines the allowed income for the DNOs. This is a benchmark based decision in which DNOs are compared on the cost per unit they realized in the past period of regulation. The observed trend in the cost per unit (presumably downward) would be used to determine the income in the coming regulation period. Over the years, this resulted in a significant reduction of costs per unit to the consumers in comparison with unregulated costs (Berndsen et al., 2012). Since 2004 (NMA, 2007) quality is part of the income regulation. Part of this quality regulation is the compensation for customers for long interruptions of supply, another part is based on the average quality of supply by means of a q factor[16]. Both aspects regard past performance.

Theoretically, such a “feedback only” regulation of income has its drawbacks. That an efficiency improvement was realized in the past does not mean it can be realized in future. Furthermore, given that an infrastructure is an inert system, it is possible to reduce costs by “mortgaging the future”, i.e. the postponement of maintenance, replacements and capacity upgrades, only to result in massive costs in the eventual failure. A typical example of how high cost of failure can become is the Auckland scenario of 1998. A series of high voltage (HV) cable failures, due to a poor condition and inadequate capacity, left the central business district without full power for more than a month (Ministry of commerce of New Zealand, 1998).

To safeguard against such disasters, several measures are taken. On the one hand, there are technical codes. These state (or make reference to the relevant norm) the technical requirements for equipment, the quality requirement for the transportation service, and planning criteria for the high voltage grid[17]. These technical codes have been in place since 2000, the start of regulation (ACM, 2014b). Technical codes are specifying the minimal requirements.

The other measure is more in the style of a management system like PAS 55. DNOs are required to prepare a (public) plan for the infrastructure every 2 years on how they will comply with the planning criteria. The first plans for the electricity grid were drafted in 2000, considering the 7 year period 2001-2007[18]. These only addressed capacity problems, hence the name Capacity plan. In 2002, a similar document had to be prepared for the gas grid, alongside with an updated version of the electricity plan. Since 2005, the plan also has to include quality issues (changing the name to Quality

---

[16] The q factor is calculated by comparing the performance of the DNOs over the past 3 years. The total income effect of the q factor over all DNOs is zero (ACM, 2014a).

[17] These planning criteria are often referred to as the n-1 and n-2 criteria. N-1 means that any component of the grid can fail without impacting supply, n-2 that any component can fail during maintenance. The difference is that for n-1 the peak load for the year has to be considered, and for n-2 the peak load during maintenance. If maintenance is planned during a low load situation, the n-2 requirement may actually be less stringent than n-1.

[18] Comparable to the seven year statement DNOs in the UK had to make.

and Capacity Document). Furthermore, the approach for risk identification and the analysis of the major risks had to be included. The review period was extended from 7 years to 10 years in 2011.

Combined, these two measures should have the effect that the DNOs manage their risks adequately. The planning criteria guaranteed a certain level of quality, and the public plans allowed the DNOs to be monitored on their compliance with the planning criteria. However, in a separate guideline (Rijksoverheid, 2005)[19] it is explicitly stated in article 13 that nonconformity against the (HV) planning criteria is allowed if the benefits do not outweigh the cost. But nowhere in the laws or regulations it is stated explicitly how this cost benefit analysis should be made[20]. A similar indeterminacy can be observed in the risk management paragraph of the quality and capacity documents. Every DNO reports the major risks, but for each one the list is different, even differently structured. This applies as well to different DNOs within the same discipline (electricity or gas) as to different disciplines within the same DNO. For reasons of comparability between DNOs it would be a great help if the same risks would be reported, but that did not happen yet.

Summarizing the paradox, there is great value in changing the way decisions are made about risk, but guidance on how to make the decisions is not given. That this did not occur in the standardization of asset management is understandable, given the diversity of asset bases to manage and the (potential) application in different legal systems[21] with perhaps different attitudes towards risk. But that it did not occur in (presumably) very comparable asset bases within the same regulatory regime seems like a missed opportunity for a faster improvement[22]. Is there perhaps a more fundamental reason why standardization of risk decision making does not occur? If so, what level of standardization could be achieved despite that fundamental problem? Thinking even beyond that, suppose a reasonable level could be achieved, what would that mean for a standardized balance between costs risk and performance for an infrastructure, towards all network operators could work? In other words, what are the boundaries of optimization in risk based asset management in this infrastructure context?

These questions will be addressed in this thesis. The thesis will be split into 3 parts. In the first part, the research will be framed. As the research took place in the real world and not in a laboratory, understanding the context is vital for appreciating the findings. Therefore, this thesis starts with specifying the research context in terms of historical development, the used technology and asset management in this historical setting. Next, the theoretical framework will be specified. Central elements in this framework are asset management and risk analysis, with a focus on the concept of

---

[19] This article is still valid (Rijksoverheid, 2013).

[20] Implicitly it seems reasonable to use the q factor. However, the q factor is based on the average performance of all DNOs over several years and not an absolute number. Theoretically, it is therefore possible that if all DNOs decide that a certain investment in reliability is not worth the cost and it is better to accept the income reduction by means of the q factor, the actual effect is zero because all DNOs move in the same direction.

[21] Despite being a British specification, the application of PAS 55 was not limited to the UK. This is further highlighted by its development into an international standard, the ISO 55000 series.

[22] As demonstrated, It did not happen in the Netherlands. No evidence of standardization of decision making in other regulatory systems was encountered during this research.

value with regard to risk. After this, the knowledge gaps will be specified, followed by the research questions and the design of the research.

The second part of this thesis describes the experiments that were conducted. All experiments but one have been published. This thesis contains the full text of those publications, with minor changes to create consistency of style throughout this thesis. To align the experiments with the line of thought of this thesis, some of the publications have been amended with an epilogue. In those epilogues additional experiments and other literature findings will be used to validate the conclusions of the experiment beyond the experimental setting.

The third part of this thesis is on the findings, discussion, conclusions, reflection and recommendations. First of all, the findings per experiment with regard to optimization are summarized. In the discussion, the findings of several experiments will be combined per knowledge gap. This will be summarized into answers to the research questions in the conclusion. After the conclusion a reflection will be made on this research, followed by a future outlook on the science of asset management. The thesis ends with recommendations for further research and recommendations for practice.

2 Research Context

The experiments for this research have been conducted at Enexis[23], a Dutch distribution network operator (DNO) for electricity and gas. Enexis serves about one third of the Dutch market. The maps below shows the area where Enexis is active, marked by 4.

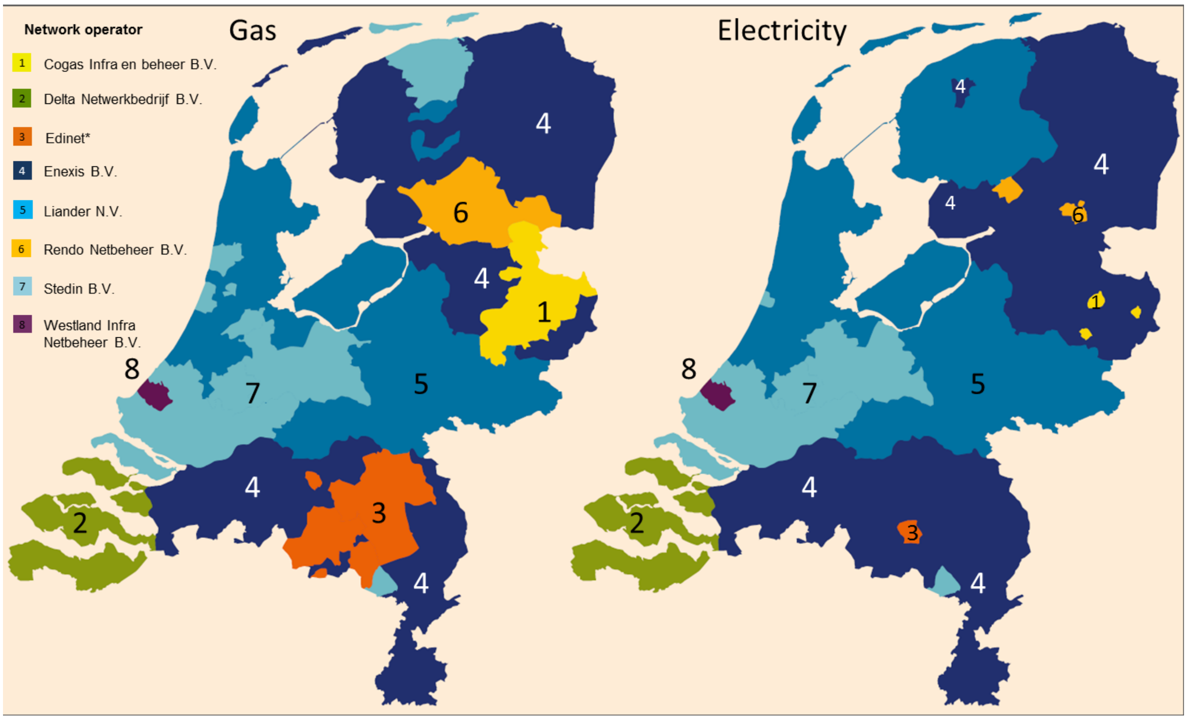


Figure 2: The energy distribution network operators in the Netherlands. Based on Energietrends 2014 (Netbeheer Nederland, 2014). Numbering by author.

The service area of Enexis is not continuous. Two more or less continuous zones can be recognized for electricity. The provinces Groningen, Drenthe, Overijssel and part of Flevoland form the area at the top of the map, whereas Noord-Brabant and Limburg form the area at the bottom. In these areas, only a few cities are served by another network operator. This situation is reversed for the province of Friesland, where Enexis only serves one city, Leeuwarden. For gas, the service area is quite different. Enexis is active in the same provinces, but not in the same municipalities. Larger parts of Drenthe, Overijssel and Noord-Brabant are served by other DNOs. On the other hand, Enexis serves a larger part of Friesland for gas than for electricity. The reason for this discontinuous service area is historical. Enexis is the result of a series of mergers, which followed a different trajectory for gas and electricity. In total 84 different companies[24] combined efforts in Enexis.

---

[23] During the course of this research the company changed names twice. The research was initiated in Essent Netwerk Noord. After an internal merger in 2004 of the DNOs within the Essent Company this changed into Essent Netwerk. In 2009, the name was changed to Enexis, in parallel with the ownership unbundling required bij the WON (Wet Onafhankelijk netbeheer, Law for independent Network Operators). For consistency reasons the name Enexis will be used throughout this thesis.

[24] This number was mentioned by Herman Levelink, first CEO of Enexis, at his retirement event.

2.1 Historical overview[25]

The oldest parts of Enexis are the gas companies of some cities, which date from the middle of the 19th century. The table below gives an overview of the gas companies in several large cities in the Enexis area.

Table 1: The establishment of gas companies in some large cities in the Enexis service area, after van den Noort (1993). Some companies started as a private company, others as a municipal company. Eventually all companies became municipal. The year in which a company became municipal (if started privately) is also indicated in the table.

City	Establishment of company	
	Private	Municipal
Leeuwarden	1845	1865
Zwolle	1848	1855
Maastricht	1849	1858
Tilburg	1853	1873
Groningen		1854
Den Bosch	1854	1890
Deventer		1858
Breda		1858

Gas in those days was mainly used for (street)lighting, often by means of an open flame. It replaced candles and oil burning lamps. At the end of the 19th century, other uses were promoted as well, like cooking, hot water and heating. Given the cost of gas, it was considered a luxury product (van den Noort, 1993, Overbeeke, 2001). The gas was produced in central gas factories, temporarily stored in a large tank (the gasometer) and distributed by means of pipelines to the end users. The system was operated at low pressure. In gas distribution, pressure is only needed to transport the gas to the users, the energy for the users is in burning the gas. The pressure needs to be high enough to deliver enough gas to the farthest user. Leakage of pipelines was a problem in the early systems, which would increase with higher pressures. Combined with technological considerations of the production and gas treatment facilities (often running at ambient pressure) and the available pumping technology a low distribution pressure was chosen. The gas distribution systems were initially built as islands without interconnection. Several technologies were available for producing gas, but most were based on gasification of coal, producing the valuable byproduct coke. Given the demand in some industries for coke (e.g. blast furnaces), the roles sometimes were reversed. A coke factory was then built, selling the byproduct manufactured gas. This often required transportation at a higher pressure, provided by pumps on site. At the feed in point to the distribution grid, pressure would be reduced by means of an automated valve. The same technology could be used for expanding the range of existing gas factories.

---

[25] The three sources mainly used for this section all had a slightly different viewpoint. Van den Noort reviewed the history of utility companies in general, Hesselmans' work was a biography of Feldman, the "founder" of the provincial organization of electricity and Overbeeke reviewed the consumers influence on the development of the gas and electricity system. The three sources generally agree on the developments. Explicit references are only made if the source makes a unique statement.

At the end of the 19th century lighting by means of electricity became an alternative to gas. The first lamps (Jablochkoff candles, arc lights) were very bright, suitable for spotlights, street lighting and lighthouses to give a few examples. For indoor household use they were too bright and the gas lamp was generally preferred. The "invention" of the incandescent light bulb by Edison opened up the opportunity for use at a smaller scale. At first, it were often private installations in which a dynamo was attached to an existing steam engine or gas motor to provide power to the lamps. Sometimes these private plants would feed a small area around them (van den Noort, 1993). These electrical systems were in structure very comparable to that of gas: a central production facility and a (direct current) local distribution grid. Sometimes storage in the form of a battery would be used. Distribution occurred at relatively low voltages. In electricity, the power delivered is the product of voltage and current. It would not suffice to get enough current to the user' it would need to be at the right voltage as well. Because of resistance in the conductors, currents for the user would create a voltage drop over the conductor. This could be overcome by using a higher voltage, reducing the current for the same power. However, light bulbs ran at 110V[26]. As no economically feasible[27] technology existed to lower the voltages at the appliance, the whole system needed to operate at 110V[28].Such a low voltage limits the economical transportation distance to a few hundred meters. Larger cross sections would allow a larger distance, but above a certain distance it would become economically more attractive to build a new production plant with its own grid. The invention of the transformer (presented at the world fair 1889 in Paris) changed this. It allowed production and transportation to operate at a higher voltage than end use required, thus facilitating longer distances. This in turn allowed for larger systems and the associated economies of scale. However, it required alternating current. Most existing installations were direct current. Yet, the cost advantage was that large that AC systems outcompeted the existing DC systems[29], and many private dynamos were replaced by a connection to the grid if that became an option. Alternating current, however, has the disadvantage that it cannot be stored[30]. The power had to be produced at the precise moment it was consumed. Given the dominant use of lighting, the peak was very sharp, leaving most power plants running at idle most of the day. Promoting other uses outside these peak hours would drastically increase the load factor of the system. Typical other uses were motors for the industry and electrical trams, but also household appliances like irons and vacuum cleaners. Development of public grids did not occur immediately after the technical possibilities arose. As stated, electricity was a competitor for gas, and many municipalities had a profitable gas company. Only by the realization

---

[26] 110V DC is (not coincidentally) about the maximum voltage that is safe to touch (Marx, 2010). Distribution grids at this voltage therefore did not require many safety precautions.

[27] Technically, a resistor in series would achieve a voltage drop, but that would result in considerable losses of energy.

[28] In some DC systems a three conductor line would be used, at +110V, 0V and -110V. This allowed putting two appliances in series. The three wire system doubled the capacity at only 50% more costs.

[29] This competition between AC and DC is often referred to as the War of the Currents (McNichol, 2006). One of the arguments in the battle was that AC was less safe, which it in fact is. Safe voltage to touch is 50V AC compared to 120V DC (Marx, 2010).

[30] This in contrast with direct current which can be stored in batteries, of gas that was stored in gasometers.

that the electricity company could be even more profitable (Hesselmans, 1995) municipalities got involved.

Advantages of electrical lighting were the brightness of the light and the option to operate it remotely. This competition stimulated innovation in gas lighting. One was the invention of the gas mantle, giving a much brighter light than an open flame. The second was auto ignition, which eliminated the need for lamplighters. Yet, this only provided a temporary advantage. A similar development occurred in the electric lamps, reducing costs and improving quality. Besides, World War 1 led to shortages in fuels of any kind. A kilogram of coal provided more light by means of electricity production and light bulb than by means of gas production and gas light. Furthermore, power plants were less discriminative with regard to the quality of coal than gas factories, which required coal with high amounts of gas. Cooking on gas, in its turn, was more efficient than cooking on coal, the dominant heat source those days. As a result, the Dutch government promoted electric lighting, cooking on gas and heating by means of solid fuels (Overbeeke, 2001). This decided the battle between electric and gas lighting in favor of electricity. The electrical grids were expanded rapidly. But because of the promotion of cooking on gas, the gas grids did also developed further. Additionally, the use of gas for hot tap water developed.

In the same period, provinces got involved. Around 1910 initiatives for provincially organized electricity systems arose, by which even smaller communities could be efficiently supplied. A benefit of a larger scale organization of electricity was that reserve production capacities could be shared[31]. Before 1921 almost all provinces would have an electricity company. The exceptions were Zuid-Holland (covenant between Rotterdam and The Hague instead of provincial Company) and Drenthe (divided between Groningen and Overijssel[32]) (van den Noort, 1993). The provincial companies took care of production and transportation, incorporating existing private and public companies. Distribution was predominantly organized on the municipal level, with support of the provincial company. However, in the 1920s some of these local distribution companies were taken over by the provincial companies because they did not put enough effort in the promotion of electricity, resulting in low numbers of connected households. The provincial organization resulted in a rapid electrification of the Netherlands(Overbeeke, 2001). The provincial companies used a similar strategy for increasing the load factor of their system as the municipal companies did before: promoting the use of appliances in off-peak hours. Around 1930, the range of appliances was expanded with electrical cooking and electrical boilers for hot tap water. The extra focus on households was to compensate for the decrease in industrial demand in the economic crisis. This shift of focus revived the competition between electricity and gas. Gas companies reacted with improved cooking and hot

---

[31] This also motivated scaling up respectively towards a national grid and the European interconnected zone. With the need to share reserve production capacity over larger areas, the need for a reliable transportation grid also developed. This resulted in the n-1 planning criterion for transmission systems, which in the early years were the current medium voltage systems.

[32] This is still visible in the structure of the high voltage grid of Groningen, Drenthe and Overijssel. Half of Drenthe is supplied by the Overijssel HV grid, half by the Groningen grid, though nowadays interconnections exist.

tap water devices. The number of connections steadily grew in this battle for the household market. The growth was temporarily interrupted by World War II, but after the war, growth continued.

In 1948, the first natural gas was found in the Netherlands, near the town of Coevorden. The local gas company took the initiative to distribute this gas to its customers. Several other finds and local distribution in other parts of the Netherlands followed. As the cost for producing and transporting natural gas are much lower than that of manufactured gas, this allowed the gas companies distributing natural gas to expand their gas grids even further.

In 1959, a large gas field was discovered in Slochteren. The amount of gas was that large, that it became a major concern to find a market big enough to sell the whole reserve in 10 to 20 years. It was foreseen that nuclear power would become so cheap that it would outcompete every other energy source. The initial plan was to sell gas to industrial clients and power plants. However, gas would compete with coal and prices would be relatively low. A much more interesting market was that of household consumption. The price then would be in the manufactured gas range. The plan was developed to roll out a nationwide grid for natural gas which would be connected to the existing distribution grids for manufactured gas, to which by that time about 75% of households was connected. Part of the plan was the transition from coal and oil towards gas for household heating. This would result in an increase of the energy consumption per gas connection. However, it did not require an capacity expansion in the distribution grids. Natural gas contains about two times as much energy per volume compared to manufactured gas. Additionally, the distribution pressure would be increased from 7mbar to 25 mbar[33]. In total, this resulted in a sevenfold increase of peak capacity. As the gas consumption for heating was assumed to be spread more evenly over the day, the load factor of the system would also increase, allowing for much more energy per unit of peak capacity.

The change in energy content and pressure was not without consequences though. First of all, it required the replacement or adaptation of the gas appliances. This was a major operation conducted during the phased, area by area, roll out. The conversion of the grids was completed in 1968, less than 10 years after the discovery of the Slochteren field. Furthermore, the increased pressure and changed composition of the transported gas led to an increase in the losses. In manufactured gas grids, losses were below 10%, but some companies that changed towards natural gas reported initial losses of 25-35% (Overbeeke, 2001). Another issue was the safety of gas. Both manufactured gas and natural gas can form an explosive mixture. Due to the higher energy content of natural gas, a lower concentration is enough[34]. At the same gas outflow, the lower explosion limit (LEL) is therefore reached faster. But because natural gas was distributed at a higher pressure, the outflow would be much larger for the same leak, decreasing the time to LEL even further. Besides, natural gas is odorless, making detection of a gas leak much harder than that of manufactured gas with its strong smell. To compensate for this, an odorant was added in distributed natural gas. Natural gas, on the

---

[33] This pressure was due to a standardization in burners for several gas qualities. To get the right air to gas ratio, a gas with a higher energy content would need a higher pressure. Decreasing the cross section of the jet (a simple replacement set) would then be sufficient to achieve the desired flame size (Overbeeke, 2001).

[34] 5-14% for natural gas, 9-24% for manufactured gas.

other hand, has the benefit that it is not poisonous, whereas manufactured gas is because of its fraction of carbon monoxide.

Following the conversion of grids for manufactured gas to grids for natural gas the gas factories were decommissioned. Some of the gas companies[35] thus lost a profitable branch. This loss however was more than compensated by the perspective of new sales, as the gas was intended for heating purposes. Many gas companies extended their service area, and even in rural areas distribution companies were established. In some cases, several municipal (city bound) gas companies combined powers in a merger so that the gas grid could be expanded into the rural area in between (Overijssel, 2013). In this period, the competition between gas and electricity evaporated. Several factors worked in parallel to achieve this. First of all, the paradigm change of gas companies towards heating meant the volume would increase by many times. If gas was only used for cooking, electricity could offer competitive prices. As the cost per unit would drop with increasing volumes, electricity could not compete for houses that were heated with gas. This was a concern to electricity companies, as it might limit their sales of household appliances, especially for cooking and hot water. This concern became reality, at least with regard to the percentage of households[36]. Yet, in the same period the number of households almost doubled, from 3 million to 5 million[37]. In absolute terms the sales thus continued. Furthermore, the total use of electricity per household grew significantly in this period, from some 600 kWh per year to some 3000 kWh[38]. Additionally, industry grew drastically. All growth was driven by the economic boom in the post WWII era. Both gas and electricity thus faced a very large increase in sales volume, mostly in non-competing uses[39]. In those market conditions, there were little incentives for fierce competition between electricity and gas. Given growth rates in the range of 10% per year[40], keeping up with the growth was challenging enough. As a result, heating, hot tap water and cooking are predominantly gas based, and every other energy use in the household is electricity based. At the end of the 1970s, the rapid growth of the economy came to an end, slowing the growth for gas and electricity. For gas, this was also the result of energy awareness,

---

[35] According to Overbeeke (2001), in the 1950s 58 gas companies had own (partial) production facilities, whereas 115 companies only distributed gas produced elsewhere.

[36] Electricity had some 10% of the market for cooking, and some 15% for hot tapwater in 1960. Over the next 20 years, these percentages hardly evolved, whereas gas grew significantly with regard to hot tap water (from 35% to 80%). In cooking, gas was already dominant in 1960 with more than 80%, according to the graphs in chapter 13 of Overbeeke (2001).

[37] Number based on statonline, CBS, historical figures, households.

[38] Current use of electricity per household is in the same range. The use of gas per household has decreased however, due to a savings program since the 70s energy crisis.

[39] Electricity is no serious contender for heating if a gas is available by means of a grid because it is much more expensive per unit of heat.

[40] Meaning a doubling of the volume every 7 years.

which led to energy conservation programs[41], resulting in a reduction per household. Energy awareness also influenced the electricity consumption, but much less.

The trend in scaling up started by the regional gas companies continued for several decades. The regional gas companies would often merge with the provincial electricity company to form a provincial energy company. In case of Enexis these were EGD for Groningen en Drenthe, IJsselmij for Overijssel (which then still contained the Noordoostpolder, currently part of the province of Flevoland), PNEM for Noord-Brabant and MEGA for Limburg. The provincial energy companies themselves would merge into larger groups. For Enexis these were EDON (EGD and IJsselmij) and PNEM/MEGA. These groups merged to form Essent. In 2010, due to the WON[42], the distribution company had to be split off and Enexis was formed.

For some regional gas companies however the order of mergers was reversed. They first merged with the municipal electricity companies to form regional energy companies. These electricity companies only covered some cities, as the rural area was electrically serviced by the provincial company. The regional energy thus covered a continuous area for gas, with some islands in which they also provided electricity. Some of these regional companies stayed independent, others joined a larger group. This explains the somewhat odd service area of Enexis and the other network operators in the Netherlands.

---

[41] Examples are insulating houses (double walls filled with insulating material, double glazing) and turning down the heat in room that were not used. These measured translated into regulations for new buildings.

[42] Wet Onafhankelijk Netbeheer, translated: Law for independent network operators.

2.2 Asset management in historical perspective

Even though asset management was not coined as a concept when the distribution grids were rolled out, in an historical overview many asset-management-like considerations can be recognized.

At the start of both gas and electricity, the service that was provided was a luxury: continuous light. In the design of the systems to provide these services several economic considerations were made, like the choice of voltage/pressure and the cable/pipe diameter in relation to the service area and the losses in servicing the customers. Other aspects of concern were the reliability and safety, though no evidence of explicit valuation were encountered, it was more of an engineering judgment on what was good enough. However, lighting was a service that was only required a few hours per day. As the installations were already there, off-peak services would not require extra investments and thus could be offered at lower prices. This pricing strategy would help increase the utilization of assets, another asset management like decision.

The development of the grid is also characterized by a constant quest for economies of scale. The product of electricity and gas was cheaper per unit if produced in large quantities, but it required parallel developments in the transportation and distribution technology. As a benefit, the larger electricity grids required less reserve production capacity per customer, thus increasing the utilization of the reserves (and potentially freeing up existing reserve capacity for normal production given the constant growth of demand).

After initial competition between the services, the markets stabilized with cooking and heating serviced by gas and light and other appliances serviced by electricity. This market segmentation was also the result of an asset management decision, but then from the customer, as it was the segmentation that best provided the end needs. The stabilization of the market allowed a different economy of scale, the merger of the gas and electricity company. In a competing market that would have been not very useful, but as most customers had both electricity and gas, services around the connections (like metering, billing and the like) could as well be handled by a single party.

Another characteristic of the development of the infrastructures is the continuous and rapid growth over a significant amount of time, until almost full coverage of the potential market was reached. This meant the main focus of the organizations was on investment in new capacity, including the financial planning to facilitate that. Maintenance, operation and replacement were not a top priority. But this lack of attention was also facilitated by the high reliability and longevity of the used components, and the fact that they (the grid components) did not need active operation to function. Only when the use of the assets was changed to avoid new investments (as in the change from manufactured to natural gas) operational issues could pop up to claim attention.

Full coverage of the markets was reached around 1980. Due to energy awareness and the economic crisis in the 1980s, the consumption per household started to decline. The only need for investments then arose from connecting new households, still growing in numbers. This meant focus in the organizations gradually changed from building new assets towards managing existing assets. This is witnessed by the rise of asset management initiatives in the infrastructure section in this same period. With regard to the future of asset management, new developments like electric vehicles and local energy production can introduce a new growth of required transportation systems. This may trigger a new change of paradigm.

In technological terms, the developments have always been in the direction of cheaper[43], more reliable and safer technologies. A constraint for any of these improvements was that they would have to work in connection to existing grids. In the primary elements (cables, pipelines, switches, valves, transformers etc) this meant that voltages and pressures could not be altered, and basically 1900s technology could still be used in the current grids. But in the secondary elements (the controls) much more degrees of freedom exist. Being less constrained, developments went much faster, from none, via mechanization, electrification to digitalization, thus enabling development in operation from manual via local automation and central operation to remote operation over the internet. It is this area where the biggest developments like smart meters and smart grids are foreseen, especially for accommodating the new types of loads mentioned.

The long history of continuous development of the distribution systems has resulted (in the Netherlands) in a very high performing system. The reliability of the (mostly redundant) grid ranks amongst the best in the world. The availability of electricity is 99,995% (on average 32 minutes interruption per customer per year), in gas the reliability is 99,99998% (one minute per customer per year). The numbers of minutes per customer is the System Average Interruption Duration Index (SAIDI), a common metric to compare reliability of grids. The energy loss in the electricity grid is about 4% of the transported volume, in gas it is about 0,1% (Wijnia and Peters, 2008). The table below shows reliability and energy loss data for electricity for a number of western countries.

Table 2: Reliability and loss data for a number of western countries (reliability data from (Council of European Energy Regulators, 2015, Australian Energy Regulator, 2014, U.S. Energy Information Administration, 2013), energy losses from (World Bank, 2015).

	SAIDI[44]	Losses [fraction of production]
Australia	225	5%
Belgium	37	5%
Denmark	21	6%
France	113	6%
Germany	31	4%
Netherlands	32	4%
UK	78	8%
USA	196	6%

---

[43] In the tender (PB-nummer 2009/S 36-052933) for power transformers by Enexis, Stedin/Joulz, Delta and Alliander (the big grid companies in the Netherlands) in 2009 the criterion was the best economical option which included the losses over the life span of the transformer. Winner was a very low loss transformer (http://www.alliander.com/nl/alliander/investors/publications/verantwoording/ketenverantwoordelijkheid/index.htm), Dutch sourcing awards sustainability 2010 (http://www.dutchsourcingawards.nl/). According to EN 50464-1, best current distribution transformer lose about 1%, though historically this was in the 2-5% range or even higher. The bigger the transformer, the better the performance becomes. The loss in High voltage transformers is more like 0,1%.

[44] SAIDI: System Average Interruption Duration Index (Short, 2002), the average time (measures in minutes) an average customer is not served per year. Given 525.600 minutes in a year, even at a SAIDI of some 500 minutes reliability would still be above 99,9%.

3 Theoretical framework

3.1 Asset Management

Asset management as a separate field of interest is relatively new. Even in 2010, asset management was still referred to as an emergent field (Amadi-Echendu et al., 2010). The oldest (scientific) records date back to the 1960s. Since then the field developed in several waves[45] both in practice and in science. Figure 3 gives an impression of the timeline of asset management. This timeline will be detailed in the next sections.

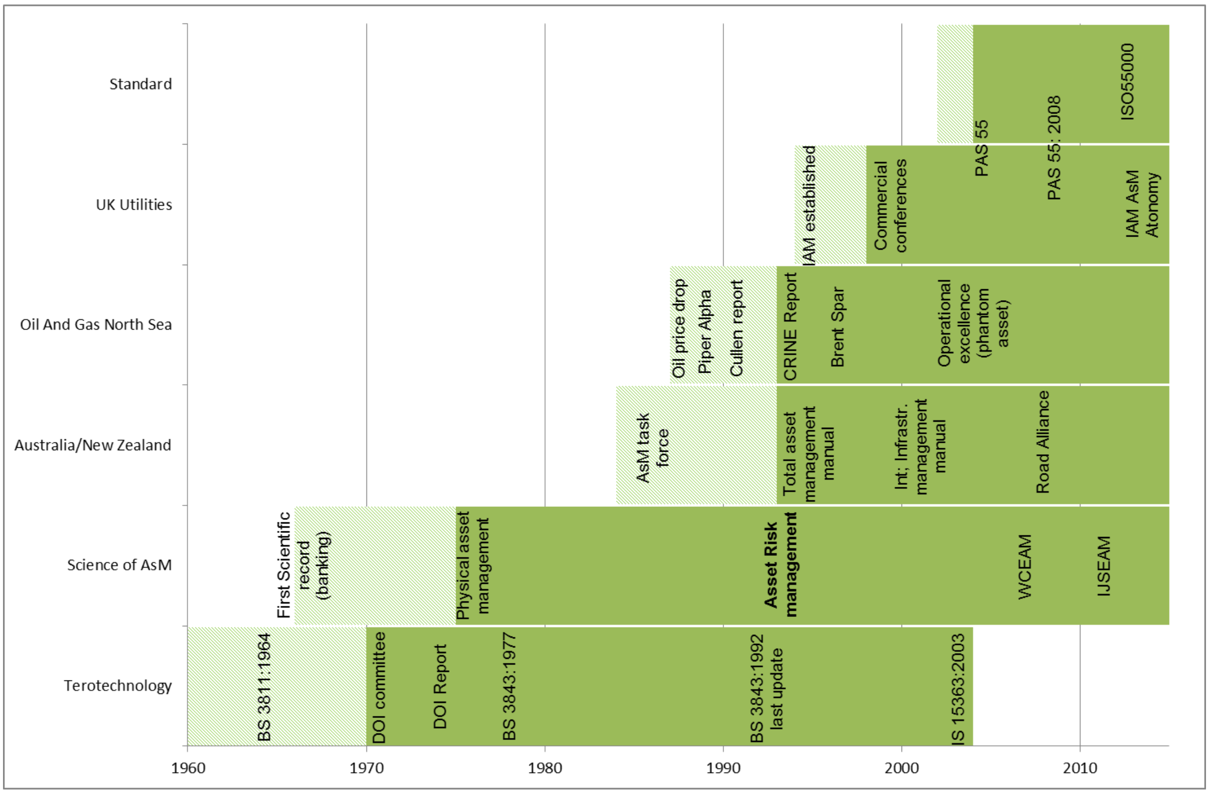


Figure 3: A view on the asset management timeline. The (scientific) origin can be traced back to the 1960s, though much of the development occurred in practice, especially in the 1980s and 1990s.

3.1.1 Scientific origin and development

The term asset management does not appear in scientific records before the 1960s. The oldest record in Scopus is from 1969[46]. In the secondary list, the oldest record with a title was from 1966[47], with the first journal article from 1967 about “[..] linear programming and optimal bank asset

---

[45] The inclusion of the developments in the North Sea into this diagram was inspired by John Woodhouse (Woodhouse, 2014).

[46] On February 23, 2015, Searching Scopus for “Asset management” OR “assetmanagement” produced 7.064 documents in the Scopus database and 6.210 in the secondary list (based on the reference tables of the documents).

[47] Though the oldest entry without a title was from 1879.

management decisions*”* (Cohen and Hammer, 1967). A similar result was achieved at Web Of Science[48] (WoS). There some 5.000 entries were found, with exactly the same journal article as the oldest entry.

Asset management in those early days was associated with the financial domain, i.e. assets as entries on the balance sheet. The first reference in Scopus with regard to physical assets only is made in 1975 in the publication **TEROTECHNOLOGY (PHYSICAL ASSET MANAGEMENT)** (White, 1975).

According to **TEROTECHNOLOGY - WHAT IT IS ALL ABOUT** (Thackara, 1975) terotechnology is:

> *Terotechnology is defined here as a combination of management, financial, engineering and other practices applied to physical assets in pursuit of economic life-cycle costs: it is concerned with the specification and design for reliability and maintainability of plant, machinery, equipment, buildings and structures, with their installation, commissioning, maintenance, modification and replacement, and with feedback of information on design, performance and costs.*

As a concept, this seems to be very close to the (current) definition of asset management. Terotechnology transforms into asset management by adding the concept of risk to the definition, expanding the values beyond the purely economical and considering all aspects more integral in a holistic approach. This is illustrated in the Figure 4. In a silo like approach, only the direct costs to the departments’ budget would be considered, whereas asset management is about including the indirect costs into optimization of the asset, though the coverage of those indirect costs may depend on the maturity of asset management.

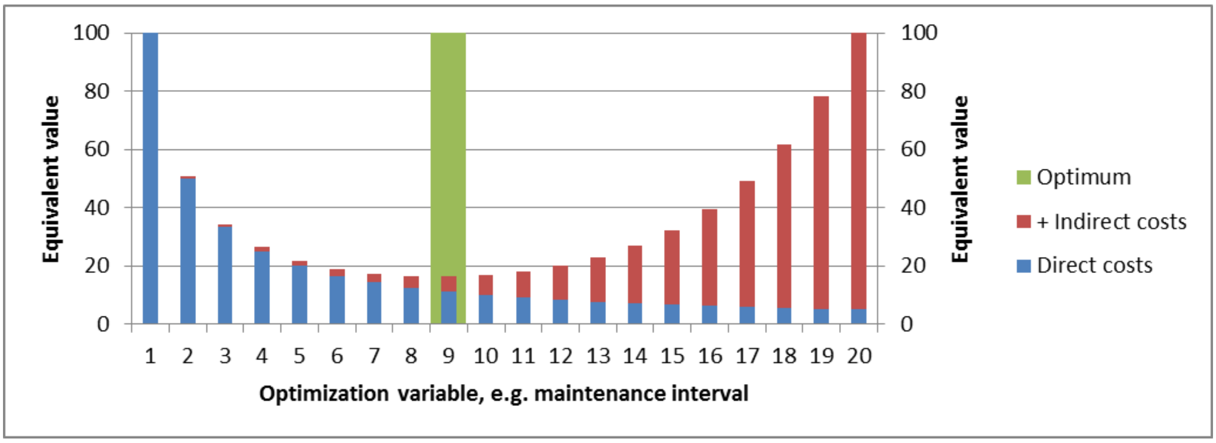


Figure 4: Optimization of total equivalent costs. The optimization variable can be the maintenance interval, replacement or upgrade moment, capacity (e.g. cable or pipeline diameter), asset quality or the like. The direct cost is the immediate cost of the asset, the indirect cost is the total of exploitation costs (energy, operation, losses), repairs and risk (in production chain and external). Planned maintenance would be a direct cost in maintenance optimization, but an indirect cost in design optimization. The optimum is achieved if the sum of those two components is minimal. Several similar examples can be found in Asset Management- an anatomy (IAM, 2014).

Searching for terotechnology however does not reveal much more documents. The Scopus database only holds 81 publications directly, and 64 secondary, dating back to 1964[49]. That record is British

---

[48] On February 23, 2015, Searching Web of Science for “Asset management” OR “assetmanagement” produced 5083 documents in total with 2185 in the core collection

Standard 3811 on maintenance management terms in terotechnology (BSI, 1964). Given that a standard needs some preceding work, the field must be older but no records were encountered.

Even though terotechnology and physical asset management are strongly linked, in scientific terms asset management has become the dominant term. The diagrams of Figure 5 show the cumulative numbers of publications on both terotechnology and asset management since 1969 as recorded in the Scopus core collection.

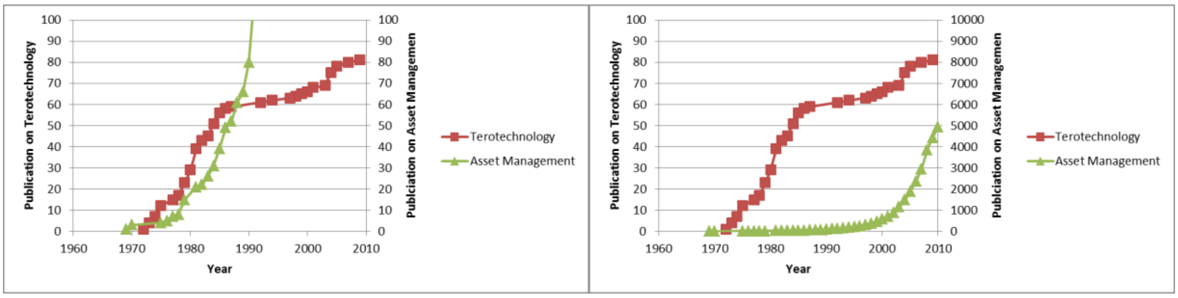


Figure 5: The total cumulative number of publications on asset management and terotechnology. The diagram on the left shows both curves at the same scale, demonstrating that terotechnology was slightly leading asset management until halfway the 1980s. Since then asset management developed much further (almost 100 times more publications) (right graph), though the 100 publications mark was only met in 1990 and the 1000 publications mark only in 2003.

Even in the professional area terotechnology seems to have died out. The last update of BS3843 occurred in 1993, with the Bureau of Indian Standards continuing their updates until 2003.

Terotechnology builds on several other disciplines. The definition of terotechnology by Thackara contains an explicit reference to designing with reliability and maintainability in mind, linking the field of terotechnology to the field of Reliability and Maintainability[50] (RAM) and its more modern descendants. These extended the scope to include aspects like Availability, Safety, Health and Environment in the concept and the acronym (Ajah, 2009, Goel, 2004). However RAM seems to focus on the technical aspects and not so much on human elements that are in scope of management systems. Originally, the objective of RAM was achieving a certain level of reliability at the lowest cost. The volume of publications on Reliability and Maintainability is comparable with that of asset management[51].

The other field to which a strong link exists is maintenance (management) and the associated prognostics and diagnostics. Many regard asset management as the professionalization of maintenance management (Wijnia and Herder, 2009) by covering the full lifecycle instead of only the operational phase. However, this view does not hold for assets like cables that are hardly maintained. Asset management then mainly is the investment decision on upgrades and

---

[49] Search conducted November 29 2013 for "Terotechnology" in title, abstract or keyword.

[50] Apparently the original name, derived from the Reliability and Maintainability Symposium, with its inaugural conference held in 1955 (62nd annual conference will be held in 2016).

[51] Scopus search on Reliability AND Maintainability in Title-Abs-Key resulted in 6337 results on feb24-2015. Searching on reliability alone yields more than 500000 documents., maintainability alone some 14000 docs.

replacements, not design of the maintenance concept. The volume of publications on maintenance is much larger than that on asset management and dates back much further[52], though many publications are in the field of medicine. Limiting the search to physical sciences in Scopus reduces the number to some 200.000 (of which some 1.000 before 1965!), much more than the number of publications on asset management. Not all publications on maintenance address management, which is also an important part of asset management. For that purpose a comparison has to be made to maintenance management. Limiting the search to maintenance and management reduces the number of entries to about 40.000, some 20% of the publications on maintenance. However, the number of publications before 1965 is only 27. Apparently, the addition of management to the field of maintenance is also only recent[53].

This observation is supported in **A FRAMEWORK FOR MAINTENANCE CONCEPT DEVELOPMENT** (Waeyenbergh and Pintelon, 2002). According to this paper, maintenance changed from a “necessary evil” only costing money to a “profit contributor” delivering value only after 1975. A similar recent development is the realization that maintenance cannot be considered as a silo, the relationships with other operating functions need to be considered. Maintenance becomes part of an integrated business concept, driven even further with outsourcing of certain maintenance activities. The relations thus need to be managed across organizational boundaries. In parallel, they see a trend from corrective maintenance, through use based maintenance to condition based maintenance, including diagnostics and prognostics. In this transition, much more attention to reliability, availability, quality, safety and the environment is given. The associated maintenance concepts developed from Reliability centered maintenance, to business centered maintenance, total productive maintenance and lifecycle approaches.

Table 3: Timeline of maintenance concepts after Waeyenbergh and Pintelon (2002).

<1950	**1950-1975**	**>1975**	**→“2000”→**
Manpower (simple)	Mechanization (complex)	Automation (more complex)	Globalization (crossing boundaries)
“Fix it when it breaks”	“I Operate – you fix” (availability, longevity, cost) PM, WO-mgnt.	RAM (Safety, Quality, Environment), CBM, CM, DOM, Multi-skilling, MMIS Asset mgnt	Optimal concept + outsourcing and ICT
Maintenance is “A production Task”	Maintenance is “A task of the maintenance dept.”	Maintenance is “(maybe) Not an isolated function”	Maintenance is “External and internal partnerships”
		Integration efforts	Maintenance meets production
“Necessary evil”	“Technical matter”	“Profit contributor”	“Partnership”
RAM: Reliability, Availability, Maintainability; PM: Preventive Maintenance; ICT: Information and Communication Technology, CBM: Condition Based Maintenance; CM: Condition Monitoring; WO: Work Order; DOM: Design Out Maintenance			

---

[52] About 500k in Scopus, more than 1M in WebOfScience, both dating back to about 1900. Scopus has better options to distinguish medicine from physical sciences, and the Scopus results were used for filtering, though WoS has older records.

[53] Though this may be only true in scientific documentation in the field of maintenance. Historical archives report much older materials. The Romans for example had an extensive system for maintaining their aqueducts, some 2000 years ago (Frontinus, 97 AD).

Furthermore, they state

> *To develop an appropriate maintenance concept, maintenance must be considered holistically. Factors that technically describe each system to maintain, as well as factors that describe the interrelations between the different systems and factors that describe the general organisational structure should be addressed. If some of the necessary aspects are not considered (e.g. due to uncareful analysis or lost data or knowledge), the maintenance concept will never reach its full potential.*

Even though maintenance is focused more on the operational phase of the life cycle of the asset, the extensions toward Design out Maintenance (DOM) and the holistic approach make it resemble asset management.

The term holistic also appeared in **CONTEMPORARY MAINTENANCE MANAGEMENT: PROCESS, FRAMEWORK AND SUPPORTING PILLARS** (Crespo Marquez and Gupta, 2006). For the definition of maintenance and maintenance management reference is made to the European Norm for Maintenance Terminology (European Committee for Standardization, 2001). According to this standard, “maintenance is defined as the combination of all technical, administrative and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function”. Maintenance management is then defined “as all the activities of the management that determine the maintenance objectives or priorities, strategies, and responsibilities and implement them by means such as maintenance planning, maintenance control and supervision, and several improving the methods including economical aspects in the organization.” According to the paper, maintenance management is frequently associated with difficulties, and a holistic approach is presented to define all maintenance management functions. The framework recognizes the strategic, tactical and operational level in activities and 3 pillars for maintenance management: IT, Maintenance Engineering and Organizational techniques. With risk analysis being an important part of maintenance engineering, these pillars very much resemble the critical enablers People Factors, Risk Management and Information Technology of asset management (Woodhouse, 2014).

The introduction of the concept of risk into asset management is also relatively recent. Before 1991 no entries were found that contained both Risk and Asset Management in the Scopus core collection. The table below contains the 10 oldest records. The topics span a number of areas, ranging from financial assets to infrastructure systems.

Table 4: 10 oldest entries with Asset Management and Risk[54].

Authors	Title	Year	Source title
Robinson R., Anderson K.	Computer systems for asset and risk management	1991	National Conference Publication - Institution of Engineers, Australia
Morris R.L., Lafitte Jr. F.	Consolidating and managing a mature portfolio	1991	
Jarvis M.G., Hedges M.R.	Use of soil maps to predict the incidence of corrosion and the need for iron mains renewal	1994	Journal of the Institution of Water and Environmental Management
Slipper M., Whipp S.	Integrated rehabilitation of water distribution systems	1994	Water Supply
Johnson S.	Strategic issues for management, reclamation and remediation of land in the mining and extractive industries	1995	Environmental Protection Bulletin
Yao Yulin, Cheng John F., Enny Philip, Guo Duanyang	Toward parallel financial computation: valuation of mortgage-backed securities	1995	Proceedings of the IEEE International Conference on Systems, Man and Cybernetics
Steed John C.	RTDE '95 - the importance of equipment reliability	1995	Power Engineering Journal
Robbens E.G.	Asset management in the construction and facilities management environments	1995	IEE Colloquium (Digest)
McMahon B.	Reliability and maintenance practices for Australian and New Zealand HV transmission lines	1995	IEE Conference Publication
Mulvey John M.	Solving robust optimization models in finance	1996	IEEE/IAFE Conference on Computational Intelligence for Financial Engineering, Proceedings (CIFEr)

After the somewhat delayed start, risk management has become a significant part of asset management, with some 10% of the papers on asset management containing a reference to risk, as demonstrated below.

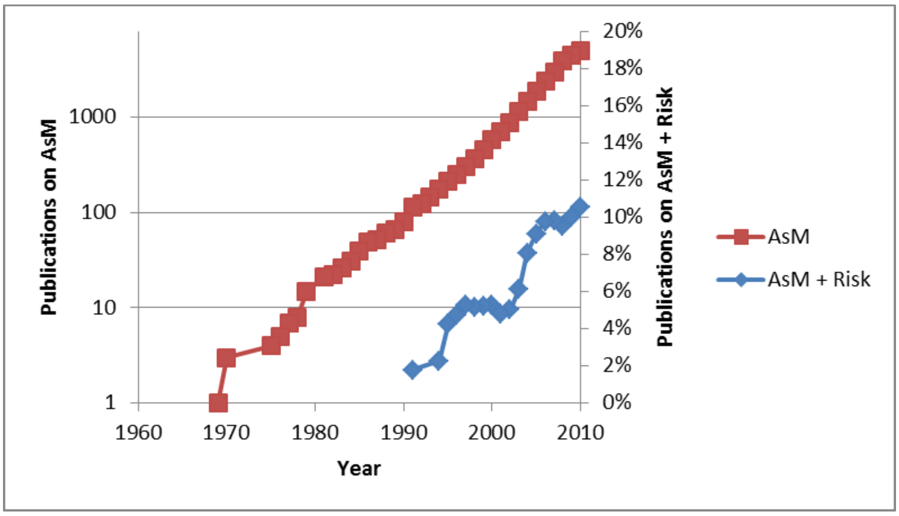


Figure 6: Publications on risk in asset management, compared to general asset management. Since about 2005 the fraction of papers on risk is about 10%. Please note the logarithmic scale for the number of AsM publications.

---

[54] (TITLE-ABS-KEY ("asset management") AND TITLE-ABS-KEY (risk)) AND SUBJAREA (mult OR ceng OR CHEM OR comp OR eart OR ener OR engi OR envi OR mate OR math OR phys) AND (LIMIT-TO (SUBJAREA , "ENGI") OR LIMIT-TO (SUBJAREA , "ENER")). This search conducted at March 4 2015 resulted in 764 records in the core Scopus collection.

Since about the turn of the century, asset management really developed into a scientific discipline with its associated community. In 2006, the first World Conference on Engineering Asset Management (WCEAM) was held in Brisbane, Australia. This conference is held annually, with some 100-200 papers per conference, forming a substantial part of the Body Of Knowledge on asset management. Additionally, the International Journal On strategic Engineering Asset Management (IJSEAM) has been established in 2012.

3.1.2 Professional developments

In the 1980s, asset management was picked up professionally in Australia (Burns, 2010). Several initiatives were employed, like the formation of task forces, conferences and several manuals that were published in the 1990s, like the Total Asset Management Manual (New South Wales et al., 1993) and the National Asset Management Manual (Institute of Municipal Engineering Australia, 1994). The focus was strongly on public assets, contrasting the more industrial focus of terotechnology. In the same period, the New Zealand Infrastructure Asset Management Manual was developed. Combining efforts with Australia resulted in the International Infrastructure Management Manual (IIMM for short) in 2000 that was adopted by the UK Institute of Asset management (IAM, 2002).

In parallel, a comparable initiative took place in North Sea oil industry. In the 1980s, the industry saw a significant reduction of profitability as a result of decreasing oil prices. On top of this, one of the worst offshore incidents occurred at Piper Alpha, killing some 167 people. In an inquiry to the causes of this incident, the Cullen Report (Cullen, 1990) concluded that this was due to a lack of systematic attention for risks. According to the IIMM, this resulted in a transition from a prescriptive regime to a more goal oriented approach with a mandatory safety case (IAM, 2002) with monitoring by the Health and Safety Executive. This can be regarded as the introduction of risk thinking into asset management (Woodhouse, 2014). Futhermore, to deal with the financial challenges the CRINE[55] initiative was employed, which resulted in ageing assets being kept in operation and thus reducing capital requirement. Over the years, the objective of reducing costs per unit further developed into operational excellence(Lynn, 2002)[56]. However, some claim that the CRINE initiative counteracted the recommendations from the Cullen report, as the rate of safety incidents did not drop after implementing the Cullen recommendations (Tombs and Whyte, 1998). Yet, in the field of safety the role of the organization has been increasingly recognized, instead of human errors of not complying with the rules (Abraha and Liyanage, 2015). Furthermore, recent years have demonstrated a significant drop in the number of safety incidents (Woodhouse, 2014) and a move towards inherently safe designs (Singh et al., 2010). Piper Alpha was not the only landmark risk event that occurred in the North Sea. In 1995, Shell got into a battle with Greenpeace over the disposal of an obsolete oil rig, the Brent Spar. Greenpeace managed to get the public involved by claiming the rig contained much more toxic materials than Shell stated, which resulted in a 20% drop of sales at the Shell petrol

---

[55] “Cost Reduction In A New Era”, Or “Cost Reduction Initiative for a New Era”, depending on source.

[56] An interesting notion in this presentation was that of the phantom asset, which was defined by the asset producing the losses. According to Lynn, the phantom refinery, containing all losses from all BP refineries, was the largest refinery BP had.

stations in Germany. This forced Shell to give in, despite having the facts on their side with regard to the contents of the rig (Löfstedt and Renn, 1997, Bakir, 2005).

The third major stream of professional development also occurred in the UK. Following liberalization, the UK utility sector discovered asset management as a means for further improvement and established the Institute of Asset management in 1994. Members of this organization were very active in the commercial conferences on asset management, organized since the end of the last century. Their contributions inspired many Dutch utilities to embark on the asset management journey as well. As mentioned in the introduction to this thesis, the first "standard" of asset management, PAS 55[57] (BSI, 2004a, BSI, 2004b), also originated from this institute. In its first edition, the specification was limited explicitly to infrastructure assets. In the second edition of PAS 55 (BSI, 2008a, BSI, 2008b), released in 2008, this limitation was dropped, and the specification applied to all physical assets. The 2008 version of PAS 55 was furthered into the Dutch norm NTA8120:2009 (NEN, 2009) specifically for the electricity and gas distribution network operators. As of January 2013, Dutch grid operators Alliander, Delta, Enexis, Rendo, Stedin and TenneT were certified against this norm. These grid operators serve the vast majority of customers in the Netherlands (97% for E, 89% for G)(Energiegids, 2013).

PAS 55:2008 was also furthered into an international standard, the ISO 55000 series[58] (ISO, 2014a, ISO, 2014b, ISO, 2014c). Whereas the 2008 version was an upgrade of the original version, the ISO version could be regarded as a redesign. This was because the ISO 55000 series was the first management system standard to be written according to the ISO template for management system standards. The idea behind this template was to facilitate organizations using more than one ISO management system standard in streamlining their processes. The template provided an opportunity to have a fresh look at the standard, instead of adhering to the original PAS 55 formulation. This also allowed for the introduction of new ideas and definitions (Hodkiewicz, 2015).

The use of the template resulted in several changes in the content of the standard. First of all, the potential scope of the asset management system has been enlarged, to include all types of assets, not only physical assets. A second noteworthy change occurred in the definition of asset management. ISO 55k defines asset management as follows (ISO, 2014a):

> *"Asset management" is the coordinated activity of an organization to realize value from assets*
>
> *Note 1: Realization of value will normally involve a balancing of costs, risks, opportunities and performance benefits*

This is a significant cleanup from the (longer) definition of asset management in PAS 55. The concept of an optimum over cost, risk and performance is no longer part of the definition, but has been moved to a note to the definition. In this move the formulation also has changed from optimizing to

---

[57] The specification consisted of two parts: PAS 55-1 containting the specification, and PAS 55-2 for guidelines for the application. If PAS 55 is used, both parts are meant.

[58] The norm consists of 3 parts: 55000, terminology, 55001, requirements and 55002, guidelines. In this thesis, the abbreviation ISO55k will be used to refer to the series. If a specific part is meant, the full number will be used.

balancing (which is not necessarily the same, even though many asset managers may regard them as equal) and the aspects have been expanded with opportunities.

Additionally, ISO 55k has a formal definition of risk, where PAS 55 only had a definition of risk management. The definition of risk that ISO 55k uses is that of ISO guide 73 (ISO, 2002), also used by ISO 31000 to which ISO 55001 explicitly makes reference (ISO, 2014b):

> *Risk is the effect of uncertainty on objectives*[59]

Unfortunately, this definition is not without discussion. In a review of the ISO 31000 standard (Leitch, 2010), the ambiguity and imprecision of many formulations is noted, with special attention to the definition of risk:

> *Taken literally this suggests a radical new focus on the way objectives are formulated but it is almost certain that the intended meaning is something else. It is something to do with the potential effect of events that are currently uncertain on the extent to which objectives are achieved.*

Furthermore, the definition does not relate to other common definitions of risk used in risk analysis literature. In a review on the development of the risk concept, Aven (2012) needs a separate category for the ISO definition as it does not match any of the other categories, which all contain multiple formulations. In the same publication the ISO definition is regarded as imprecise, further substantiated in a separate publication (Aven, 2011).

In hindsight, these professional developments seem to follow a similar pattern. The concept is adopted, developed at a high intensity during a short period, until a level is reached that is good enough in practice. Beyond this point, development tends to slow down, because the benefit of improvements is not very clear[60]. These improvements (if any) address refinements of current practice, like continuous improvement advocated by the management standards. But continuous breakthroughs based on deeper understanding of what is happening does not seem to occur. A similar comment was made by Jonsson (Jonsson, 2000) on the development of maintenance management:

> *Although, proper maintenance approaches exist, neither maintenance practice nor theory are fully developed.*

One of the omissions is the

> *[….] lack of maintenance management configurations, such that could be useful to improve the understanding of the underlying dimensions of maintenance, and that could explain the effects of preventive maintenance and integrating maintenance into manufacturing.*

If asset management is to be developed into a true science, such a deeper understanding should be at the core of such a research program.

---

[59] Note 1 to this definition is about effect: “An effect is a deviation from the expected — positive and/or negative”. Apparently, risk in the standard is not limited to bad things, but can be about upsides as well. That seems to overlap with the inclusion of opportunities in note 1 of the definition of asset management.

[60] As argued by Hodkiewicz (2015), the value of complying with ISO55001 itself is also not very clear, nor is that of any management system.

3.1.3 Current state of asset (risk) management[61]

The table below gives the 10 most cited articles on asset management and risk.

Table 5: Most cited sources in Scopus for "asset management" and "risk" in engineering (search date March 4 2015). Only the 8th and 9th entry are not about infrastructure assets. The number of citations is very low compared to many other sciences. Apparently these is no standard paper yet.

Authors	Title	Year	Source title	Cited by
Brown R.E., Humphrey B.G.	Asset management for transmission and distribution	2005	IEEE Power and Energy Magazine	50
Schneider J., Gaul A.J., Neumann C., Hografer J., Wellssow W., Schwan M., Schnettler A.	Asset management techniques	2006	International Journal of Electrical Power and Energy Systems	47
McGill W.L., Ayyub B.M., Kaminskiy M.	Risk analysis for critical asset protection	2007	Risk Analysis	35
Selih J., Kne A., Srdic A., Zura M.	Multiple-criteria decision support system in highway infrastructure management	2008	Transport	32
Christodoulou S., Deligianni A.	Neurofuzzy decision framework for the management of water distribution networks	2010	Water Resources Management	31
Stewart M.G., Rosowsky D.V., Val D.V.	Reliability-based bridge assessment using risk-ranking decision analysis	2001	Structural Safety	27
Abu-Elanien A.E.B., Salama M.M.A.	Asset management techniques for transformers	2010	Electric Power Systems Research	23
Feng D., Gan D., Zhong J., Ni Y.	Supplier asset allocation in a pool-based electricity market	2007	IEEE Transactions on Power Systems	23
Emmanouilidis C., Liyanage J.P., Jantunen E.	Mobile solutions for engineering asset and maintenance management	2009	Journal of Quality in Maintenance Engineering	22
Moglia M., Burn S., Meddings S.	Decision support system for water pipeline renewal prioritisation	2006	Electronic Journal of Information Technology in Construction	22

All of these are relatively recent, with the oldest in this list dating back only to 2001. These publications address several topics, though most of them are related to infrastructure issues, and especially decision support, decision making and the like.

Interestingly, according to the most cited entry (Brown and Humphrey, 2005),

> *Risk management is perhaps the most misunderstood aspect of asset management.*

They state that executives view the aspect as financial risk management, to which physical risk management is seemingly unrelated as it is about undesirable events. Those physical risks are most often only addressed in project approval, in terms of what could go wrong if the project was not approved, but that misses the risk of all the other assets. However, their suggestion is to think of risk as not meeting performance targets, similar to the (interpreted) ISO 55k definition. This is followed by a proposal to include confidence intervals in performance targets. However, their view on risk misses the event notion characteristic for physical risk.

---

[61] The purpose of this section is to provide an overview of the body of knowledge used in the science of asset management. It is therefore based on a citation count, which inevitably is biased towards older publications. Newer publications may be better aligned with the research in this thesis, but they are (not yet) generally recognized.

The second paper on the list regards asset management especially for electricity grids. In the paper several techniques like RCM, statistical fault analysis and simulation are mentioned. However, their view of asset management as defined in "Asset management means operating a group of assets over the whole technical life-cycle guaranteeing a suitable return and ensuring defined service and security standards" does not align with the idea of asset management being a management system.

Of the two publications in this list not on infrastructures, **SUPPLIER ASSET ALLOCATION IN A POOL-BASED ELECTRICITY MARKET** (Feng et al., 2007) considers risk in a classical financial perspective, though it is on power systems. The paper **MOBILE SOLUTIONS FOR ENGINEERING ASSET AND MAINTENANCE MANAGEMENT** (Emmanouilidis et al., 2009) considers assets in a much broader range, and focuses on the application of wireless solutions for remote management complex, high risk and capital intensive asset. This allows the formation of knowledge networks regardless of the geographical location, so that decision support can be provided to local operators doing the actual maintenance. However, no mention is made to what is meant by a high-risk asset, nor if the cost of remote monitoring is justified by the risk reduction achieved.

Also interesting is the third paper in the table on **RISK ANALYSIS FOR CRITICAL ASSET PROTECTION** (McGill et al., 2007). The paper provides a framework for risk analysis for terrorism for critical assets, including an example with a fully monetized cost benefit analysis[62]. The authors also make a notion of the method being applied to a portfolio of assets. They conclude the total risk of the portfolio can be obtained by summing the risks of the assets, but that this does not account for interdependencies between assets. These are likely to exist, as adversaries are likely to shift their attention to weaker assets in the portfolio once protection is employed for some, thus influencing the risk profile. Furthermore, they note that their framework is for strategic risk management, that is supporting investment decisions, though it can be modified to support tactical risk management.

With regard to decision support in asset management, Gulski (18 contributions) and Smit (12 contributions) contributed to the largest number of publications. Their field of interest is assets in the electrical transmission and distribution systems, especially condition monitoring. Their most cited paper, **PD KNOWLEDGE RULES FOR INSULATION CONDITION ASSESSMENT OF DISTRIBUTION POWER CABLES** (Gulski et al., 2005), on providing support for selecting the right components for inspection, maintenance or replacement. However, their assessment rule is formulated in technical terms, not the values typical for asset management.

Summarizing the findings on asset risk management, there is no consistent view on risk. On the one hand it is claimed to be the most misunderstood aspect of asset management, but on the other hand some highly quantified framework is presented. But there are also publications that are about risk without specifying what is meant. Given that the definition of risk in the asset management standard is criticized for not being precise, some further research into risk and risk analysis is needed.

---

[62] In their Expected loss table, fatalities are not monetized, but a monetization factor is given for fatalities. Unfortunately, no source is given for the loss conversion factors.

3.2 Risk and risk management

Risk, as a concept, can be traced several 100s of years back (Aven, 2012). In scientific literature, accessible by scientific search engines like Scopus and Web Of Science, records date back to 1832[63]. The majority of these publications is in the field of medicine (roughly 50%), whilst engineering only accounts for some 160.000 records[64], about 5% of total. Within this engineering subset, the journal Risk Analysis is the largest contributor, with more than 3000 contributions. However, the Society for Risk Analysis, publishing this journal, was only established in 1980, with the journal starting in 1981. Many other large contributors seem to focus on specific fields of risk (e.g. safety, reliability) or specific fields of engineering (optical engineering, advanced materials, applied mechanics, transportation). This suggests that systematic research on risk and risk management in general is also relatively recent.

Limiting the scope of the subset to risk management reduces the number of records by a tenfold to about 17000 entries, with only 17 entries preceding 1981. Many papers within this risk management subset are about managing specific risks, and not so much about the activity, framework or methodology of risk management. In browsing some highly cited papers in the risk management subset, the observation can be made that if the risk management process is mentioned, it is generally only by naming the phases with only little (if any) reference to scientific literature (Jonkman et al., 2003, Haimes et al., 2002, Hallikas et al., 2004). The table below relates the 4 phases of Jonkman to the 4 phases of Hallikas and the 6 questions of Haimes (numbering of steps by author).

Table 6: Relating several definitions of the risk management process.

Step	Jonkman	Hallikas	Haimes
1	Qualitative analysis: definition of system and scope, identification of hazards, failure modes and scenarios		
2		Risk identification	What can go wrong?
3	Quantitative analysis: probabilities and consequences	Risk assessment	What is the likelihood of that happening? what are the consequences?
4	Risk evaluation: decision on risk tolerability		
5		Decision making and implementation of risk management actions	What are the available options? What are the associated tradeoffs? What are the impacts of current decisions on future options?
6	Risk control and risk reduction measures: determining the measures to reduce the risk, control (inspection, maintenance warning systems) of the risk		
7		Risk Monitoring	

Even though there seems to be some agreement on the order of the activities or questions, there is no agreement on the precise content of each of the phases. Except for the border between step 2 and 3, separating identification from analysis, everything else is only partially aligned. This suggest there is no well documented, scientifically supported definition of the risk management process, only interpretative descriptions of common practice.

---

[63] In a basis search on Risk, Scopus revealed some 2,7 million records dating back to 1832, Web of Science revealed 5 million records dating back to 1883. Further remarks with regard to the body of knowledge are based on the Scopus collection.

[64] 163.478 on March 31 2015.

With regard to good practice, other sources are available. The ISO standard on asset management refers explicitly to the ISO standard on Risk Management, ISO 31000[65]. Another description of good practice is the COSO[66] framework for enterprise risk management. This framework was in place at Enexis during the experiments[67]. These 2 descriptions of good practice will be reviewed in some more detail[68]. The differences between them (supplemented with the differences of table 6) will be used to identify the relevant scientific debates for the definition of the risk management process. For each of these debates the relevant perspectives will be reviewed.

3.2.1 Frameworks for risk management

According to ISO 31000, risk management is “the coordinated activities to direct and control an organization with regard to risk”. The standard starts with the principles, from which a risk management framework is derived. Risk management is implemented by means of a risk management process, which exists of a number of steps. Figure 7 shows those steps and their relations.

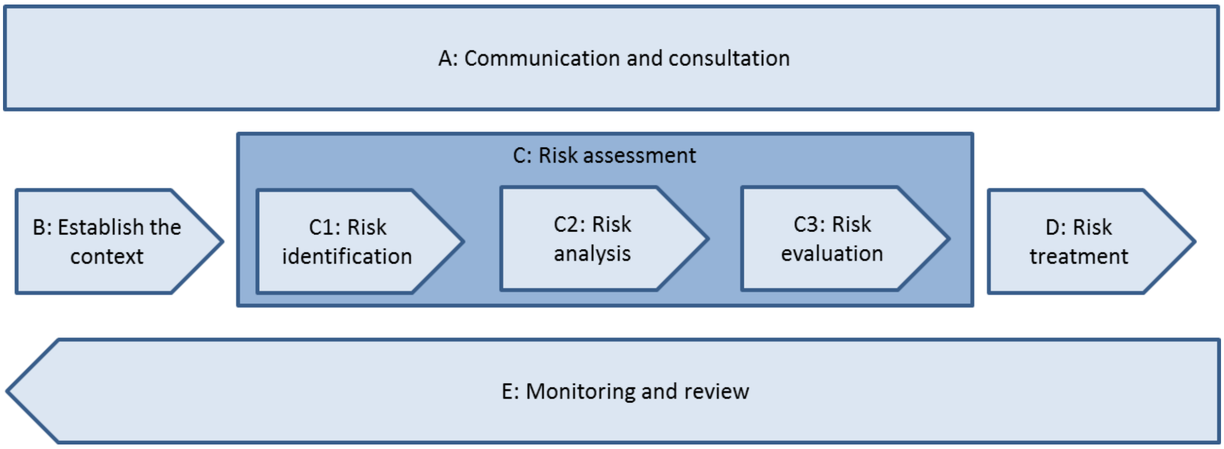


Figure 7: The risk management process after ISO 31000 (ISO, 2009a).

Risk management starts with establishing the context in which the risks are managed. This is amongst other things about defining the risk criteria and the approach to be followed in decision making on risk. After establishing the context, risk are identified, analyzed and evaluated. These three steps are grouped into the activity Risk Assessment. If the risk exceeds the criteria, treatment

---

[65] The accompanying document ISO 31010 (ISO, 2009b) on risk management techniques contains many examples of practical tools (including a short description), like RCM, FMECA and RPN. Description of these tools is outside the scope of this section, but will be part of chapter 7 and 10.

[66] Committee of Sponsoring Organizations of the Treadway Commission (COSO).

[67] According to https://www.enexis.nl/over-enexis/investor-relations/en/corporate-governance/risk-management, accessed March 31 2015, it still is in place.

[68] The selection of these two frameworks may seem arbitrary, as many others exist. However, the two frameworks are used in a differential analysis. The contrast between them is large enough to identify the relevant scientific debates.

options are developed, selected and implemented. Development of risks and effectiveness of treatments is monitored, reviewed and fed back if necessary (hence the reversed arrow). To align with the stakeholders, communication and consultation takes place within and over all activities.

The COSO framework (COSO, 2004b) as shown in Figure 8 defines risk management as a process:

> *Enterprise risk management is a process, effected by an entity's board of directors, management and other personnel, applied in strategy setting and across the enterprise, designed to identify potential events that may affect the entity, and manage risk to be within its risk appetite, to provide reasonable assurance regarding the achievement of entity objectives.*

The framework recognizes several types of objectives (strategic, operational, reporting and compliance), several layers of the organization (entity, division, business unit and subsidiary) and uses 8 steps in the risk management process:

1. Internal Environment
2. Objective Setting
3. Event Identification
4. Risk Assessment
5. Risk Response
6. Control Activities
7. Information & Communication
8. Monitoring

The types of risks, layers of the organization and the steps of the process are often represented as sides of a cube to demonstrate they are different but interrelated views on enterprise risk management as shown in Figure 8.

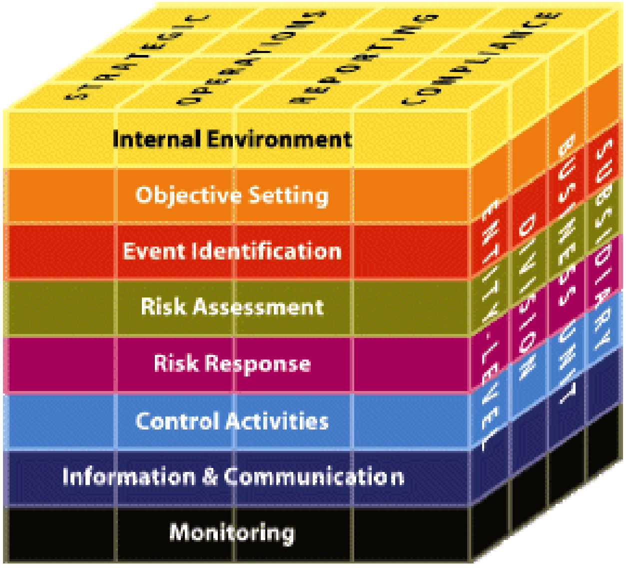


Figure 8: the COSO framework, after (COSO, 2004b).

3.2.2 Comparing the frameworks

COSO and ISO are more alike than they are different. Most differences are in organization of the content, not in the content itself. The COSO definition for example is longer than the ISO definition, but it reflects some concepts that are very similar to that what ISO regards as principles.

Table 7: Comparing ISO and COSO.[69]

Process step	**ISO**		**COSO**
I	B Establish the context		1 Internal Environment
II			2 Objective Setting
III	C risk assessment	C1 Risk identification	3 Event Identification
IV		C2 Risk analysis	4 Risk Assessment
V		C3 Risk evaluation	
VI	D Risk treatment		5 Risk Response
VII			6 Control Activities
VIII	A Communication and consultation		7 Information & Communication
IX	E Monitoring and review		8 Monitoring

With regard to the organization of the risk management process, ISO and COSO are also very much alike, as can be seen in table 7. The first impression may be that ISO is more compressed in the organization of the process than COSO, as several ISO activities bundle two COSO activities. *Establishing the context* of ISO equals *Internal Environment* plus *Objective Setting* of COSO, *Risk assessment* covers both *Event Identification* and *Risk Assessment*, and *Risk treatment* is both *Risk Response* and *Control Activities.* On a more detailed level, however, ISO also splits a single COSO activity. COSO *Risk Assessment* equals ISO *Risk analysis* and *Risk evaluation*[70]*.* The net result thus is that ISO contains 7 steps and COSO 8 steps. A similar superficial difference can be seen in the way the standards deal with communication. ISO explicitly regards communication as an activity affecting the whole process, whereas in COSO it is an activity in the process. COSO however recognizes that risk management is not a strictly serial process and that any component may influence the other component in a multidirectional and iterative process. All in all, the processes thus are very much alike and both processes can be mapped on each other.

However, there are some more significant differences as well. The first major difference considers the concept of risk that is used. COSO limits identification to “internal and external events affecting achievement of an entity’s objectives […] distinguishing between risks and opportunities”. Opportunities then exit the flow and only risks continue through the process. Risk thus appears to be reserved for events with negative impacts. In ISO, identification considers a broader range than events only:

> *The organization should identify sources of risk, areas of impacts, events (including changes in circumstances) and their causes and their potential consequences. The aim of this step is to generate a comprehensive list of risks based on those events that might create, enhance, prevent, degrade, accelerate or delay the achievement of objectives.*

This clearly differentiates risks from events. Furthermore, there is no distinction between risks and opportunities, as the list of risks is also based on events that enhance achievement of objectives, an opportunity in COSO terms. Opportunities are addressed in ISO, but only secondary as “the risks

---

[69] To align ISO and COSO 9 process steps are needed, whereas both frameworks have less steps themselves: COSO 8 and ISO 7.

[70] Interestingly, ISO and COSO use the term risk assessment for different concepts: in COSO it is limited to analysis and evaluation, whereas ISO expands the activity to include identification as well. This precisely creates potential for miscommunication that standards are supposed to prevent.

associated with not pursuing an opportunity". Risk in terms of COSO thus is limited to downsides, whereas ISO regards any deviation from the planned course as a risk, whether positive or negative[71].

The second major difference between the frameworks can be found in risk decision making. Both start with selecting risk for development of treatment options, using concepts like acceptable and intolerable risk. This is followed by selecting the best treatment option for these risks, in which types of responses like avoidance, reduction, sharing and acceptance have to be considered[72]. However, the aspects to be considered differ. ISO decision making on the treatment is about balancing costs and benefits of treatments, considering multiple values and stakeholder views (and even stakeholder involvement in decision making) on acceptability of risk and treatment and the resulting residual risk[73]. Even the need to implement economically unjustifiably treatments is mentioned. As a contrast, COSO seems to limit the scope of the decision to the costs and benefits of the entity. That is, no reference is made to any external stakeholders or non-entity values (like social responsibility) as ISO does. On the other hand, COSO explicitly mentions the portfolio view, stating that the overall residual risk should be within the entity's risk appetite. ISO does not require a portfolio approach, it only recognizes the potential benefit of combining several treatment options. Even though nothing in ISO forbids a portfolio view, the whole seems to be more focused on dealing with individual risks. ISO thus considers more aspects per individual risk, but not so much the combined effort of all risks.

Both these points will be reviewed from a scientific point.

3.2.3 The theory on the concept of risk

Despite its importance, there is no consistent view on risk in neither asset management nor in risk management. Unfortunately, this is also the case in the scientific field of risk analysis. According to Aven (2012) "There is no agreed definition of the concept of risk." Many have reached similar conclusions by listing several conflicting definitions of risk. Slovic and Weber (2002) for example list four definitions, whereas Beer and Ziolkowski (1995) list thirteen definitions. In his paper **THE RISK CONCEPT—HISTORICAL AND RECENT DEVELOPMENT TRENDS**, Aven (2012) traces the use of risk back more than 300 years and combines the many uses into a list of 9 categories, as shown below.

---

[71] However, according to Leitch (2010) this inclusion of upsides in the concept of risk is not maintained throughout the standard, as the language used in the section on risk treatment only makes sense if risk refers to downsides, like COSO does.

[72] ISO is not very precise in this respect. It first states that treatment options have to be considered for unacceptable risks, effectively removing acceptance as a viable option. But one of the treatment options to be considered according is retaining the risk, which would only make sense if the risk was not unacceptable to start with.

[73] Residual risk is a concept that only makes sense for negative impacts, contradicting the claim of ISO 31000 that is it also about upsides.

Table 8: Different definitions of the concept of risk, after Aven (2012).[74]

NR	Definition of the concept of risk
1	Risk=Expected value (loss) (R=E)
2	Risk=Probability of an (undesirable) event (R=P)
3	Risk=Objective Uncertainty (R=OU)
4	Risk=Uncertainty (R=U)
5	Risk=Potential/possibility of a loss (R=PO)
6	Risk=Probability and scenarios/Consequences/severity of consequences (R=P&C)
7	Risk=Event or consequence (R=C)
8	Risk=Consequences/damage/severity of these + Uncertainty (R=C&U)
9	Risk is the effect of uncertainty on objectives (R=ISO)[75]

In the paper, the hypothesis is formulated that

> *there has been a gradual change from rather narrow risk perspectives based on probabilities and expected values, to broader not-probability-based definitions with a sharp distinction between risk as a concept and how this concept is measured.*

The arguments for this distinction between the concept and the metric[76] can be found in the conclusion of **ON THE ONTOLOGICAL STATUS OF THE CONCEPT OF RISK** (Aven et al., 2011): Risk exist as a concept, even when it is not quantified. Measurement is needed for risk assessment, but not for the general concept.

> *"Risk should also exist as a concept without modeling or any other tool. We face risk when we drive a car or run a business even if the probabilities are not specified. For risk assessment we need the probabilities, but not as a general concept of risk. In this way we obtain a sharp distinction between risk as a concept (a), and risk descriptions (assessments) (c) which could be based on models (b).*

To underline this distinction between the concept and the way it is measured, it is proposed to use uncertainty in the definition of the concept, and to use probability as a potential measurement of uncertainty (Aven and Renn, 2009, Aven, 2010). The main argument is that it is useless to speak about the probability (in the meaning of relative frequency[77]) of unique events.

---

[74] This tables list definitions of the concept of risk. This is different from the perception of risk, as discussed on the following pages.

[75] As mentioned in the comment on the ISO 31000 definition of risk, all of these categories except the ISO definition contain several formulations as an example.

[76] In chapter 11 of this thesis, a similar distinction is made. In risk identification workshops, the term risk was used for the things that could go wrong, whereas the risk level (high, medium, low etc) was used to express the combination of probability and severity of the consequences. To emphasize the difference between risk and risk level, the latter was reformulated as exposure.

[77] In the theoretical debate about the definition of risk, even the concept of frequency is challenged. Cox claims that there is no definition of frequency possible that satisfies two basic conditions (Cox, 2012), though the conclusion is rejected both from a more theoretical viewpoint (Yellman, 2012) and from a practitioners viewpoint (Garrick, 2012).

3.2.3.1 Objective versus subjective risk

The argument that risk exists as a concept, even when it is not quantified suggests an objectivists approach to the concept of risk, where risk is part of reality and the consequences and likelihood of adverse events can be objectively quantified.

In their paper **PERCEPTIONS OF RISK POSED BY EXTREME EVENTS,** Slovic and Weber (2002) take a strong position against the assumption of objective quantification in risk assessment.

> *Much social science analysis rejects this notion [of risk being objectively quantified by risk assessment], arguing instead that such objective characterization of the distribution of possible outcomes is incomplete at best and misleading at worst. These approaches focus instead on the effects that risky outcome distributions have on the people who experience them. In this tradition, risk is seen as inherently subjective [..] It does not exist "out there," independent of our minds and cultures, waiting to be measured. Instead, risk is seen as a concept that human beings have invented to help them understand and cope with the dangers and uncertainties of life. Although these dangers are real, there is no such thing as "real risk" or "objective risk."*

In many cases (like the Brent Spar mentioned previously) the public response to risks is not aligned with the objective assessment. In a paper on the attitudes toward technological risk Fischhoff et al. (1987) link the expressed preference (of lay people) for risks to various characteristics of those risks, and conclude that many of the characteristics are correlated and that all can be expressed into two fundamentals, the "unknown risk" factor and the "dread risk" factor. According to this paper, the risk perception of lay people is strongly related to the dread risk factor.

In the paper **PERCEPTION OF RISK** Slovic (1987) contrasts this with expert risk perception which tend to correlate with expected value. A main conclusion of the paper is that "[..]'riskiness' means more than 'expected number of fatalities'." Furthermore, it mentions that some risk debates are not about the risk itself, but that risk is used as an excuse for other concerns:

> *Research implies that some of these debates may not even be about risk [..]. Risk concerns may provide a rationale for actions taken on other grounds or they may surrogate for other social or ideological concerns. When this is the case, communication about risk is simply irrelevant to the discussion. Hidden agendas need to be brought to the surface for discussion [..]*

This observation applies very much to the Brent Spar. The debate was on distrust about the contents of the Brent Spar, general distrust in oil companies, the idea of using the ocean as a dumping ground, amongst many other things, but not on the actual risks in any of the options for dealing with the Brent Spar.

3.2.3.2 The constructive approach

The criticisms against "objective risk" are also present in the report **UNDERSTANDING RISK: INFORMING DECISIONS IN A DEMOCRATIC SOCIETY** (Stern et al., 1996), on the process of risk characterization (regarded as an analytic deliberative process) as a prelude on decision making on risk where they state that

> *"Risk characterization should be a decision-driven activity, directed toward informing choices and solving problems."*

In the section on the idea of risk characterization this is furthered by

> *"the purpose of risk characterization is to enhance practical understanding and to illuminate practical choices"*

However, nowhere in the report a formal definition of risk is given. On the contrary, in the section on Principles for risk Characterization they explicitly state that

> *"Relevance to a decision, and therefore to a risk characterization, cannot be determined a priori by a formal definition of risk."*

Risk, in this view, is then more like a potential problem that requires decision making. If someone regards something as a risk, it is worth the debate.

3.2.3.3 A pragmatic resolution

Even though risk is not formally defined in the report by Stern et al. (1996), the six cases on risk analysis and characterization are all about the impact human activities have on other humans and the environment. This can be formalized into a definition of risk, for example as given by Klinke and Renn (2002):

> *We define risk as the possibility that human actions or events lead to consequences that harm aspects of things that human beings value [..]*

Klinke and Renn recognize that in economic theory risk refers both to gains and losses. However, given that they are discussing risks to human health and the environment they believe the limitation to negative impacts is "more in line with the average understanding of the concept of risk". Furthermore, they note that whether an impact is positive or negative is a human judgment which is inherently subjective and not to be derived from the impact itself.

Other strong arguments for keeping upside (i.e. opportunities) and downside risks apart can be found in the prospect theory of Kahneman and Tversky (1979). People tend to value gains differently than losses. This even holds if the difference between gain and loss is only the formulation of the decision problem (Tversky and Kahneman, 1981). This suggest that in decision making on a portfolio of decision problems, opportunities and risks will not be treated equally and options for manipulation exist if risks are formulated as opportunities (like the opportunity to mitigate a risk) and vice versa (the risk of a missed opportunity). The viable option then is to limit portfolio decision to either upside or downside risks. Given that risks are generally understood as negative, a limitation to downside risk only seems the reasonable choice. This view will be used throughout this thesis: Risk is the potential for undesired outcomes to happen.

3.2.4 Theory on risk decision making

As mentioned, the two reviewed frameworks on risk management, ISO 31000 and the COSO, agree to a large extent on risk decision making. Both use the concept of intolerable risk, which should be mitigated until the residual risk is within tolerance. Furthermore, they both regard risk decision making as weighing costs and benefits of the risk treatment options. In more detail the frameworks differ. COSO seems to take an internal business/economic perspective, whereas ISO regards much more aspects as relevant for decision maker, including stakeholders views on risks and their view on acceptability of potential measures. ISO also considers the need to implement economically unjustifiably treatments, like when no efficient intervention exists for intolerable risks.

3.2.4.1 Available approaches

This recognition of other viewpoints than only the economical is not new. In an overview of available methods for social risk management Merkhofer (1987) recognizes "three distinct, internally consistent theories of normative decision making: cost benefit analysis, decision theory and social choice." These theories are relevant in several aspects. First of all, they serve "as a source of (implicit) rationales". Their second purpose is to provide a basis from which to derive the procedures to be used in the various decision aiding approaches. And thirdly, they form the basis from which to derive criteria for evaluating approaches as right or wrong. The table below summarizes the most relevant differences between these theories.

Table 9: Comparing decision making theories, based on table 5 of Decision Science and Social Risk management (Merkhofer, 1987). Additions (scope, criterion and timing) come from same source.

	Cost Benefit Theory (CBT)	**Decision Theory (DT)**	**Social Choice Theory (SCT)**
Conceptual basis	Economic efficiency	Axioms of individual choice	Axioms of social choice
Method of analysis	Comparison of aggregated value of estimated consequences of alternative actions	Determination of logical implications of alternatives, information and preferences of decision maker	Derivation of group decision from acceptable mechanisms for incorporating individual preferences
Perspective on value	Total monetary equivalent as determined by economic actors in a free market	Responsibility of decision maker, objective is consistency	Social preference derived from "equitable" synthesis of preferences of impacted parties
View of uncertainty	Objective characterization of environment	Subjective beliefs of the individual	Product of individuals coping with erratic environment
Scope of alternatives	Often only yes or no for a specific option	Comparing multiple options	
Decision criterion	Maximize expected net present value	Maximize expected utility	Maximize expected welfare
Valuing outcomes over time	Discount future outcomes		
Perspective of decision maker	Technical (impersonal)	Individual	Group/social

According to Merkhofer, these theories span a space of implicit rationales. Decision aiding approaches may be positioned anywhere in this space. Merkhofer also links perspective conflicts to the theories. Going from Cost Benefit Theory (CBT) to Decision Theory the perspective changes from objective to subjective. The difference between Decision Theory and Social Choice is the perspective of the individual versus the perspective of the group. And moving from Social Choice to Cost Benefit, the debate is on equity versus economic efficiency. By moving away from a well-defined theory, an approach can be tailor made for the decision problem at hand, compensating weaknesses of one theory with strengths of the others. However, in this mixing of elements of conflicting theories, the approach may lose its internal consistency. This for example happens when the subjective concept of intolerable risk is mixed with the objective notions of costs and benefits, like the risk management frameworks do.

The decision aiding approach and its associated procedures is not the only perspective on risk decision making. Morgan (1993) for example approaches decision making from the rules that govern decisions. He recognizes three broad classes of rules: utility based, rights based and technology based. Utility based rules focus on maximizing net benefit and thus are highly aligned with the framework of Cost Benefit Theory. Right based rules are based on the idea that "there are certain things that one party cannot do to another without its consent, regardless of cost or benefits". Many environmental and safety laws are based on this principle, in the sense that they limit the risk to which humans may be exposed. This aligns with the theory of social choice, and even decision theory if applied by an elected official, but not with decision theory if applied by the party doing things to another party. The technology based rules are not aligned with any of the frameworks as they focus on what is technologically possible instead. Typical criteria are Best Available Technology (BAT) or As Low As Reasonably Achievable (ALARA).

In their report **COPING RATIONALLY WITH RISK** RIVM[78] (2003) expands these three categories with a fourth one: the precautionary principle. Especially with regard to environmental risks this principle is used increasingly, based on the Rio Declaration(United Nations, 1992). In that declaration, the principle is formulated in principle 15:

> *In order to protect the environment, the precautionary approach shall be widely applied by the States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing costeffective measures to prevent environmental degradation*

The principle is often understood as risk averse, adhering to evolutionary principles of "better safe than sorry". The use of this principle tends to correlate with decision making under large uncertainties.

3.2.4.2 Criticisms[79]

Neither of the presented theories or principles is without its criticism[80]. The major comment on Cost Benefit Theory is that it is very difficult to price risk. Only a few risks are actively traded (e.g. derivatives in the financial markets), but most risks (e.g. with regard to safety, the environment and reputation) are not[81], making their value notoriously difficult to establish. Indirect methods like hedonic pricing (deriving the value of risk like pollution from changes in observable prices for e.g. real estate) are not very precise as they measure many influences at the same time. Direct methods like the stated preference for the equivalent amount, either to prevent (willingness to pay, WTP) or to accept (willingness to accept, WTA) the risk, lend themselves for strategic behavior (in the form of absurdly high values) as no actual transaction takes place.

Furthermore, it may be doubted if every unit of exposure[82] has the same value given moral considerations. It makes a huge difference in the willingness to pay if it is about some statistical fatality risk (WTP may drop down to zero) or about a real person struggling for survival in a dangerous situation like someone stranded on the moon (WTP virtually unlimited). This not only applies to the monetary equivalent value, but even between human lives themselves. Compare crash-landing a plane in a sparsely populated area instead of in the middle of a city limiting the fatalities on the ground (pilots who do so are generally posthumously honored) with sacrificing a healthy individual to save several people waiting for the transplantation of an organ (doctors who do

---

[78] RijksInstituut voor Volksgezondheid en Milieuhygiene, the National Institute for Public Health and the Environment.

[79] As both risk management frameworks have a strong tendency towards cost benefit analysis, this will be the focal point of this section.

[80] Each method or approach suffers from difficulties in predicting the effects of the measures, modeling biases and so on, but those problems do not differentiate between methods.

[81] Insurance for those types of risks only covers the financial part.

[82] Exposure is used here (in line with chapter 11) to indicate the risk level.

so would be prosecuted for murder). Other arguments than just economic efficiency may become much more important like decision principles based on equality, rights and available technology, or the question whether risks were taken voluntarily or not. The formulation of the comparison may even be of significance, as Tversky and Kahneman (1981) have shown. Interventions are thus not necessarily comparable on a single scale of expected value based efficiencies. This is scientifically widely recognized, see for example Aven (2009) who argues that risk is more than expected values or Hanson (2004) that the price of a risk may depend on circumstances and that thus no universal value can be used.

Modeling all these aspects into some function representing this subjective value would be extremely complicated and would move power away from elected officials with the associated public accountability to some technical elite making the analysis[83]. Instead it can be more effective to use a subjective approach like Decision Theory or Social Choice. Many actual regulations are rather based on process design (embedded in the democratic process) than on cost benefit analysis. However, this does not result in undisputed results either. As RIVM (2003)shows, measures with regard to public health demonstrate a very large variation in efficiency. As a metric they use euro per QALY, a Quality Adjusted Life Year [84]. The best measures (often in prevention) are relatively cheap, sometimes resulting in net cost reductions per QALY. Environmental measures on the other hand are often very expensive, up to more than 1 million euro per QALY. This is not unique for the Netherlands. In a study on the cost effectiveness of life saving interventions Tengs et al. (1995) reviewed over 500 measures. To assure comparability between interventions, they categorized the interventions according to intervention type, sector of society, regulating agency and prevention stage. They found an even larger spread in the cost per life-year saved, both within and between categories. The interventions ranged from a net cost saving to a cost of more than 10^{10}\$ per saved life year, as displayed in Figure 9.

---

[83] The point the subjectivists make is that this modeling would be value laden and not objective.

[84] A QALY is one year of extra life with full functionality. If limitations occur, the extra year does not count as 1 QALY.

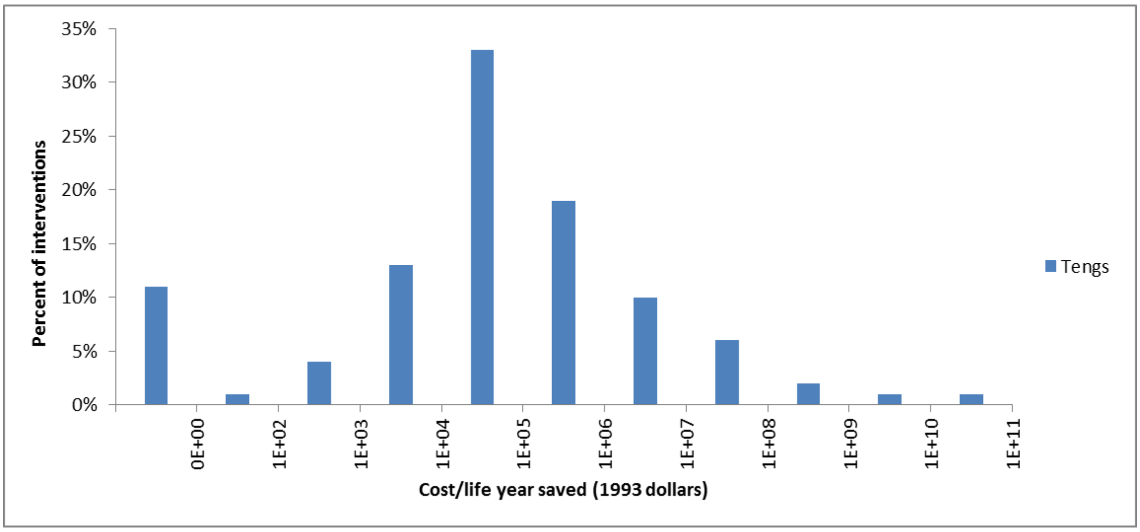


Figure 9: Distribution of cost per life saved for 587 interventions after Tengs et al. (1995).

Tengs et al. (1995) conclude that by shifting resources so that the marginal benefit per dollar spent is equal over all programs much more lives could be saved[85]:

> *This compilation of available cost-effectiveness data reveals that there is enormous variation in the cost of saving one year of life and these differences exist both within and between categories. Such a result is important because efficiency in promoting survival requires that the marginal benefit per dollar spent be the same across programs. Where there are investment inequalities, more lives could be saved by shifting resources. It is our hope that this information will expand the perspective of risk analysts while aiding future resource allocation decisions.*

RIVM (2003) recognizes this potential, but also adds the notion that this is only from a limited utilitarian definition of efficiency and other arguments as mentioned before may have been more important. However, it is hard to see how this would work for very comparable risks, other than that in applying those arguments any consideration with regard to efficient use of resources was lost[86].

This concern with regard to the rights based rules is strengthened in the review by RIVM (2003) of the Dutch limits on negligible and maximum tolerable risk to human life, being a mortality rate for an individual of 1 in 100 million and 1 in a million per year respectively[87]. For groups of risks the combined mortality probability may be a factor 10 higher. RIVM recognizes that there is no scientific ground for these limits other than that they are (almost undetectably) low compared to other risks individuals encounter. After all, the average mortality rate is in the order of magnitude one in a 100

---

[85] RIVM states in the report Coping Rationally with Risk that Tengs et al claim over 200.000 life years per year, but this claim cannot be found in the publication.

[86] Interestingly, the spread can also be used as an argument against cost benefit analysis, as apparently thorough deliberation (only documents explicit about the efficiency were used) results in a large spread whereas CBT requires a uniform value. However, it just may be that no reference was made to the peers, i.e. other interventions in the same category. It then just demonstrates that numbers in isolation have very little meaning.

[87] The number 1 in 100 million per year means that people exposed to this risk for their entire life have a probability of dying due to this risk of about 1 in a million (assuming a 100 year life), the tolerability level results in a 1 in 10.000 probability.

per year (given a life expectancy of roughly 100 years). Using such an arbitrary number to justify huge expenses may not be right, even if the limits have been established in a well-designed democratic process. Technically, the reverse might also occur, in the sense that effective measures are not taken because the required levels are already reached.

Similar concerns can be raised for technology based decision principles like BAT and ALARA. Available technologies to reduce specific risks may be more expensive than justifiable given options to mitigate other risks with similar impacts. But the lack of available technologies may also drive acceptance of a very high risk level, whereas it may just be wiser to cease the activity creating the risk. It has to be recognized though that both principles have a utilitarian notion. BAT is not independent of the ability to pay, thus putting a limit on the maximum cost. ALARA in general refers to costs being not disproportionate to achieved risk reduction. Other concerns can be raised about the moving target BAT and ALARA pose given technological advances. Either compliance has to be judged against the BAT/ALARA at the moment of design which threatens principles of equality, or installations have to be updated continuously which can be very costly.

With regard to criticisms between Social Choice and Decision Theory, an exchange of opposing viewpoints can be recalled between Hazelrigg (Hazelrigg, 2010) and Frey (Frey et al., 2010, Frey et al., 2009) on the value of the Pugh Controlled Convergence Method[88]. Hazelrigg regards the Pugh process as a form of social choice which is axiomatically flawed by Arrows impossibility theorem, whereas Frey argues that there is no voting involved and that the argument of a solution scoring best on all relevant criteria but not overall is a very theoretical stance difficult to envision in real life other than that some important criterion is then missed. In reviewing the debate, Reich(2010) showed there truly was a division into two camps with little exchange of ideas. He commented that the scientific debate “should expose the assumptions, the interpretation of results, and potential limitations of different positions.”

The concerns regarding the precautionary principle develop along a slightly different line of thought. Besides debates about what a serious or irreversible threat is and what is meant with full scientific certainty, the main point is that it is not a robust principle. How can it be certain that measures are cost effective if there are large uncertainties surrounding the risk? Furthermore, even though a risk averse strategy may be sound for an individual, the performance of a portfolio (i.e. society) is generally better if decisions are based on expected values, a risk neutral approach. Precaution may also create ethical dilemmas. Suppose a genetically modified (GM) crop was certain to relieve hunger in large parts of the world, would it be right to limit that just because it is GM with potential (but not proven) undesired side effects? In the end, every human activity impacts the environment and that ceasing all activity until it is clear no harm will be done would most likely result in an economic collapse.

Perhaps the most important lesson to be learnt from the criticisms on the approaches for risk decision making is that any approach can go wrong, especially if applied on a risk in isolation. Some explicit consideration of the right approach to be used may be very useful, as may be approaching risk management as a portfolio decision.

---

[88] A design method based on the improvement of a selected starting point

3.2.4.3 Selecting the right approach

Morgan (1993) recognizes that there is no correct choice of principle, as the right set of rules depends on the underlying values of individuals and society. Morgan also stresses that it is “critically important that the decision frameworks are carefully and explicitly chosen and that these choices are kept logically consistent”.

The view that there is no correct choice is not commonly shared. In order to shed some light on selecting the right approach for the right decision problem, Merkhofer (1987) proposes criteria for evaluating the approaches. These can be grouped into internal and external criteria. Internal criteria are Logical Soundness, Completeness and Accuracy, whereas the external criteria are Practicability and Acceptability. For any decision problem (following his taxonomy for characterizing risk decision problems) the available approaches (the three extremes plus three intermediates) can be scored on those criteria, to establish which is best. This may result in logically inconsistent approaches for different decision problems within the same decision maker. Unfortunately, only the structure for selection is given, not the answer.

The recommendation (RIVM, 2003) gives on the right approach for risk management, the risk ladder[89], is largely based on the paper **A NEW APPROACH TO RISK EVALUATION AND MANAGEMENT: RISK BASED, PRECAUTION BASED AND DISCOURSE BASED STRATEGIES** by Klinke and Renn (2002). In this paper, risks are classified according to nine criteria by means of a decision tree.

Table 10: Criteria for evaluating risks based after Klinke and Renn (2002).

Criteria	Description
Extent of damage	Adverse effects in natural units such as deaths, injuries, production losses
Probability of occurrence	Estimate for relative frequency of a discrete or continuous loss function
Incertitude	Overall indicator for different uncertainty components
Ubiquity	Defines the geographic dispersion of potential damages (intragenerational justice)
Persistency	Defines the temporal extension of potential damages (intergenerational justice)
Reversibility	Describes the possibility to restore the situation to the state before the damage occurred (possible restorations are e.g. reforestation and cleaning of water)
Delay effect	Characterizes a long time of latency between the initial event and the actual impact of the damage; the time of latency could be of physical chemical, or biological nature
Violation of equity	Describes the discrepancy between those who enjoy the benefits and those who bear the risks
Potential of mobilization	Understood as violation of individual, social, or cultural interests and values generating social conflicts and psychological reactions by individuals or groups who feel inflicted by the risk consequences; these could also result from perceived inequities in the distribution of risks and benefits

If risks do not exceed thresholds on these criteria, they are qualified as normal risks for which routine management is the right approach. The specification of these thresholds is regarded as one of the

---

[89] RIVM uses risk ladder whereas Klinke and Renn use risk escalator. Both metaphors have their pros and cons. A ladder represents the difficulties of climbing compared to the automated stairs of an escalator. However, the concept of escalation, growing in importance/magnitude/intensity, is captured very well by the term escalator. In this thesis preference is given to the notions of growing in importance and the term risk escalator will be used. This also helps in differentiating from risk ladders that are used for indicating risk severity.

main tasks of the regulating body. Risks that exceed any of the criteria are then assigned to one of the risk classes, based upon Greek mythology.

Table 11: Risk classes based on Greek mythology after Klinke and Renn (2002).

Risk Class	Characteristics	Typical examples
Sword of Damocles	Large damage potential, but very low probability, both relatively well known	Nuclear energy, natural hazards like periodic floods
Cyclops	The damage potential is high and well known, but the probability is largely uncertain	AIDS and infectious diseases, earthquakes
Pythia	Both damage potential and probability highly uncertain, though there is some understanding of the causal relations	Instability of West-Antarctic ice sheet, BSE
Pandora's Box	High ubiquity, persistency and irreversibility. Estimates for damage potential and probability are not only uncertain, the causal relations are not proven plausible	POPs, CFCs, Ecosystem changes
Cassandra	Risks for which both damage and probability are high and relatively well know, but with a delay between the initiating event and the occurrence of consequences, providing an opportunity for ignoring or denying the risk	Climate change, loss of biodiversity
Medusa	Risks for which damage and probability are low and well known (thus within the tolerability area), but that are nevertheless perceived as high risks	Electromagnetic fields

These classes can be plotted in a risk map, as shown below.

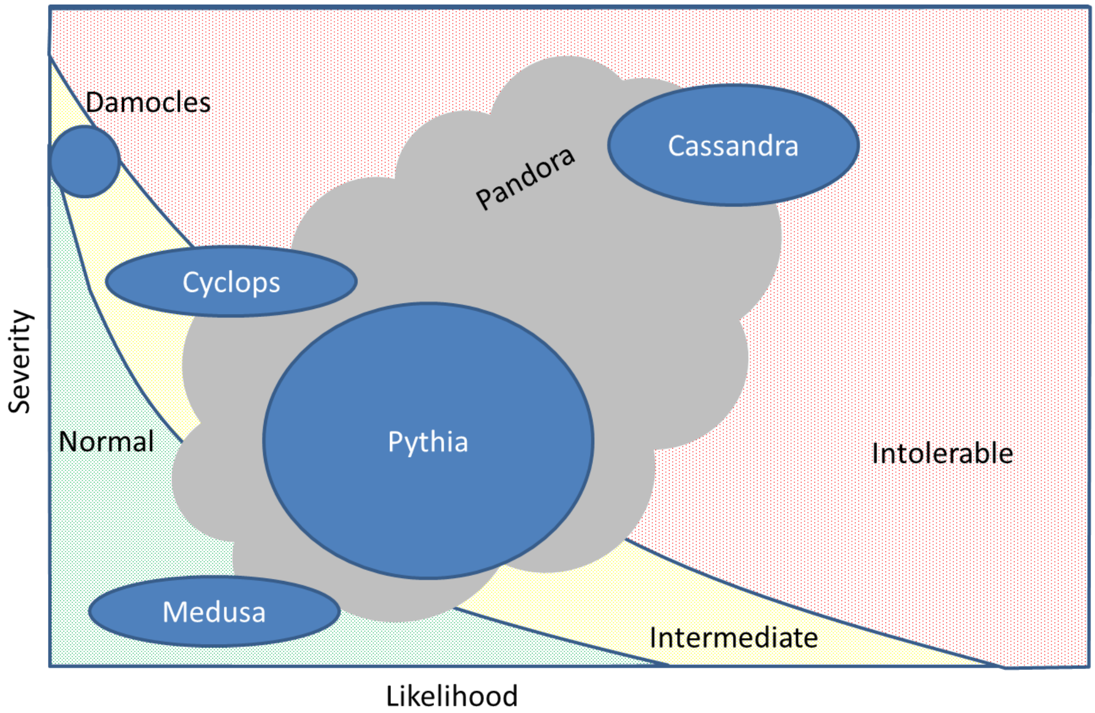


Figure 10: Plotting the risk classes in a risk map after Klinke and Renn (2002). Please note, in order to align with other risk diagrams in this thesis, the axes of severity and likelihood have been exchanged compared to the original.

The right decision making strategy for a risk depends on the key management challenge of the risk. Klinke and Renn distinguish three different challenges: complexity, uncertainty and ambiguity. The corresponding strategies for these challenges are risk based, precaution based and discourse based strategies. In the table below, the risks classes, management challenges and strategies are linked together, along with notions on the instruments, involved actors, type of discourse and type of conflict.

Table 12: Linking management challenges and decision approaches, free after Klinke and Renn (2002). The order of categories is the risk escalator. Effort in decision making increases from left to right, as more and more elements are added.

	Normal	**Complex**	**Uncertain**	**Ambiguous**
Management challenge	Balancing costs and benefits	The difficulty in identifying and quantifying causal relations	A range of problems from variation, measurement errors, ignorance and indeterminacy	The variability of interpretation of the same data
Decision Strategy	Routine operation	Risk based	Precaution based	Discourse based
Risk classes	Normal risks	Damocles and Cyclops	Pythia and Pandora	Cassandra and Medusa
Risk management Objective	Efficiency	Effective, efficient and adequate protection	1) Resilience 2) Fair distribution	Social acceptance
Instruments	Cost benefit analysis	Standards, risk-risk comparisons, cost-effectiveness, risk-benefit analysis	1) Containment in space and time, substitutes 2) Negotiated rule making, decision analysis	Advisory committees, citizen panels, value tree analysis
Actors	Agency staff	+ External experts	+ Stakeholders like industry and affected groups	+ Public representation
Type of discourse	Internal	Cognitive	Reflective	Participatory
Type of conflict	None	Cognitive	+ Evaluative	+ Normative
Required (additional) action	None	+ Scientific risk assessment	+ Risk balancing	+ Risk tradeoff and deliberation

Selecting the right approach is critical. Approaching everything as ambiguous would result in a waste of effort, but approaching every risk as normal could result in social mobilization, escalation and mistrust in regulating bodies. However, it should be reminded that

> *The desired result of risk management effort is not to reduce all risks to zero but to move them into the normal area, in which routine risk management and cost-benefit analysis becomes sufficient to ensure overall safety and integrity.*

The risk escalator thus not only works in adding elements when difficulties in risk decision making increase, but also could work the other way around. Ambiguous risks could (partly) be reduced to uncertain risks, then to complex risks and finally to normal risks, for which routine risk management could take over. A key question that is not answered yet is if this risk escalator could be applied “on the run”, starting with Cost Benefit Analysis (CBA) for every risk and then moving up the escalator if the outcomes are not accepted, which could allow more efficient decision making. However, it might be very difficult to notice CBA is not producing the right results from within the application of CBA as according to the axioms CBA is producing the right results. Furthermore, participants may not trust the decision process if there is already a CBA based conclusion produced by the organization, potentially requiring another party to do the decision making. This may lose the benefits gained in increasing the efficiency of decision making in the first place.

3.3 Reframing asset management for infrastructures

Asset management developed from terotechnology, which has clear roots in the production optimization of (commercial) plants. In the 1990's a move towards infrastructures was made, with the first standard on asset management (PAS 55-1: 2004) being limited to physical infrastructure assets. The successor of this standard (PAS 55-1:2008) relaxed the limitation to all physical assets, and the current standard (ISO 55000 series) includes all assets. It has to be recognized that the standards describe a management system, which focusses more on process than on content[90]. As processes can be presumed to be more or less independent of assets being managed, strong arguments for generalization of the concept of an asset can be found.

Yet, it also has to be recognized that not all assets are equal in terms of optimal asset strategies. Some assets may be designed for single use, in which case the development of a maintenance concept is not very useful. This is not necessarily limited to simple assets, even very complex machinery like a rocket launching a satellite into orbit is essentially a single use asset. On the other extreme are assets that are meant to last virtually forever, like the Yucca Mountain Nuclear Waste Repository with an required compliance time of 10000 years (National Research Council, 1995)**.** For these kind of assets the design of the system to keep the asset functioning is vital. Other assets may fall anywhere in between those extremes, from days (drills in mining), a few years (mobile phones, computers), ten years (vehicles, plant machinery, road surface), to several tens of years (cables, pipelines, buildings, routes) or even more than 100 years (dikes, bridges). If for a portfolio of assets certain strategies are very dominant, some influence on the asset management system may be expected.

3.3.1 Initial steps in recognizing the importance of asset characteristics

A starting point for this line of thought can be found in the work of Stacey (1990), who recognized that the risk involved in business decisions is affected by the length of the planning period, because the level of uncertainty in the prognosis increases for longer periods. Stacey distinguishes three environments in the future development. The first environment is the Stable Region or Closed Future, where business predictions are reasonably certain, with uncertainty measured in percentages. The second environment is referred to as the Probabilistic Environment or Contained Future. Predictions are not very certain, but some knowledge exists with regard to the amount of uncertainty, and order of magnitude estimates tend to be correct. The third environment is the Volatile Environment also referred to as the Open-ended Future. In this environment predictions are exposed to deep uncertainty and may be orders of magnitude wrong.

---

[90] In fact, all management systems complying with the new ISO template are highly similar, whether on assets, quality management, the environment etcetera.

Komonen builds on this idea of accelerated future uncertainty (Komonen et al., 2008, Komonen et al., 2010., Komonen et al., 2012). The business environment is modeled by recognizing the drivers behind the volatility, which are the dynamics of the market in which the asset is operated and the dynamics of the technology that is used, as shown in the framework below.

Market

Dynamic	**Specific Features e.g.** • Determine economic life • Short economic life-time • Life Cycle Profit (LCP)-approach required • Increase flexibility • New asset concepts	**Specific Features e.g.** • Determine economic and technical life time • Short economic life-time • Short pay-back time required • LCP-approach required • Manage dynamics • New asset concepts
Stable	**Specific Features e.g.** • Long economic life-time • Long pay-back time • Increase life time • Life Cycle Costing (LCC)-approach • Continuous improvements	**Specific Features e.g.** • Short technical life-time • Determine technical life time • LCC-approach • New asset concepts • Improve technical performance
	Stable	Dynamic

Technology

Figure 11: The influence of various business environments on asset strategies after Komonen et al. (2012).

In dynamic markets, the product lifecycle is very short. Typical examples are fashionable items (merchandise for events, clothing, gadgets) and products with rapid technological improvements (electronics like mobile phones and computers). The expected lifespan of these products may be much longer than their time on the market. The dynamics may originate in changing needs of the market, like merchandise for an event is not needed after the event. Another potential source of dynamics may be the scarcity of resources, like the need to produce petrol from coal or gas if oil is in shorts supply. However, as some may put it, the stone age did not end because of the lack of stones. Sometimes a substitute becomes available that fulfils the need behind the product in different way. The market for photographic films did not collapse because there was no longer a “need” to make pictures, but because the introduction of digital cameras provided a preferred alternative.

As a contrast, in stable markets, products have a very long lifecycle. These are typically commodities that are not consumed consciously, for example staple foods like bread, rice and potatoes, or energy products like petrol, electricity and gas. For these products it is almost inconceivable that the demand for them will ever run dry. Yet, the life span of the actual product may be very short, for example only single use for the mentioned foods and energy products. Dynamics of the market are thus not necessarily related to the lifespan of the products sold in that market.

The other dimension in the framework is dynamics of the technology for the assets with which the markets are served. Part of these dynamics is related to the technical asset life, as mentioned in the introduction of this section. But developments in the production technology may render assets obsolete well before their technical end of life. The typical example is the computer. Technological developments go that fast that only after a few years new equipment will outperform the old equipment by factors[91]. Developments resulting in efficiency improvements would have a similar impact, like the change from gas to electricity for lighting.

The optimal asset strategy depends on the dynamics of those business environments. For assets with a highly dynamic technology, it does not make sense to have strategies in place (like maintenance) that prolong the technical life. Instead, increasing the performance at the cost of wear and tear and thus technical life may be the right option. This especially holds in stable markets, where a new asset will be built alongside the old one because it simply performs better and the performance increase pays for replacing the asset. Optimally, the old asset would fail precisely at the moment of transfer to the new asset. However, if the market is also dynamic, optimally the end of asset life would coincide with end of product life. This may be very difficult to realize, especially because during the product life cycle the asset should be as reliable as possible in order not to miss any opportunities. It thus may make sense to make the asset more robust so that the asset will not fail during its useful life, but at the same time guarantee that the investment in the asset will be recaptured well before the end of useful life, by means of a life cycle profit approach. But having robust assets in a dynamic market opens perspectives on flexibility. Being able to adapt the asset so that a new product can be made is a way of extending the useful life. Part of the asset is then stable technology, but parts (like the mask for printing event logos on merchandise) will be dynamic, to be replaced at the end of product life. But in some cases the product of the asset may be aimed at a new market itself, like the focus on using gas for cooking and heating, after electricity took over the market for lighting. If both market and technology are stable, longer payback times become acceptable. The typical example is the transportation of gas and electricity by means of pipelines and wires. The product is basically the same since it was introduced more than 100 years ago, and some of the assets from the start were still in use until recently[92]. For assets operated in this combination of environments, it is useful to extend the technical life as much as possible, for example by means of maintenance strategies that are continuously improved.

3.3.2 Refining the framework

Even though the framework of Komonen et al. (2012) allows a differentiation of asset strategies dependent on dynamics in market and technology, it does not recognize characteristics of the assets themselves (outside of technical lifespan) which also can drive the way the assets are managed. For some assets maintenance may have little meaning. This is the case for example for cables, where the dominant internal failure mechanism is breakdown of the insulation, and nothing can be done to restore degradation. If a failure occurs, a certain length of cable will be replaced[93].

---

[91] Given Moore's Law, capacity doubles every 18 months, resulting in more than a factor of 10 in five years

[92] In 2007, the last part of the direct current system in New York which started in 1882 was converted to alternating current (Lee, 2007)

[93] Whether this is maintenance or replacement depends on the definition of the asset. If the cable is the asset, it is replacement, but if the link is regarded as the asset replacement of a few meters of cable is corrective maintenance of the link.

This point was addressed during a EURENSEAM[94] meeting in Sevilla in november 2009[95]. In that meeting, several important characteristics of assets were identified that would drive the way the assets would be managed. Added to this was the notion that the assets could differ in the dominant phase of their life cycle. For some assets, the whole behavior would be determined by their design, whereas for others the operation would be key. All these aspects and notions were grouped into a framework consisting of 4 quadrants, as shown below. The quadrant of management was not a result of the identification of important aspects, but added to be able to link the way the assets were managed to the asset characteristics.

Table 13: Integral asset management framework, after Van der Lei et al. (2010).

Asset characteristics	**Environment**
Function (commercial or public) Behavior (discrete(batch), continuous or passive) Technical life span Location (fixed/moving, point or distributed) Technology (mechanical, civil, electrical, software)	Dynamics of the market Dynamics of technology
Dominant phase of Life cycle	**Management**
The Life cycle phase (concept, design, manufacturing, assembly, commissioning, operation and maintenance, and disposal) in which (most) of the asset value is determined	People Key performance indicators Budgets Processes Systems

The framework was applied in a series of interviews with infrastructure managers in the Netherlands. All managed a portfolio of assets differing over all aspects. Based upon these interviews several notions on asset management could be made:

1. Different limitations were encountered for asset life span:
   a. Technical: Assets breaks down but function is still required
   b. Market development: There is no longer for the function the asset provides
   c. Technological: A better alternative is available for providing the function of the asset
   d. “Unlimited”: Repairs are always possible and cheaper than replacement, though limitations b and c could still be encountered
2. Most of the assets faced a technical end of life (limitation a) with a life span typically measured in 10s of years. Technological end of life (limitation c) was mentioned in all interviews but only as applying to a very small fraction of the asset base, like the (digital) controls.
3. For passive, non-operated assets like cables, pipelines and civil constructions the asset value was determined in the design of the asset. The maintenance need (if any) was addressed separately. For active, operated assets like water purification plants, the focus was much more on the operational phase, as the operation can be improved within a design with small modifications.
4. The majority of the assets were not actively operated.

---

[94] European Research Network on Strategic Engineering Asset Management

[95] For a full description of the discussion and the results see Van der Lei et al. (2010)

5. Most assets served a public function and could impact much more values than only those of direct importance for the owner.

Combining these points results in the notion that infrastructure asset management may be something very different from other forms of asset management. Due to their passive characteristics and the environment in which they are operated the asset managers are focused on design. For the (few) assets that need maintenance this is considered separately[96]. Operational improvements are for most infrastructure assets unthinkable. This is in strong contrast with asset management for active assets, for which operational improvement combined with maintenance is the key focus area.

3.3.3 Asset management for energy distribution infrastructures

The idea of infrastructures being a separate class of assets to manage was furthered in **A SYSTEMS VIEW ON INFRASTRUCTURE ASSET MANAGEMENT** (Herder and Wijnia, 2012). Several specific characteristics of infrastructures were listed that set them apart in asset management. Furthermore, the notion was added that managing existing infrastructures is mainly about managing risk of those infrastructures. This notion of risk management being the key element in infrastructure asset management was developed into a risk based reference model for infrastructure asset management (Wijnia et al., 2014a, Wijnia et al., 2014b, Wijnia and de Croon, 2015).

Reflecting back to the infrastructures for the distribution of electricity and gas this may even be more so. As demonstrated in the research context, in the early phases of development the focus was on expansion. Part of this considered connecting new customers, but another part was about increasing the utilization of the assets, by promoting products to be used in the off peak hours. With market penetration reaching a virtual 100% and a very high utilization rate of assets the potential in this area has dried up. There is still growth, but that is because of a growth of customers and a growth of consumption per customer. Both are more or less autonomous with the numbers in the order of magnitude of 1%. As a consequence the attention of the asset manager turned towards managing the existing asset base. The decisions typically concern the maintenance concept and interval for the few assets that require so, the timing of investment for upgrades (to accommodate growth) and replacements (to deal with asset ageing), and the choice between technical alternatives. These can all be regarded as optimizations. Given the high reliability of existing grids, neither of these decisions would impact sales significantly. Maintenance and replacement decisions would impact the direct failure probability of assets, and upgrade decisions would impact the probability of overload induced failure. All these failures could result in outages with potential (economic) damage for the customers[97] and a potential for dangerous situations[98]. This makes the asset management decisions in the energy distribution infrastructures cost and risk only considerations.

---

[96] However that differs per type of asset: roads for example do require significant maintenance even whilst being passive. In case maintenance is organized in a separate department with a separate budget asset management may be regarded as the professionalization of maintenance (Wijnia and Herder, 2009), even though the maintenance costs in the total cost of ownership still may only represent a small fraction.

[97] In general the costs of interruption for the customer per unit of unserved energy is much higher than the price paid per unit of energy.

Given that the assets perform a public function and are operated in the public domain, this may pose special requirements on managing those risks. When the criteria as specified by Klinke and Renn are applied, none of the thresholds (even though they are not yet defined) seems to be violated.

Table 14: Confronting the criteria specified by Klinke and Renn (2002) with risks in the energy distribution infrastructure.

Criteria	**Description**	**Applicability to energy distribution infrastructure risks**
Extent of damage	Adverse effects in natural units such as deaths, injuries, production losses	Limited, often only economic damage or nuisance because of outages. Fatal incidents are possible but often limited to single fatalities.
Probability of occurrence	Estimate for relative frequency of a discrete or continuous loss function	Large incidents are (very) rare both with regard to reliability and safety
Incertitude	Overall indicator for different uncertainty components	Low, risks are well understood. Small incidents occur very often
Ubiquity	Defines the geographic dispersion of potential damages (intragenerational justice)	Uniform, any user is at comparable risk
Persistency	Defines the temporal extension of potential damages (intergenerational justice)	None, effects do not persist beyond the repair
Reversibility	Describes the possibility to restore the situation to the state before the damage occurred (possible restorations are e.g. reforestation and cleaning of water)	High, in general damages can be restored
Delay effect	Characterizes a long time of latency between the initial event and the actual impact of the damage; the time of latency could be of physical chemical, or biological nature	None, failures have an immediate effect
Violation of equity	Describes the discrepancy between those who enjoy the benefits and those who bear the risks	Little, most users are at comparable risk and benefit from the infrastructure
Potential of mobilization	Understood as violation of individual, social, or cultural interests and values generating social conflicts and psychological reactions by individuals or groups who feel inflicted by the risk consequences; these could also result from perceived inequities in the distribution of risks and benefits	Little, the whole society benefits directly from using the infrastructure assets that may pose a risk. Vulnerable users can protect themselves.

This means that most risks (with perhaps a few exceptions) in the energy distribution infrastructure would classify as normal risks. The right approach for decision making on ins for those risks should be cost benefit analysis with risk reduction as the benefit. Given adequate pricing of risk, the mentioned decisions the asset manager faces can be developed into formal optimizations, like the general form introduced in the section on the scientific origin (Figure 4).

---

[98] Safety in normal operation is covered by many safety regulations, often even extending to single failures or single errors. Appliances are insulated to prevent touching live wires, but even if a life wire is touched in most homes the ground fault interrupter switches off power before injuries occur.

4 Research design for risk based system optimization

4.1 The current state of knowledge

As defined by the ISO standard, asset management is the coordinated activity to realize value from assets. This normally involves balancing costs, risks, opportunities and performance benefits. Many case studies in asset management literature prove significant improvements in any (or even all) of those aspects. For infrastructure assets, and especially those used in energy distribution, this balancing act can be simplified, because there are very little opportunities to increase the (gross) value delivery (given full market coverage) or increase the performance benefits given performance is virtually at 100%. Net value delivery then only can be increased by improving the cost - risk balance.

Several frameworks[99] exist for balancing cost and risk. Cost Benefit Theory is one of the frameworks, but Decision Theory and Social Choice theory are others. The right framework depends on the characteristics of the risk for which the decision has to be made. According to the risk escalator, the more uncertain, complex and ambiguous a risk is, the more stakeholders should participate in decision making so that their (subjective) preferences can be captured. As demonstrated, the risks in the energy distribution infrastructure score very low on the criteria that determine the position on the risk escalator. Even though the thresholds for those criteria are not yet defined, the low scores are a very strong indicator that most risks in the research context can be categorized as normal risk. For these normal risks, cost benefit analysis should be the right approach[100]. Furthermore, it is also noted that the purpose of the more participatory approaches is to bring the risk back into the category of normal risk so that decisions can be made using Cost Benefit Analysis (CBA). Along with notions on the critical importance of consistency within decision making this indicates that cost benefit analysis should be the base method for decision making in energy distribution infrastructure asset management. However, there needs to be awareness that in special cases CBA may not produce acceptable results and that it is not certain that “escalating” CBA to more participatory methods will be accepted in those cases.

As demonstrated in the previous chapter, many decisions in the research context can be regarded as optimization problems. There is ample documentation that such optimizations are at the core of asset management[101]. However, in those optimizations risk is limited to economic risk with the non-financials excluded from the optimization. They are part of a separate consideration. According to Cost Benefit Theory these non-financials can be priced and included in the optimization, but there is no generally agreed pricing scheme for those aspects. This problem was already mentioned in the

---

[99] Used here to indicate a consistent theory of decision principles and associated methods

[100] Given the mentioned lack of improved performance benefits and opportunities in this context, benefits should be understood as equivalent with risk reduction

[101] See for example **Asset management – an autonomy** (IAM, 2014) with numerous examples of economic optimization

introduction to this thesis. Assuming an adequate pricing scheme can be developed[102], integration of these non-financials into optimization is technically not a problem, they then just become one of the costs to determine the optimal decision.

Unfortunately, that assumption is not self-evident. As Tengs et al. (1995) showed (see chapter 3.2.4.2), actual decisions on life saving interventions demonstrate a very large spread in value per unit of risk. With such uncertainty in valuation, optimizations can go anywhere. The suggestion Tengs et al. make to increase the overall efficiency is to align the interventions to a more uniform marginal benefit per dollar spent per program. Decision making on an intervention then is not limited to the cost and benefits of the intervention itself, but also considers other comparable interventions. This is more like a portfolio approach with the selection of the right set of interventions given the budget.

Such a portfolio approach would be a first step in the transition from optimization of individual decisions to whole system optimization, as it would result in the best achievable performance given the budgets and the interventions to select from. Whether that would be the best achievable result as such would depend on the question if the right problems were addressed for intervention development in the first place. The best set of solutions for a set of irrelevant problems is not likely to be optimal in a broader sense. Reasonable assurance that the most relevant problems are addressed can be achieved by implementing a risk identification and evaluation process, as advocated by both ISO 31000 and COSO. This can be regarded as “procedural” optimization, with optimal solutions for the most relevant issues[103]. “Formal” optimization of the system (like the total equivalent costs as a function of the planned budget) would only be possible if the total risk would be known, including all potential measures[104].

At this point it is important to recognize that Cost Benefit Theory does not provide a framework for identifying the most relevant problems/risks, simply because it is a differential method. It compares two or more options, to determine which one has the highest net present value[105]. In this comparison the absolute value does not matter. The reviewed standards of good practice (ISO 31000 and COSO) use evaluation criteria based on the combination of probability and impact, for example by means of a risk matrix[106]. In the evaluation the concept of unacceptable risk is used, for which interventions are required[107]. The evaluation criteria can be aligned with Cost Benefit Theory, by

---

[102] The other requirement is that the non-financial impacts can be modeled with reasonable accuracy, but this applies to the financial impacts as well.

[103] Optimizing the most relevant problems in a continuous cycle can be regarded as continual improvement.

[104] These conditions are (and cannot) be fulfilled completely in real life. This was used as an argument to omit the concept of optimization from ISO55k (personal communication 2 members of ISO committee) whereas it was part of PAS 55. However, on individual decisions optimization has meaning, as demonstrated earlier.

[105] If CBA is used to judge a single proposed action, it may appear if an absolute judgement is given, but in fact a comparison is made with (continuing) the current situation as the base alternative.

[106] Also used in technical analysis like FMECA and RCM

[107] As a reminder, unacceptable does not necessarily have the same meaning in both standards. It can be really intolerable with a requirement to reduce it to the tolerable zone no matter what, or just not blindly acceptable.

ranking risks on their total equivalent value, but that is neither common practice nor well documented in literature. This means alignment of risk evaluation with decision making on interventions is a separate challenge.

The evaluation criteria for risk (expected value based or not) only provide a ranking of identified risk, but not the identification of the risks themselves. Getting a complete overview of all risks is not helped by the debate on the concept of risk. As argued in chapter 3, there is no agreement on what risk precisely is and many different definitions are used. One of the fundamental debates is on whether risk is purely about undesired outcomes (e.g. COSO), or that it should include desired outcomes as well (ISO). Another debate is whether risk is the event that produce undesired outcomes, or that risk is a measure for the expected misery (risk = probability*impact). Some even regard risk as the variation in the total amount of misery. The third fundamental debate is about subjective versus objective risk. This is not so much about what risk is, but more if the risk can be assessed objectively, as in modeling risk many subjective choices (like clustering, used metric, hidden assumptions) are made that in the end result in judgment on the (un)acceptability of the risk. Yet, it seems many of those debates consider non-normal risks. Within the normal domain, a pragmatic solution is to regard risks as events with the potential of producing undesired outcomes. The risk severity is then represented by the expected amount of misery, which can be (given pricing of all impacts in the normal domain) the monetary equivalent expected value summed over of all impacts. Unfortunately, even for this limited subset of the concept of risk no proven and well documented methods for getting an overview exist. There even is no agreement on for which values impacts have to be considered. It seems a procedural approach (as defined by the standards of good practice) is the best there is.

4.2 Knowledge gaps for risk based optimization[108]

Summarizing these considerations, the risks in an energy distribution infrastructure can mostly be classified as normal risks for which cost benefit analysis is the right method for judging interventions. In many decisions this takes the form of optimization. The additive character of net present values allows summing of the value of the outcomes of those decisions. Therefore at least theoretically “formal” optimization of the “normal” part of the system should be possible, for example in total equivalent value versus planned costs. The scientific knowledge on such system wide optimizations in asset management is very limited, both with regard to “formal” and “procedural” optimization. At the moment, the debate on the development of scientific knowledge is even mostly lacking. This applies both to the prerequisites for optimization as to the achievable results as well. Below, the most important knowledge gaps are listed, sorted from more prerequisite like to more result like.

- **Risk management value system**: This determines what is good and what is bad. For asset risk management, this value system should consist at least of 3 elements: the values to be considered, risk evaluation criteria and risk monetization factors. The values to be considered provide guidance for risk identification. If a value (e.g. esthetics) is not considered as important, it does not make sense to identify risks that may impact that value. In a way, these values are a prerequisite for risk. If there is no value, it cannot be harmed and there is

---

[108] In the context of this thesis, risk based optimization (abbreviated as RBO) is founded on the principles of Cost Benefit Analysis. For reasons of readability the reference to CBA is often emitted.

no risk. The risk evaluation criteria help in prioritizing attention. There may be much more risks identified than can be managed, and it is important to select the most relevant risks. This possibility to prioritize is a prerequisite for procedural optimization, finding the best solution for the most relevant problems. The monetization factors allow pricing of risk and risk reduction, and thus are a prerequisite for optimizations. If these factors exist as a fact of the world then theoretically only one global optimum would exist. But given uncertainties and subjective choices in modeling the risk value, this cannot be assumed. **The** optimum then becomes **an** optimum[109] for the used assumptions, valuations and constraints. At the moment (as mentioned in the introduction to this thesis) no generally agreed upon value system is available for the energy distribution infrastructure, nor is there agreement on how to capture and represent that system. This means that even the existence of an optimum in such a confined part the world can be debated and thus the concept of optimization itself.

- **The risk position**: This is the total expected amount of misery in the system represented by a list of risks. Understanding the risk position is a requirement for whole system optimization. Without such an understanding, relevant risks could have been missed in intervention design, and the actual total amount of misery may be very different from what the optimization predicted. No practical method currently exists for capturing the total volume of risk in a system, not even for normal risks. A procedural approach for identifying and evaluating risks on their relevance is the best there is. This can be enough for procedural optimization (consistently making the best choice within constraints) but it does not allow any debate on whether the constraints are right, in the sense that they maintain an acceptable risk position.
- **Acceptance**: The extent to which the stakeholders can agree with the outcomes of risk based optimizations. Acceptance is required for optimizations to deliver value, as without acceptance the outcomes of optimizations will not be implemented. This holds for all levels of optimizations. The knowledge gap on acceptance considers both the risks that are normal according to the criteria but for which stakeholders may not accept CBA as the right method, as the acceptance of using CBA as a first step in the risk escalator to be followed by more stakeholder involvement for non-normal risks
- **Potential net value of optimization**: The relative or absolute improvement in total equivalent value that risk based optimization can deliver to an organization. This may be very dependent on the context in which optimization is implemented. There is little knowledge on the conditions under which optimizations do or do not deliver value.

4.3 Research goal and research questions

Adhering to the principles of CBT, Risk Based Optimization (RBO) should only be practiced if it has a positive net value. This means the results should be positive, the results should be acceptable and

---

[109] Optimization is finding the input variable that gives the best value on an goal function, as demonstrated in the example on the optimal maintenance interval. Formal (objective) optima only exist if the goal function is not disputed, like mathematical optimizations (e.g. shortest path) or physical optimizations (optimal power transfer) but if subjective choices or uncertainties are involved this objectivity cannot be maintained. The optimum is then only valid for the used frame of reality. In general many different frames are possible, all with (potentially) a different optimum, hence the use of an optimum.

the cost of optimization should be below the value of the results. For decision making on implementing RBO this should also be known beforehand with reasonable certainty, as no asset manager should be willing to commit large amounts of resources for a highly uncertain result. The research goal of this thesis is to contribute to the body of knowledge on those aspects, formulated as follows:

> *To understand the achievable results, costs and limitations of and conditions for risk based optimization in asset management in the energy distribution infrastructure*.

In terms of process, the goal is to initiate or inspire the debate within the asset management community on what is possible with risk based optimization.

The pivotal point for achieving this goal is the recognition that the knowledge gaps are not independent in real world implementations of risk based optimization, where feasibility is of concern. Filling in the prerequisites in a very simple way (e.g. only considering cost and financial risks) enhances the practical feasibility, but it will be virtually certain to limit acceptance if other values than finance are impacted. Furthermore, such a basic implementation would give a very distorted view on the added value of RBO as many relevant aspects are neglected. On the other hand, a very sophisticated monetization and prioritization scheme, combined with an extensive model of the risks may just be too complex for any organization to handle. RBO then would only be executed on a very limited scale, limiting the potential value again. It even may be applied incorrectly, potentially destroying value. Somewhere in between these extremes a representation of value system and risk position may exist that allows delivery of maximal value. But that means that the level of sophistication of the prerequisites becomes an optimization variable for the value RBO itself can deliver.

At the current state of knowledge it is not possible to formalize such an optimization. It would require knowledge on the net value different levels of sophistication of the prerequisites could deliver, including a deep understanding of the factors driving the value delivery. At this moment there is not even full knowledge on the value any level of sophistication can deliver via formal, whole system, risk based optimization, neither theoretically nor practically. It is not even certain that such a formal optimization is practically feasible[110]. The key research question for this thesis therefore is about the feasibility of such a formal whole system optimization.

> To what extent is formal risk-based-whole-system optimization feasible in managing assets of an energy distribution infrastructure?

Feasibility can be used to indicate both the theoretical possibility as the practical achievability. The theoretical possibility is about the existence of an optimum. This requires that for all assets and risks that are managed, cost benefit analysis is acceptable as decision method (i.e. only normal risks), that for all decisions the same monetization factors apply (i.e. full consistency of the value system) and that the risk position of the system can be known (i.e. expected misery equals actual misery). If these

---

[110] The practical feasibility of procedural optimization of the whole system is not doubted, given the cycles of continual improvement required in the standard for asset management.

conditions do not fully apply, an optimum would only exists partially. The theoretical extent of feasibility then is the fraction of the portfolio for which these conditions apply.

To be practically feasible this optimum must be approachable with reasonable certainty given limited resources. The term reasonable certainty is used here to indicate that the errors made due to imperfections in the optimization method are acceptable in relation to the errors due to the fundamental uncertainties surrounding the optimization. Approachable means that the optimization must be scalable in terms of commitment of scarce resources. No decision maker adhering to the principles of asset management should be willing to provide the means to do a whole system optimization if there was no practical evidence in the form of smaller scale pilots or experiences in very similar organizations that such an optimization would deliver value with reasonable certainty. This is only possible if robust methods exist, for which increased effort results in a predictable increase of the value of optimization.

Given the considerations on the characteristics of risks for infrastructure assets, the base assumption is that an optimum exists for a very large fraction of the portfolio. The research is therefore approached as if an optimum exists for the full portfolio, reducing feasibility to the practical possibility of approaching that (assumed) optimum by means of a robust method.

This results in the following four sub questions:

1. What is an adequate representation of the value system that facilitates both CBA as the selection of the most important risks?
2. What is an adequate representation of the risk position?
3. What is the effectiveness of applying Risk Based Optimization by means of these adequate representations?
4. How robust is this effectiveness of Risk Based Optimization?

Representation is used here in the double notion of content and process, i.e. the image and the method for making that image. Effectiveness is used to indicate a combination of added value and acceptance. Added value without acceptance is not effective, as is acceptance without added value. Robustness is used here to indicate both the insensitivity to errors made in the optimization (either due to (numerical) uncertainty, misunderstanding resulting from complexity or even applying the wrong decision approach for the decision at hand) as to the ability to improve the optimization after disclosure of these errors.

4.4 Research planning

A straightforward way of planning the research would be to answer the research questions (and fill the underlying knowledge gaps) in the order in which they were presented[111]. First defining the value system, then establish the risk position, optimize the interventions for those risks within the relevant constraints and test for the acceptance and value that optimization would deliver. However, that approach would be at significant risk of not reaching conclusions. Neither the value system nor the risk position are facts of the world, they are imposed (subjective) structures upon reality. As a consequence, there is no natural end point in representing them in the sense of a universal truth.

---

[111] That order is (not coincidentally) aligned with the process steps in risk management

Risk monetization factors, risk prioritization criteria and the risk position can always be represented differently (both in terms of resolution as in structure[112]), but which representation is adequate can only be answered by considering the quality of the decision outcomes it delivers. The stop criterion thus would be situated outside the research question at hand. In line with the conclusion of the National Research Council in **UNDERSTANDING RISK** (1996) that risk characterization is a decision driven activity, it may be necessary to reverse the order of steps. By understanding the decision that can or needs to be made in the optimization, the risk(s) and uncertainties that drive the outcome of that choice can be identified, and based on those the relevant impacted values can be derived.

Furthermore, the straightforward approach would not necessarily provide any insight in practical feasibility, consistency, acceptance and value. Suppose that the problem of no natural end point would be resolved by limiting the effort per step, preferably in a number of cycles of improvement. The first round then necessarily be highly abstract with assumed relations between inputs and outputs in the form of a system response[113]. But because of being abstract, it would not provide direct answers to be validated, accepted and implemented, and thus could not deliver value. That means that it could not be used as a predictor for acceptance and value of the next round. Only when the system would be modeled with enough resolution to result in practicable answers, acceptance and value could be assessed, and used as a predictor. But once practicable answers would be given in a whole system approach, there would not necessarily be a need for a next round as the whole system would have been covered already. The approach thus would presumably violate the requirement of being able to predict the value of the optimization before committing resources.

In response to these two potential flaws of the straightforward approach a more incremental and exploratory approach was adopted. This effectively reversed the risk management process, by starting at the decision and ending with the value system. Furthermore, it would work bottom up instead of top down by developing from risk based optimization for single assets via procedural optimization of the portfolio to formal system optimization. Staging the experiments in this order assured robustness of the approach: if an experiment did not provide a valuable outlook for the next experiment, the research project could be redesigned. To ensure practical feasibility these “experiments” were conducted in the asset management department of Enexis. The planned timeline of the research is shown in Figure 12. This will be followed by a description of the planned experiments in more detail.

---

[112] Resolution is used here to indicate the number of values and risks to be considered. Risks for example can be modeled for asset groups, assets, asset function or parts. Structure is used to indicate the clustering of values and risks like grouping assets to type, age and location.

[113] For example in system dynamics modeling, of which **The Limits to Growth** (Meadows et al., 1972) is an (in)famous example

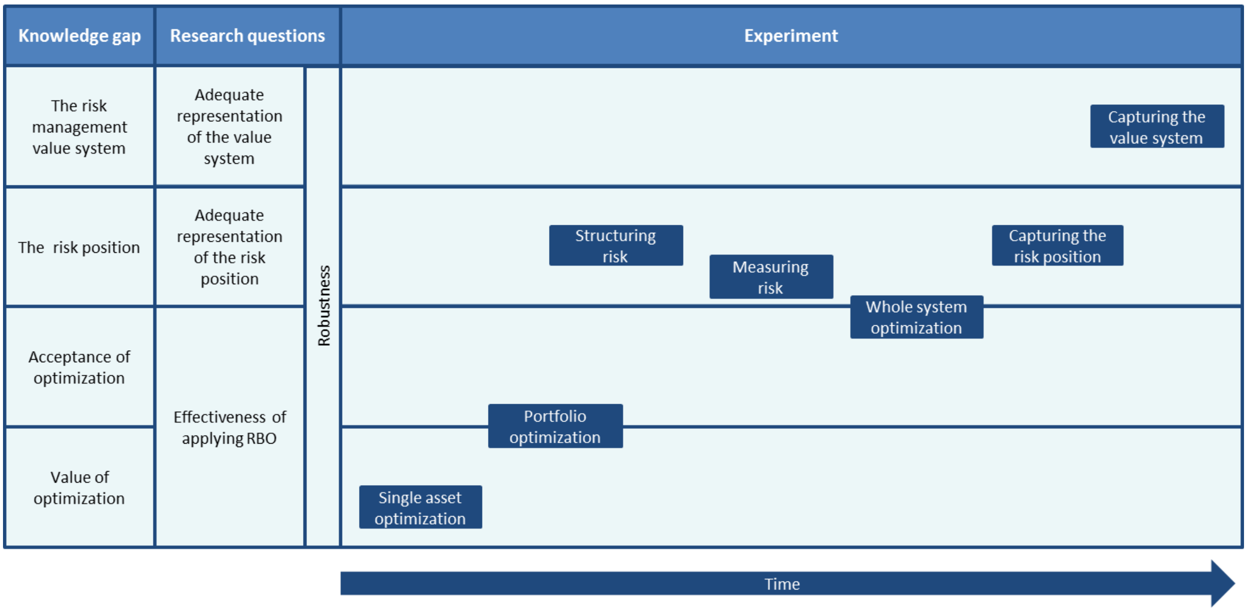


Figure 12: Timeline of the research with the experiments plotted against the knowledge gaps and research questions.

The first experiment in the research plan considers the potential value of risk based optimization for single asset decisions. Only if this value is large enough there is a justification for a wider rollout of RBO, and in a way for continuation of the research project. The method used in this experiment is contrasting several more risk based evaluations of options (including real options) with classical technical evaluations. The decision problem in this experiment is the optimal timing of a circuit upgrade in the electricity distribution grid. The value system in this experiment is limited to two aspects, finance and reliability. A literature based monetization factor is used to compare these aspects. The value system holds no risk evaluation criteria. This experiment is documented in chapter 5.

The second experiment expands the single asset decision to a risk based portfolio decision, selecting the best set of interventions for the annual plan. This is a procedural optimization. The key topic of interest of this experiment is acceptance. Due to the requirement of practical feasibility annual plans are prepared in several teams, often with predetermined budgets. The experiment first performs a ranking of interventions within the teams, based on the risk reduction per unit of cost (the yield) on 4 values: economics, reliability, safety and compliance. This provides a clue towards acceptance within limited settings. The next step of the experiments is the integration of the results of the teams into a single ranking, potentially shifting budget between the teams. That would be a true test with regard to acceptance. Additionally, the experiment allows testing of the relation between whole system cost and performance by varying the budget constraint in the portfolio decision. This provides insight in the potential value of such procedural optimizations. This experiment is documented in chapter 6.

To expand the procedural optimization to a formal optimization knowledge on the risk position is required. This quest for an adequate representation is split up in several experiments. The first experiment considers the method for representing the risk position, by comparing several methods in their expected success for structuring the identified risks. This is documented in chapter 7.

The next experiment then provides a test for this method. Key research goal is to see if the proposed method helps in representing the total risk position by means of a limited set of risks. The risk position used for this test was the total incident risk of the gas grid. As the incidents in the gas grid mainly consider safety and economic costs, this experiment also provides an option for determining the exchange rate between those values. This experiment is documented in chapter 8.

The fifth experiment is about formal optimization of the whole system within constraints, as an alignment of optimizations of individual assets. This brings together the considerations on optimizing of the first experiment, the portfolio approach of the second experiment and the representation of the risk position, resulting from experiments 3 and 4. The value system in this experiment is a combination of experiments 1 and 4. The research goal of this experiment is mainly feasibility of such a system optimization and acceptance of the outcomes, based on relatively simple representations of the value system and the risk position. The experiment in that sense also allows calibration of the used monetization factors by means of their impact on system performance. Additionally, by expressing the total value of the asset base with the metrics in the value system, this experiment allows optimization of the constraints, at least in the form of strategies to adjust the constraints. The decision problem in this experiment is the optimization of the replacement strategy for the whole asset base. This experiment is documented in chapter 9.

The planned experiments so far only model the risk position from a single perspective on risk. In experiment 4 this is the incident perspective, in experiment 5 the asset ageing perspective. Theoretically, many more perspectives could be relevant for understanding the whole risk position. This is addressed in the sixth experiment, on the feasibility of adequately representing the risk position with an unstructured collection of risks with multiple perspectives. If the same method as identified in experiment 3 is useful in doing so, it can be regarded as a method for representing the risk position. The results of this experiment are documented in chapter 10.

The final experiment is on aligning the monetization factors used in optimization of interventions with criteria for selecting the risks for which interventions should be developed. This experiment combines insights from experiment 5 on whole system optimization with the unstructured overview of risks in experiment 6. The resulting rules for design and use of such a value system are documented in chapter 11.

The experiments are documented in part 2. In part 3, the key findings for the experiments will be repeated. This will be followed by confronting these findings with the literature on the knowledge gaps. After this, the research questions will be answered in the conclusion section of this thesis.

PART 2: EXPERIMENTS

The chapters in this part describe the experiments that were conducted to answer the research questions. Most of the chapters have been published before. For each chapter, the original publication is indicated. Some minor changes were made to align the publications with the rest of this thesis. These consider the title, naming convention, styles and layout. The references per publication are moved to the reference list of this thesis, for an adequate overview. In some cases the published paper does not entirely cover the research question as relevant for this thesis. In those cases the publications have been amended with an epilogue with other (published) results of experiments.

5 Single asset optimization

This chapter was published before as **OPTIONS FOR REAL OPTIONS: DEALING WITH UNCERTAINTY IN INVESTMENT DECISIONS FOR ELECTRICITY NETWORKS (WIJNIA AND HERDER, 2005)**[114]**.**

5.1 Introduction

According to the European regulations the electricity markets have to be liberalized. In this system competition is supposed to drive down costs and improve service levels. This competition presumably will not work for the networks, because the huge investments needed to build a competing network can never be recovered by gains in operating efficiency. Networks are `natural monopolies`. Because real competition is not feasible, regulators simulate competition by posing income cuts on the network companies. These cuts can be substantial. In the United Kingdom the average income cut for the network companies in the year April 2000/March 2001 was 21,8%. The total average income cut over the period 1995-2001 was more than 50% (Electricity Association, 2000).The level of the income cut is determined by the efficiency improvement of the best player. In this situation it pays to be the most efficient network operator. Reacting to this incentive the network companies started to rationalize their expenditure patterns for the network. This is called "Asset management" by most companies. For asset management different definitions exist, although the differences in work practices are almost negligible. All asset managers try to postpone investment, reduce operating cost and improve the performance of the network.

To realize those seemingly conflicting targets at the same time two approaches can be followed regarding the investment opportunities[115]:

1. *Optimize the individual opportunity*. In this approach investment proposals are scanned for unnecessary or contra-productive activities. For example, Preventive Maintenance is based on the bathtub curve, according to which components have a high failure rate right after their start up (infant mortality) or after a certain period (wear out). However, while 72% of systems suffer from infant mortality, only 11% shows wear out problems(ALM-43-7494-C). In 68% of the cases maintenance would do more harm than good by introducing a new infant mortality phase.
2. *Optimize the portfolio of investment opportunities*. In this approach the effort is not spent on improving the individual proposals, but on searching for better opportunities in the network. For example, if the regulator sets a target for the Customer Average Interruption Duration Index (CAIDI) (Short, 2002), this approach would identify all opportunities for improvement and start with the ones with the highest yield (CAIDI improvement per Euro spent).

Of these alternatives the first is relatively straightforward, as it does not require a new approach to valuing the investment opportunities. The second one does. It requires a framework with which one can value the different aspects involved in infrastructure decision making, and there are many. All asset managers have somehow introduced a multi criteria analysis in their decision making. This MCA can be a cost benefit approach, a cost-risk approach, or any other variety of MCA available.

---

[114] Wijnia, Y. C. & Herder, P. M. 2005. Options for real options: Dealing with uncertainty in investment decisions for electricity networks. *International conference on Systems, Man and Cybernetics.* Hawaii.

[115] In the context of this paper maintenance is also an investment opportunity.

Sometimes the MCA is performed at the individual investment, sometimes it looks at the portfolio[116], a few look even at the dependencies within the total expenditure portfolio[117]. Although these new valuation frameworks have brought tremendous efforts (Wijnia, 2004), most decisions are still plagued by large uncertainties in the prognoses. In financial asset management options are used to reduce the risk of uncertainties, and real options can be used for real asset management. However, so far real options are not commonly used in valuing the investment opportunities in electricity networks.

In this chapter we will explore the possibilities for the use of real options in valuing investments opportunities. To do so we will first describe the system in which the decisions about the infrastructure are made. Secondly, we will introduce a common decision problem for network managers. For this case we will use a number of decision criteria (both technical and economic), all resulting in a different outcome for investment timing. We then identify the uncertainties that drive these differences. A few of these uncertainties have been resolved over the years. Others have not. Finally, we will look at some opportunities for the real option theory to value solutions to overcome those remaining uncertainties.

5.2 System description

In this paper we look at decision making about investment opportunities in electricity networks. In this section we will describe the system in which the decisions are made and specify what we mean with an investment opportunity.

5.2.1 Investment decisions

Investment decisions in electricity infrastructures are quite different from financial investment decisions. Financial investments are supposed to deliver a series of future cash flows that, adjusted for the opportunity cost of capital, has a higher value to the investor than the initial payment(Brealey and Myers, 2000). An investment for which this is the case is said to have a positive Net Present Value (NPV). The decision rule is very simple: only make investments that have a positive NPV. Investment decisions in electricity networks cannot be based on such a simple rule. There are 3 major causes for this deviation:

1. It is very difficult to allocate cash inflows to specific network elements. Most elements can be removed individually without effects on the cash inflow, but removing all elements will stop the cash inflow. Most investments in the network do not directly induce new income, but only create the opportunity to connect new customers.
2. Investments that do induce new income are new connections. However, to protect the customer against monopolistic behavior the network companies have the obligation to connect anyone requesting it to the network against predetermined tariffs, even if the new connection has a negative NPV.
3. Investments often deliver benefits on other values than the financial one, for example on safety, sustainability or reliability. For those values the agreed upon monetary equivalent is

---

[116] PRIMO, a portfolio optimizing tool developed by TNO telecom and Essent Netwerk.

[117] For this a tool like “dependency” can be used (www.dependency.com).

only an order of magnitude. For example, a fatality is mostly valued between 1 and 10 million euros. This is not nearly accurate enough to be used in Net Present Value calculations.

Because of these differences with normal investment decisions we will call decision making in electricity networks risk management. This risk management consists of two control loops, as shown in the picture below.

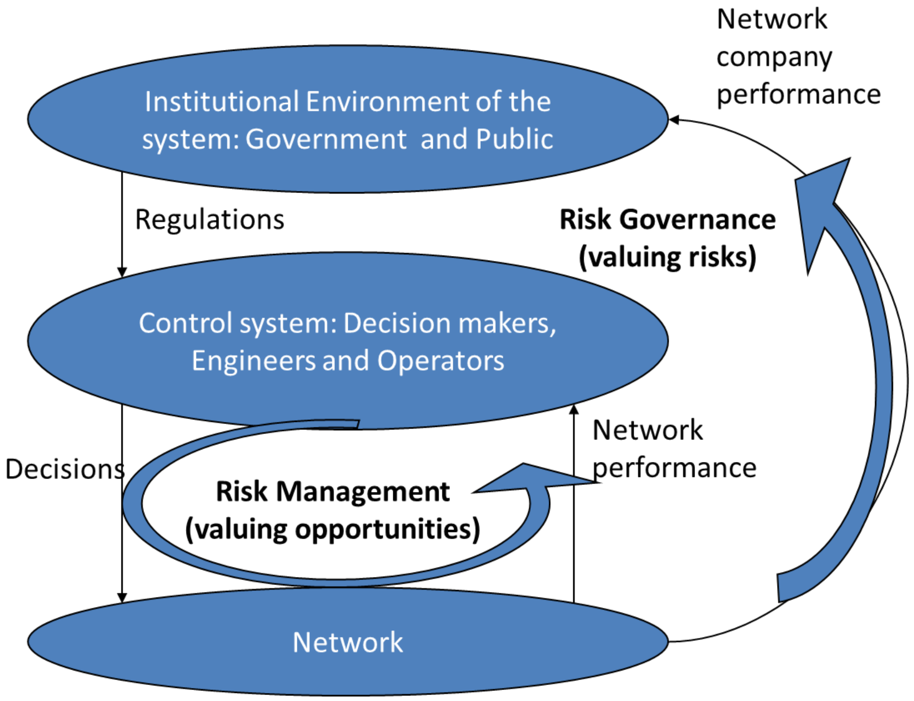


Figure 13: The risk management model.

The first control loop is risk governance in which is “decided” about the monetary equivalent of the non-financial values. Decided is put between quotation marks, as it often is no explicit valuation, but an implicit valuation that can be derived from the decision to implement certain regulations. If the costs and the benefits of those measures are known, the monetary equivalent can be calculated. Risk Governance involves multiple actors. The second control loop is risk management. In this loop the costs, the projected (non-financial) benefits and the monetary value of those benefits are calculated and tested for a positive NPV. For this calculation the non-financials have to be discounted at the same rate as the financials (Office of Management and Budget, 2001).

5.2.2 Investment opportunities

The problem space in the risk management cycle is very large. Although we only look at one case study, we do not want to tempt the reader into thinking that decision making in infrastructures is easy. To get a feeling for the immensity of the problem space the table below shows some dimensions of the problem space with the upper and lower limit on the dimensions (Wijnia and Herder, 2004).

Table 15 Dimensions and upper and lower limit of the problem space.

Dimension	Lower limit (example)	Upper limit (example)
Consequence of failure	Voltage sag	European blackout
Failure duration	10 ms (voltage sag)	Weeks (France 1999)
Probability (1/yr)	<10-6 (nuclear meltdown)	>10000 (medium voltage interruptions)
Intervention Size	100 € (switchgear maintenance)	>1 G€ (new high capacity power plant)
Complexity	Single criterion threshold (load level exceeded)	Multi criteria weighted sum (maintenance concept)
Scope	Technical (device settings)	Ethical (can we continue operating overhead lines if this might increase the probability of leukemia)
Time horizon	Tomorrow (new customer application)	25 years (network design for distributed generation)
Actors	Single actor single objective	Multi actor multi objective
Risk Perception	Perception in line with objective risk analysis	Perception deviating from objective risk analysis
Uncertainty	Almost certain	Uncertainty about consequences and likelihood
Ambiguity	Shared objectives	Conflicting objectives

In this table we can see 11 dimensions with upper and lower limits that can differ more than a factor 1 million. This makes it very difficult for the decision maker to keep the portfolio of activities consistent, or if that is not possible, at least coherent.

5.3 Case introduction

The case we will look at is a fairly common capacity problem. A village of 4.500 customers is connected to the high voltage grid with two medium voltage cables. These two cables are designed to be redundant, meaning that the village will not be interrupted if one of the cables fails[118]. The power consumption in the village is growing. A question every asset managers asks is when the cable section should be upgraded with a third cable. Figure 14 shows the basic outline of the problem.

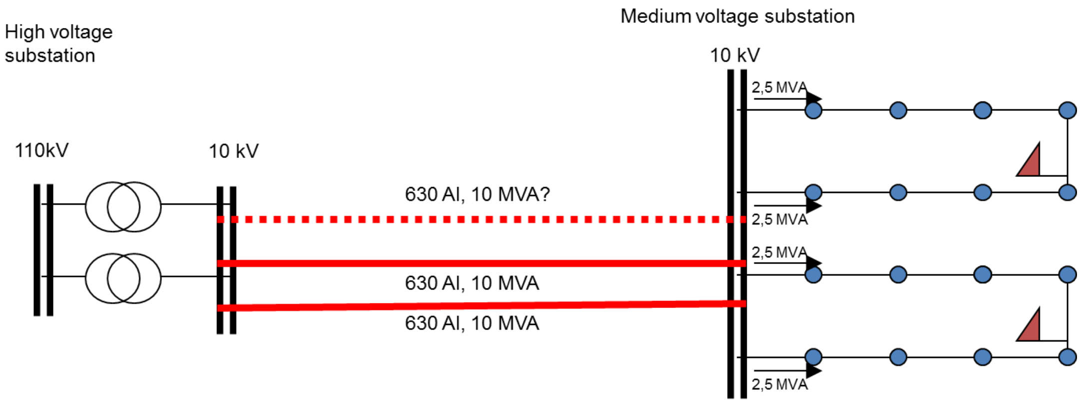


Figure 14: The capacity decision problem.

---

[118] A single cable has a failure rate of about 8 hours every ten years. The likelihood of the second cable failing in this 8 hour window is again 8 hours per 10 year, resulting in a total likelihood of about once every 10.000 years with an average duration of 4 hours (half the window). As both cables can be the first to fail the combined frequency is doubled. This number does not account for common cause failures, in which a single event takes out both cables.

In this diagram we see a redundant high voltage substation, to which the medium voltage substation is connected with two cables which have a nominal rating of 10 MVA[119]. The total load connected to the substation is 10 MVA and grows with 3% a year[120]. Figure 15 shows the development of power consumption in the last years and the prognosis of future demand.

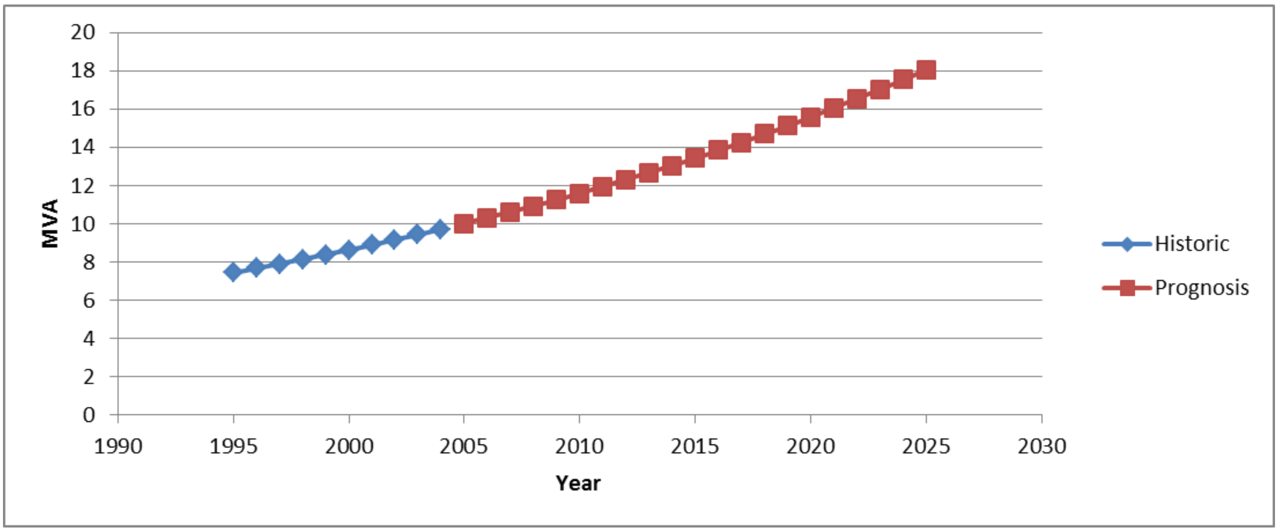


Figure 15: Load development.

5.4 Technical valuation

5.4.1 The classic approach: nominal rating

The most safe decision approach is to upgrade the circuit with a third cable once the nominal rating of a single cable is reached. This means regardless of the circumstances, the remaining cable will not be overloaded in case the other cable fails. In the example above this would mean to upgrade the circuit in the year 2005. This safe investment strategy therefore is an expensive one, as investments are made relatively early.

5.4.2 The enhanced approach: dynamic rating

The failure mechanism behind overload is temperature related degradation of the insulation material. This means the overloaded cable will not fail immediately after an overload occurs, but only if the temperature has reached a critical level. As the cable and the surrounding ground need time to heat up, an overload could be accepted for a limited time. Figure 16 shows the maximum allowed load levels (MALL) for some periods (IEC, 1985b, IEC, 2002).

---

[119] The real power capacity (in MW) is about 90% of the MVA rating.

[120] The actual growth rate of electricity consumption is estimated between 1% and 3% annually, depending mostly on economic development.

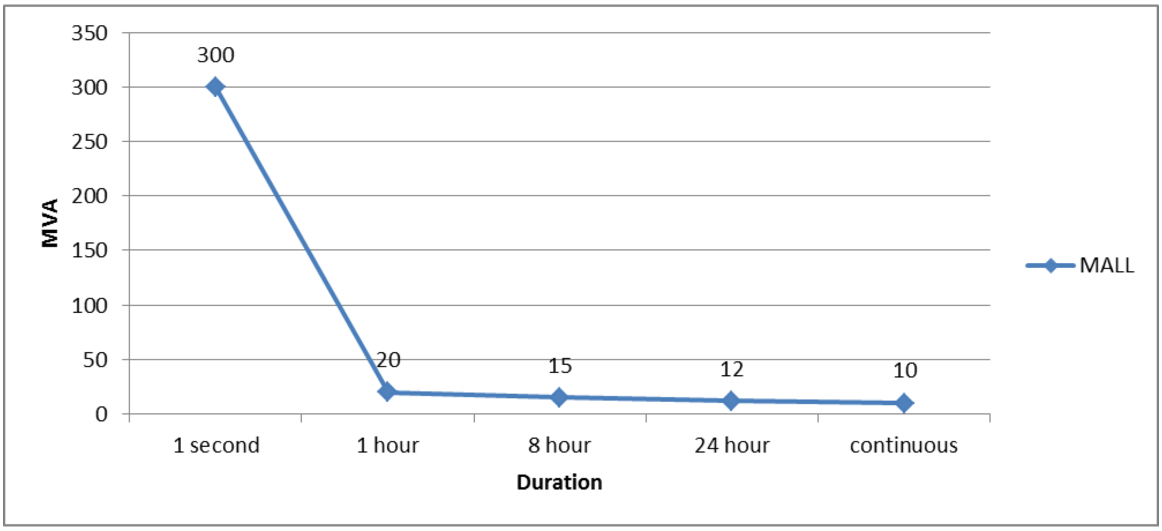


Figure 16: Maximum allowed load levels (MALL) for the electricity cable used in the example.

All failures in the cables will be repaired within 24 hours. In Figure 16 we can see that the cable can bear a load of 12 MVA for 24 hours.

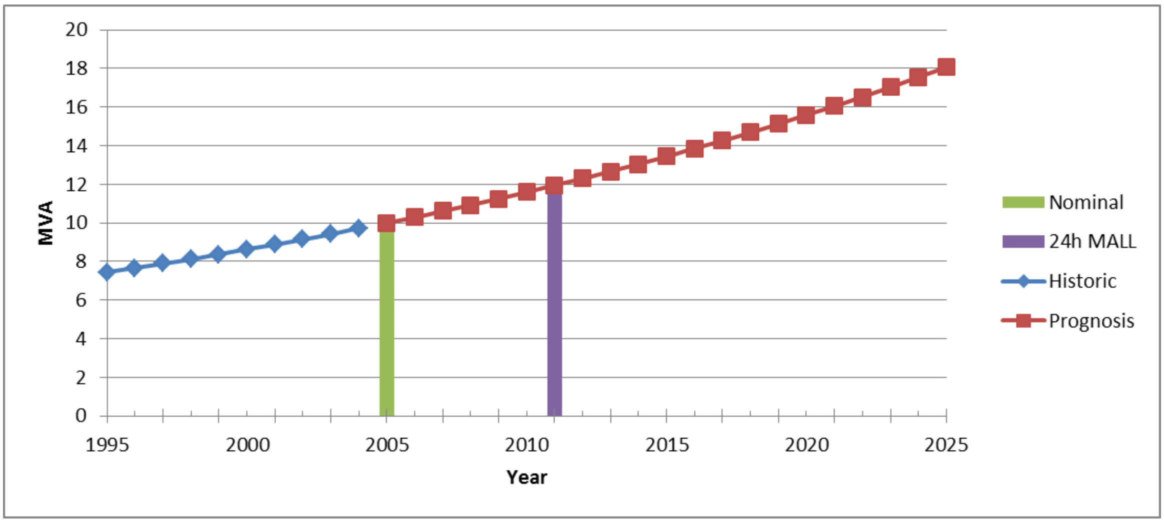


Figure 17: The delay of investment by changing the criterion from continuous to dynamic.

This means the circuit does not need upgrading before the load level has reached 12 MVA. This will only happen in 2011. We see in Figure 17 the replacement of the allowed nominal current by the allowed dynamic current delays the investment by 6 years. Only in case the cable will not be repaired within 24 hours there will be a slight increase in the risk.

5.5 Economic valuation

In the previous two decision approaches the decision criterion was technical. Could the customers be supplied if a failure occurred in one cable? In the decision no attention was paid to the value of the new cable. To value an investment in reliability a figure is needed for the value of reliability. As no real market exists for reliability this is a very difficult task. A figure commonly used for valuation of reliability is the economic damage caused by an interruption. However, this is a highly volatile number. A report by KEMA (KEMA, 2004) showed values ranging between 0,65 EUR/kW and 80

EUR/kW. For the quality regulation in the Netherlands the figure to be used is 0,25 EUR per customer minute lost (CML). With an average of 2 kW per customer and an average interruption duration of two hours this equals 15 EUR/kW, well within the specified range. In this paper we will use the figure expressed in EUR/CML, as the interruption duration in the case deviates a factor 4 from the average interruption duration in the Netherlands.

To use this economic value in the decision we need a model to calculate the expected unreliability. In the base case we will assume that the cable will fail if its load exceeds the previously specified value of 12 MVA. The value of an interruption will be 4.500 customers times 480 minutes times 0,25 EUR/CML equals 648.000 Euro. Such an interruption is expected once every 5 years[121]. The expected annual value of an interruption is then 129.600 euros. The economic value of the cable would be the net present value of this figure over 30 years (the life expectancy of the cable), resulting in a value of 2.111.040 euro[122]. As the costs of a third cable are about 1.000.000 euro[123], the benefits clearly exceed the costs.

If in case of a failure only the part of the load that exceeds the maximum allowed load level is interrupted, the economic damage could be limited very much. This is called load shedding. In Figure 18 the cost of load shedding are compared with the annuity of the investment (30 year, 4,5% real interest) in a new cable. For the maximum allowed load level the safe value of 12 MVA is used. We can see that there is no cost of load shedding until 2012, the year in which the load exceeds 12 MVA. Load shedding is the low cost alternative until the year 2025. This means the use of a very simple risk calculation method will postpone the investment in a new cable until the year 2025, a delay of 20 years compared with the classical approach.

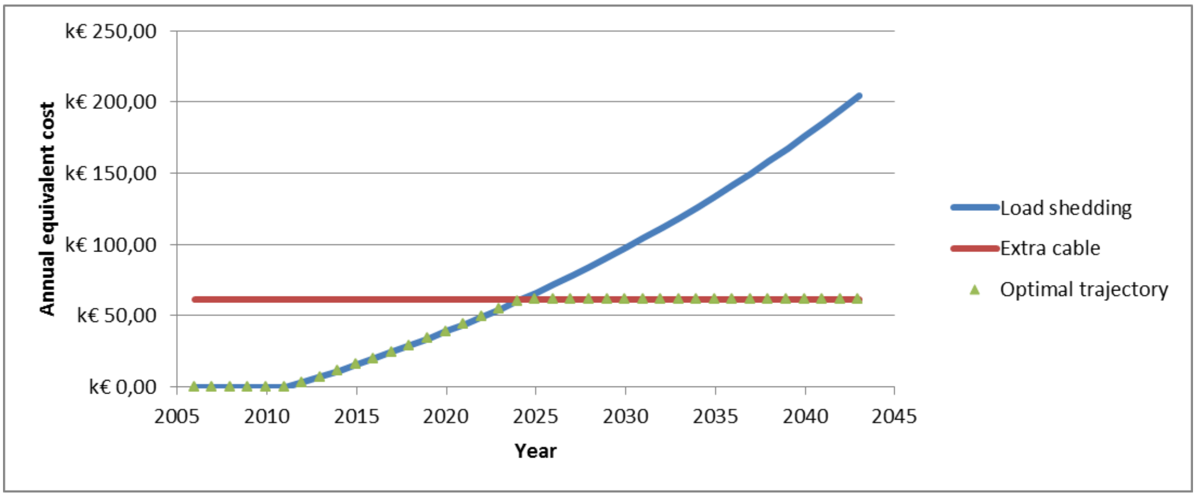


Figure 18: Optimal timing of investment under load shedding.

---

[121] The failure rate of a single cable is 1 per 10 years. As there are 2 cables, the failure rate of the circuit is 2 per 10 years or 1 per 5 years.

[122] Based on a real interest rate of 4,5%.

[123] Cables cost about 100.000 euro per kilometre

5.6 Valuing load shedding

The previous section showed a possible delay of investment in new capacity of about 20 years. This would solve the expected problem of the income cuts for the network managers. They even would not have to think about investing in new capacity for at least 15 years. However, this only holds if load shedding is valued equally as the unintended power interruptions. There is evidence this is not the case. After the energy crisis in California the market was restructured to prevent any further rolling blackouts. The costs involved in this restructuring was a thousand times more than the expected economic damage resulting from those rolling blackouts(de Bruijne, 2004). This would result in a value of 250 euro/CML. If this value is used for calculating the optimal investment moment it will not result in an extra delay compared with the dynamic rating example. In the following sections we will not consider involuntary load shedding any more. Instead, we will look further into the uncertainties of the risk model.

5.7 Uncertainties

In the previous examples we did assume certainty in modelling the decision. Unfortunately we are not at all certain about the behavior of cables and loads in reality.

The first major deviation is the load pattern. The model assumed a continuous load level. In reality the load shows a pattern. In Figure 19 the load pattern of the described case study is shown. The maximum load of 10 MVA is only reached one hour a year. The average load level is only 5,25 MVA. The peak load occurs only in blocks of 4 hours, so with the current load pattern the limitation to the 24 hour MALL is very risk averse. Even the use of the 4 hour MALL combined with the maximum occurring load level would be risk averse.

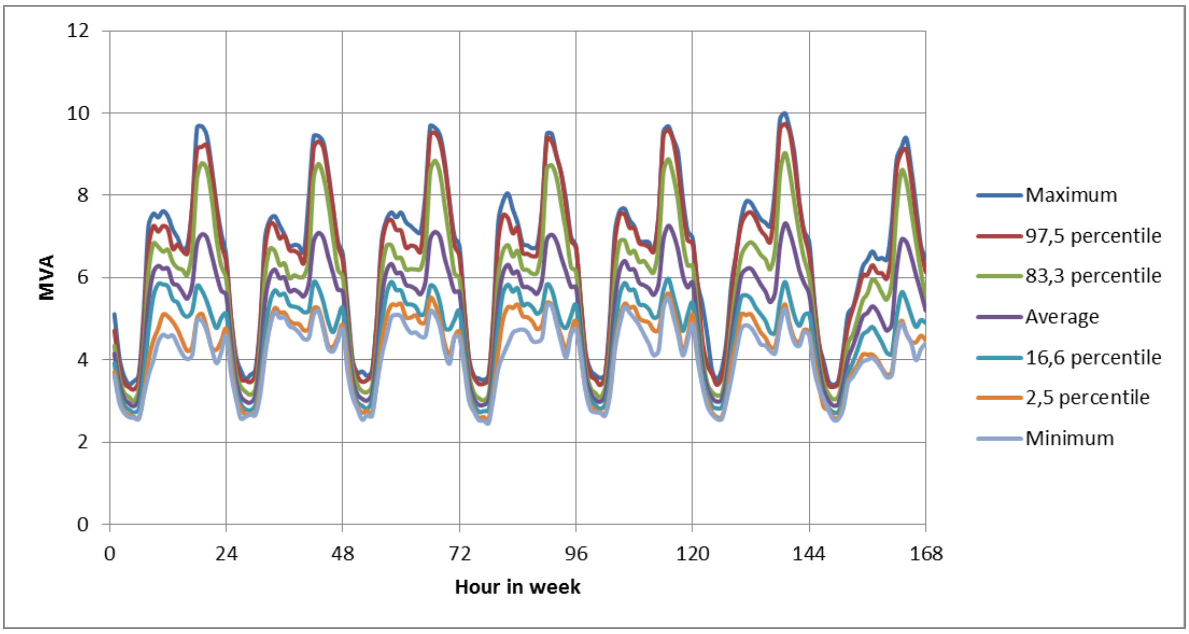


Figure 19: Actual load pattern for a village recorded at Enexis. The diagram shows the distribution per hour of the week. For example, the maximum load of 10 MVA occurred at hour 140 of the week, which is between 19.00 and 20.00 of a Saturday. However, for 43 (or 83%) of the 52 weeks the load in this hour 140 did not exceed 9 MVA.

The second deviation is the timing of the incident. In the model we assumed that the interruptions were evenly distributed over the years. This is in accordance with cost benefit analyses(Sugden and

Williams, 1978) and net present value calculations(Brealey and Myers, 2000). However, in reality either a large interruption occurs, or no interruption at all. There is a reasonable likelihood that the circuit will not experience any failure in its lifetime. For a circuit with a failure rate of 5% per year the probability of no interruption at all in 30 years is more than 20%. It is clear that in an perfectly working circuit the redundant cable has no value.

The third deviation is the actual economic damage caused by the interruption. As we have seen the damage can vary by a factor 100, depending on the type of customer connected. Optimizing the reliability on the average value would lead to a too low quality in areas with a high concentration of vulnerable customers, and vice versa.

The last major deviation is that the Maximum Allowed Load Levels are safe figures, at which a cable will not fail for certain. The model assumed that cables would fail for certain if the load exceeded the figure. This is most likely not the case. Unfortunately, the relation between load level and failure rate is not known very well. In some cases the load can be up to 200% of the MALL without causing any damage, in other cases the cables might fail below the continuous level, if hot spots[124] are present in the circuit.

5.8 Uncertainties resolved

In this section we describe the methods that are currently in use to resolve two of the uncertainties.

5.8.1 Load pattern

A way to overcome the uncertainty in the load pattern is to replace the peak load characterization with a thermal equivalent load. This is the continuous load which dissipates the same amount of heat in the cable as the load pattern. As heat dissipation is proportional to the squared current, the thermal equivalent is the Root Mean Square (RMS) value of the pattern. This could be evaluated against the MALL. In the example, the maximum has a RMS value of 6,9 MVA. This is shown in Figure 20.

---

[124] A hot spot is a section of the cable in which the heat dissipating capacity is limited. This might be caused by the proximity of other cables, tree roots, air ducts or soil with a high heat resistance.

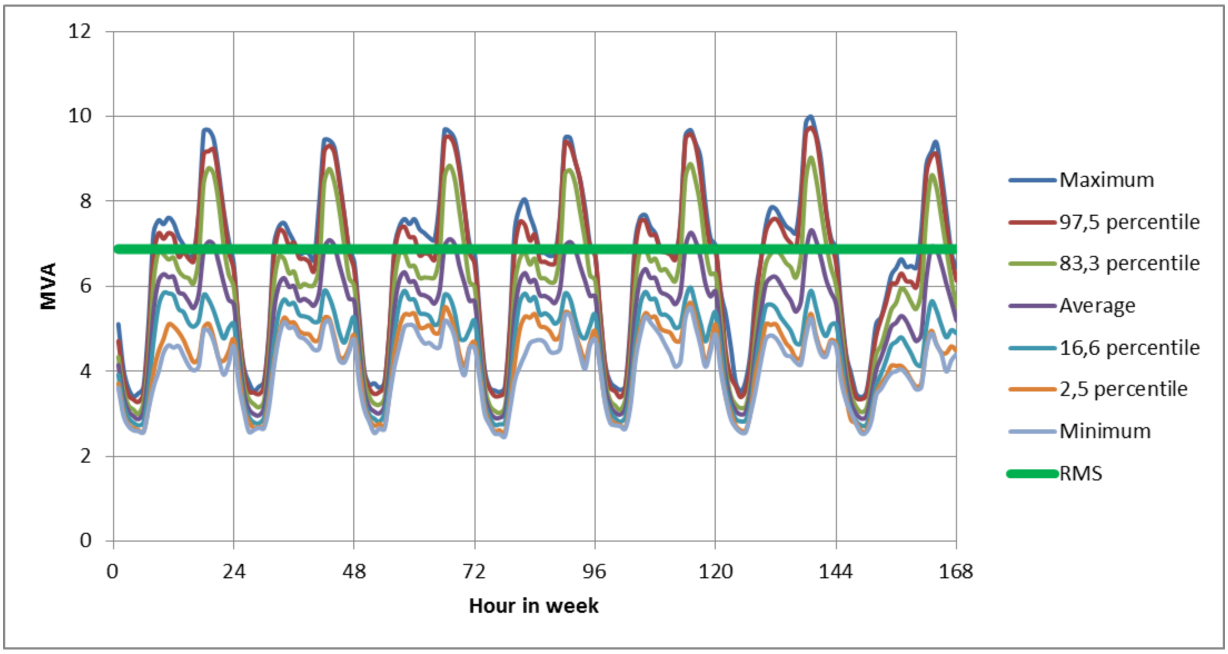


Figure 20: The RMS value of the maximum.

If this RMS value is used in combination with the 24 hour MALL, it is highly unlikely that in case of a cable failure the other cable will fail because of an overload. Figure 21 shows the optimal investment timing.

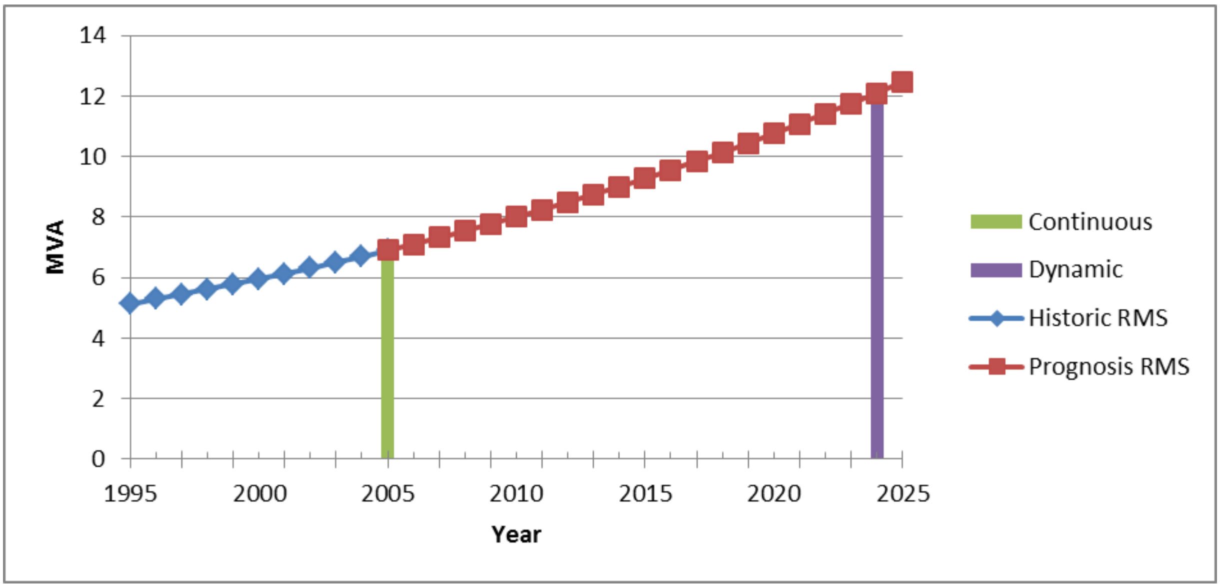


Figure 21: Investment timing based on RMS value of the load.

The use of the RMS value delays the investment until the year 2024. This is again a technical judgement, in which no attention is paid to the costs of the investment, or the expected costs of an interruption.

5.8.2 Incident timing

In our base case model we assumed that every year a small interruption would occur, whereas in reality it would be a large interruption or no interruption at all. For a single circuit with a failure rate of 5% per year the probability of a worthless redundancy for the next 30 years is more than 20%.

Fortunately, if one looks at multiple circuits at the same time, the behavior of this set is represented correctly by the sum of the average interruption sizes of the individual circuit. As networks are a large set of circuits (the three large network companies in the Netherlands each have 2,5 million customers, corresponding to 500 circuits as described) this uncertainty automatically resolves itself if looked at from the portfolio perspective.

5.9 Options for real options

Of the 4 major uncertainties 2 are more or less resolved currently, which was described in the previous section. However, 2 major uncertainties remain. The first one is the true economic damage caused by an interruption, the other is the true capacity of the cable. In this section we will describe a few solutions that may require a real option valuation.

5.9.1 True economic damage

An approach to get a better estimate of the true economic damage in case of an interruption is to characterize the customers in the area for their vulnerability. However, as the valuation of the involuntary load shedding might differ by a factor 1.000 from the true economic damage, it seems reasonable to question this approach for its validity.

Another approach is the use of load management contracts. A load management contract is an agreement between the network company and a customer, in which the customer commits itself to reduce power consumption when requested for by the network company. The network company pays a fee for this right. However, negotiating such a load management contract is quite a challenge, as it is vulnerable to all sorts of strategic behavior. For example, a customer would try to get a fee as high as possible, regardless of the economic damage caused by the interruption. The price of the contract therefore would not be a measure of the true economic damage, but it would be a true measure of the value of the interruption. The limitations in this game are this expected damage as the lower limit and the worth of the investment postponement for the network company. To be able to play this game correctly the network company needs to value the investment postponement, for which the real option theory seems the appropriate instrumentation. However, it is not clear yet to those network companies how the theory should be applied.

5.9.2 True cable capacity

The true cable capacity is mostly determined by hotspots. At the moment the technology is available to measure the temperature of the cable real-time. This would allow the operator to monitor the cable for overload situations, and thus preventing a failure caused by such an overload. The costs of such a built-in thermometer are relatively modest. If such a technology was used in combination with the current design practices, it would become valuable in 20 to 30 years. However, its worth is yet unknown.

5.9.3 Circuit upgrading options

A final opportunity for the real option theory is to value a cable which can be upgraded relatively cheaply. For example, using a 20 kV cable instead of a 10 kV cable allows for doubling the capacity if needed. It would require two 10/20 kV transformers at both ends of the cable, but for long stretches of cable this would cost less than a new cable. The question is if the value of the option to expand relatively cheaply exceeds the extra cost of the cable.

5.10 Conclusions

Network managers try to postpone investment in new capacity because of the income cuts posed upon them by the regulator. If the investment strategy for new capacity was based on maximum allowed load levels for continuous loads and the occurred peak load there are lots of opportunities to delay investment. This can be reached without any economic valuation, just by using common sense to review the investments. However, not all uncertainties can be resolved by using common sense. In case of load management contracts, cable monitoring equipment and circuit upgrading options the real option theory seems to be the appropriate tool to value the investment opportunities. If the real option theory could be used to value these uncertainties properly, it would be a great support for the decision makers of the future.

5.11 Epilogue

The optimization of timing showed a large sensitivity to the precise definition of the risk, and thus demonstrated the relative indeterminateness of the optimum. An optimal moment is essentially only the midpoint of a window of opportunity. Similar results were found in the optimization of the replacement moment of assets (Chapter 9), which was worked out in more detail (with reviewing real options) in “dealing with uncertainty in asset replacement decision” (Wijnia, 2013)

That imprecision of the optimum not only applies to timing. The network used in the case in this chapter was also used in a slightly modified form[125] in the paper on integrating sustainability into asset management (Wijnia and Peters, 2008). The consideration of that paper was not the timing of the investment, but the choice of technology for constructing such a circuit.

Table 16: Comparing TCO for different cable diameters (after table 13 from (Wijnia and Peters, 2008))

	Cable 240 Al	Cable 400 Al	Cable 630 Al
Cable costs /m	15€	20€	25€
Total cable costs (6*4500m)[a]	405 k€	486 k€	567€
Excavation costs (2*4500 m)[b]	135 k€	135 k€	135 k€
Energy loss (GWh/yr)	0.31	0.19	0.12
Energy costs	16k€	10k€	6k€
NPV energy loss [c]	310k€	190k€	120k€
Total footprint (NPV)	1951	1457	1335
Footprint value[d]	195 k€	146k€	134 k€
Total Economic costs(a+b+c)	**850 k€**	**865k€**	**930k€**
Total business costs(a+b+c+d)	**1045 k€**	**1011k€**	**1064k€**

NPV calculated for perpetuity at 5% in real terms. There is no load growth in the system

As can be seen, in pure economic perspective the cost difference is some 10% between the two extremes. If the environmental impact is added, the difference is reduced to 5%. This is much smaller than the uncertainty that surrounds decisions. For example, the NPV of the energy losses was calculated using a 5% discount rate, the WACC at that time. Currently, the cost of capital is more like 2,5%, which would double the NPV of the energy losses, and reverse the order of alternatives completely. Assuming a load growth of 1% a year would have even more dramatic effects, as it would not only increase the cost of energy losses, but also would require an extra cable, which could be avoided by using a bigger cable right now. The key drivers in the decision are thus assumptions about

---

[125] Cables were shorter, no load growth.

the future, not facts about costs and benefits. It has to be stressed that if differences between alternatives are not that large, any of the alternatives would suffice. It is just that some options might be slightly better in certain circumstances. Cost benefit analysis helps to identify reasonable options, the technological window of opportunity, but within this window other arguments and beliefs determine the choice.

Combining both the overload case (assuming a growing demand) and the economic evaluation of energy losses results in a very interesting result as presented at the urban deltas conference in Sydney 2013 (Wijnia and Herder, 2013). From an economical viewpoint, the construction of an extra cable would only be justified if the economic value of the reduction of energy losses was larger than the cost of that extra cable. For the numbers used in the examples, this typically only will be reached if the cables are used to their full capacity in normal operation. This is shown in the left diagram of Figure 22. The value of the loss reduction equals the annual equivalent investment costs around 20 MVA, the full capacity of the circuit. The right diagram of Figure 22 gives the overload risk. This is the product of the probability of failure of any of the cables (10% per year), the probability that within the mentioned 24 hour repair time of the cable the load will exceed the allowed level of 12 MVA for at least one hour and the consequences of the resulting interruption. Below this level of 12 MVA there is no overload risk. This risk maxes out above a certain value, as the load then will always exceed 12 MVA in any 24 hour window.

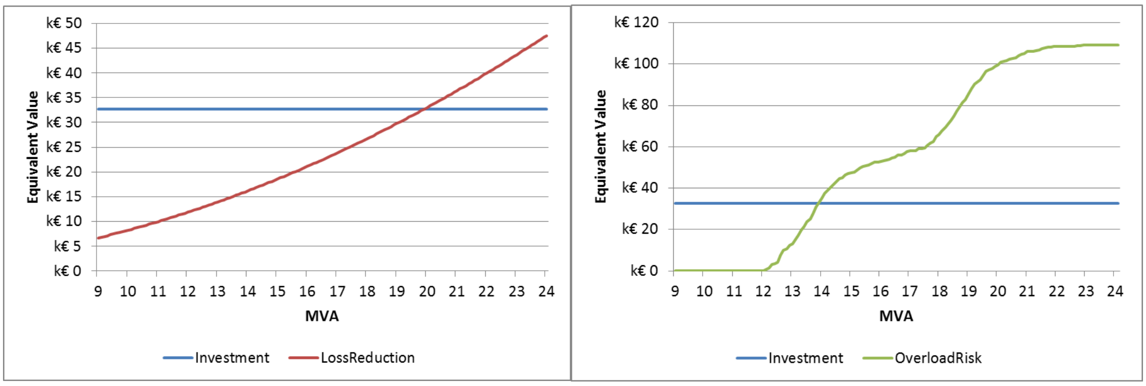


Figure 22: Loss reduction and (worst case) overload risk compared to the annual equivalent investment costs of a new cable as a function of peak power in MVA

One extreme (the best case) in decision making is to neglect the overload risk (equivalent to not valuing the consequences). This would result in an optimal load for upgrading the circuit of (in this case) 20 MVA for the circuit. In reality, there would be some risk, so this really is an extreme. It cannot be better. The other extreme would be to take the right diagram as a fact and add the costs to that of the energy loss. That would result in an optimal upgrade value of something close to the allowed level. Again, this is an extreme, as the cable needs time to heat up in order to fail, which will not happen if only for 1 hour the allowed load level is exceeded. The failure risk thus is lower than the calculated curve. Figure 23 holds optimizations for both scenarios. The difference between those scenarios is quite significant. This may provoke research for a better understanding of the failure behavior under overload.

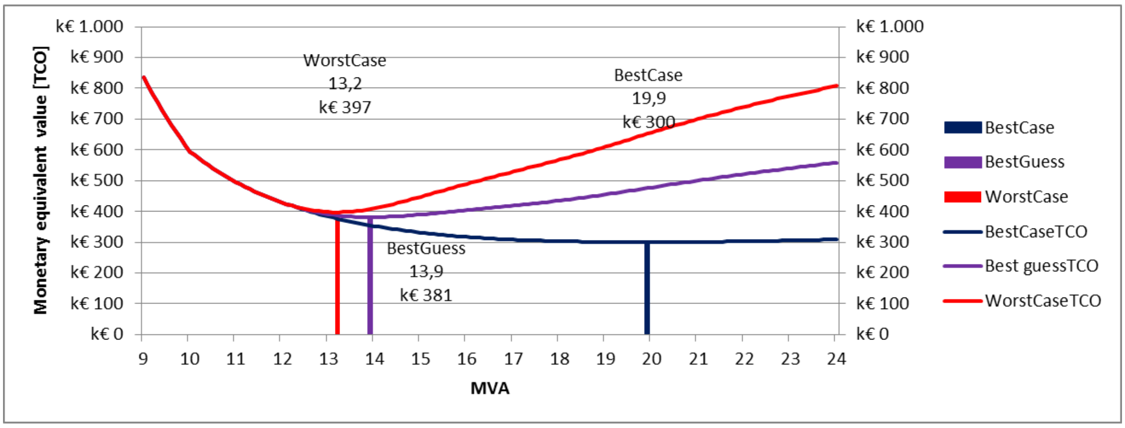


Figure 23: Comparing optimizations for the worst case, best guess and best case scenarios as a function of peak power in MVA.

However, reality must be in between those extreme scenarios. As a best guess the average between worst case and best case is taken, assuming 50% of the worst case risk is real. Plotting the optimization curves relative to their respective optimal costs reveals that there is an upgrade level that performs reasonably well in any of the two extreme scenario's. This is shown in Figure 24. Upgrading at 14,8 MVA will not be more than 11% more expensive than the best possible cost in either the best or the worst case. This enclosed optimum, where the value loss against best case and worst case scenarios is equal, can be regarded as the option with minimal maximal regret. As reality will be in between the extreme scenarios, the difference with "real optimum" will be less. For example, against the best guess scenario the value loss of the enclosed optimum is only 2%. This potential deviation from the "real optimum" is much smaller than many other uncertainties surrounding the decision. It may be argued that this is a lucky combination of input variables, but the results hold for a wide variety of combinations. This means that even without knowing the precise failure mechanism very robust optimization decisions can be made.

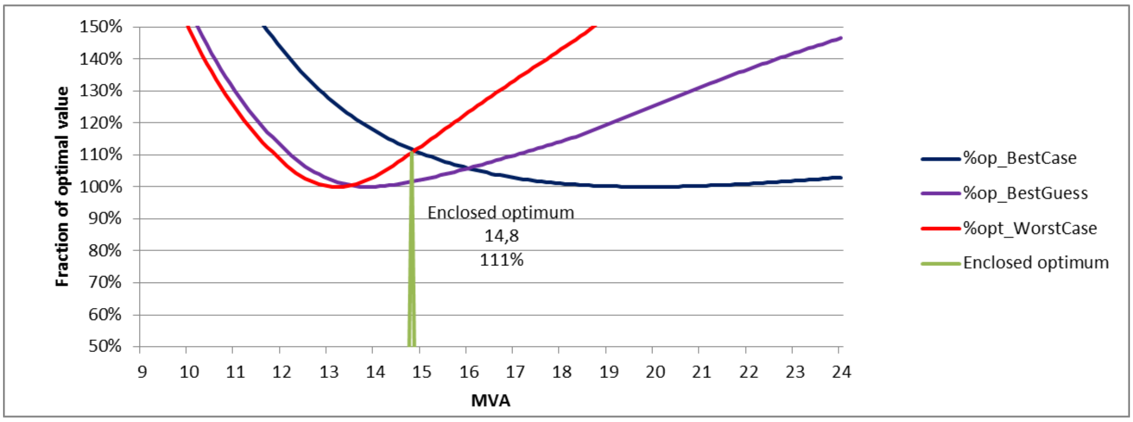


Figure 24: Development of relative value of optimization per scenario.

This again demonstrates that trying to find the real optimum with a cost benefit approach easily turns into a debate about beliefs and assumptions, whereas from a practical viewpoint any alternative within the window of opportunity suffices. This can be regarded as the paradox of decision making: the more alternatives resemble each other in terms of total value, the more difficult it becomes to choose between them on a rational basis and the less important it is which alternative is chosen.

6 Portfolio optimization

This chapter was published before as **PRIORITIZING INVESTMENT: THE VALUE OF PORTFOLIO DECISIONS IN ELECTRICITY INFRASTRUCTURE MANAGEMENT (WIJNIA AND WARNERS, 2006)**[126]

6.1 Introduction

All over Europe the energy distribution network companies are confronted with tighter budgets, as the regulators place efficiency targets on those monopolists (Electricity Association, 2000). The distribution companies react both by true efficiency measures (can the same result be achieved for less money) as by effectiveness measures (is the investment really necessary for the company). The latter question is the most difficult to answer, as most investments in the distribution system are measures to control risks that have a relatively small probability and large consequence. Therefore, for each individual investment it is safe to assume that it will not be a problem to postpone the investment one year. However, for the system as a whole the risk will increase substantially. This means it is vital to include a quantitative risk measure with a minimum yield ("risk reduction per euro") requirement into the investment selection. Unfortunately, this is not straightforward. Determining the risk for an individual investment is complicated by uncertainties (Klinke and Renn, 2002, Merkhofer, 1987, Morgan and Henrion, 1992), making the risk assessment more of an art than a science(Morgan et al., 2000). Therefore, the yield requirement is meaningless if it is to be more than a order-of-magnitude type, but the yields of the investment proposals usually are in the same order of magnitude. Besides, setting a yield requirement does neither tell in advance whether the budget will be exceeded, nor if the resulting system performance is still adequate. A possibility to overcome these last two problems is the portfolio decision(Brealey and Myers, 2000). The decision is then made over the total of investment proposals, which shows directly if the budget is violated. If risk is included into the portfolio decision it can show the effects on the system performance, but this requires that the proposals can be ranked to their yields. As in most cases the ranking is less sensitive to uncertainty than the absolute figures, this has the additional benefit of increasing the certainty about the best value for money. Finally, if the figures are reliable enough, the portfolio approach facilitates the discussion about the level of the budget. This paper describes the approach used to fulfill the requirements of the risk based portfolio decision and the results a Dutch energy distribution company achieved with it.

We start with the analysis of the design challenge. In section 3 we formalize this design challenge into equations and a tool. Finally, in section 4 we share the experiences with the new decision making approach. We end this paper with a few recommendations for further development.

6.2 Making decisions

Policy development can be seen in many different ways, ranging from a rational design process, mixed scanning, garbage can decision making to rounds of negotiation. Although all views agree on the difference between problems and solutions, the debate is on the chronology. Some assume that problems exist for which we invent solution, but others think the solutions exist for which we invent

---

[126] Wijnia, Y. C. & Warners, J. P. 2006. Prioritizing investment. The value of portfolio decisions in electricity infrastructure management. *29th IAEE Annual International Energy Conference 2006: 'Securing Energy in Insecure Times'*. Potsdam.

problems. What is important is the conceptual difference between problems and solution. We use this difference to distinguish between two levels of decision making in policy development. On the first level, problem valuation, it is decided how much one is willing to spend on the problem, on the second level the best solution within the constraints is chosen (solution valuation). Those levels are intertwined, as the problem valuation determines the feasibility of the solution, but the (non)existence of solutions might very well alter the problem valuation. In case of a distribution network, the decisions are repeated year after year, and experiences from past years feed the process. We therefore see decision making as two intertwined cycles, as presented in Figure 25.

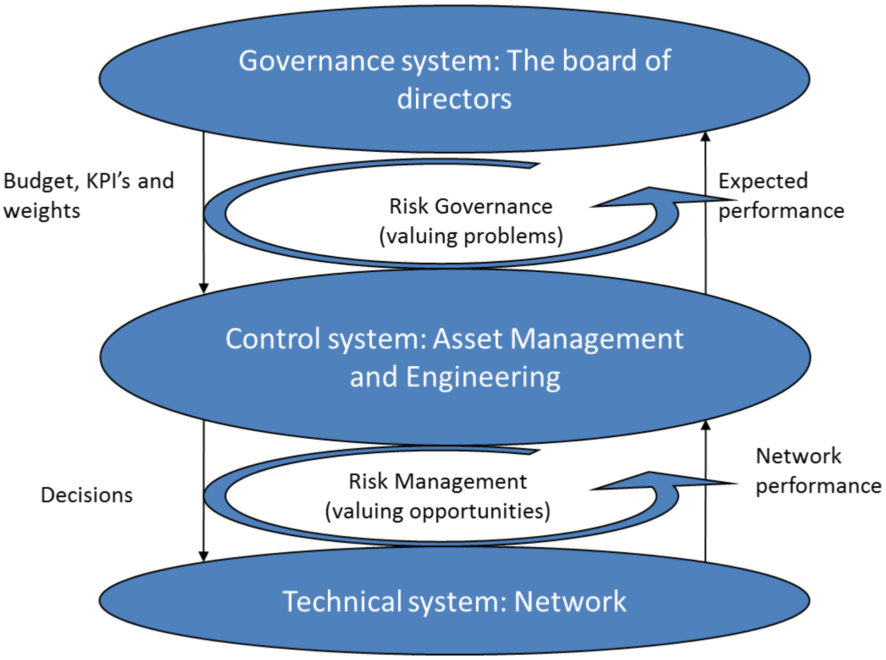


Figure 25: The policy development model, based on Wijnia and Herder (2005).

Due to its evolutionary nature, the process of policy refinement is necessarily slow. Unfortunately, the changes in the environment seem to demand a much faster policy adjustment. Therefore a redesign of the decision making process is required. Instead of setting a policy and waiting for investment proposals to show up, the chain could be reversed, by gathering as many potential investment proposals as possible. To structure this search, the process started with a thorough risk analysis (Wijnia, 2005). However, this process redesign also required a new type of decision. Whereas in the previous process investment proposals can be reviewed on a case by case basis, in the new process a selection out of hundreds of proposals is needed. A complication is that the proposals are on different disciplines, different values and different problems, so that no single individual can be expected to master the overview to make a proper selection. Therefore, a more systematic, quantitative approach seems required. In the next section we describe our approach.

6.3 Problem formalization

Many methods are available for supporting decision making and setting priorities over seemingly incomparable alternatives, e.g. multi-criteria decision making techniques (Wijnia, 2005). There is no perfect method, and so the selection of a method depends on a number of criteria. Important criteria for us are:

1. to support integral and centralized decision making.

2. to objectify the decision making.
3. to minimize the extra expert effort (for example, no endless lists of pair-wise comparisons).

In this section we describe the way in which we approach the portfolio selection.

6.3.1 Process

Our ultimate goal is to decide which projects to carry out, such that the overall yield is maximized, under specific budget constraints. In order to do this, we must clearly define what we actually mean by this, i.e. a measure must be defined to specify the yield of an individual project.

We propose the following approach:

1. Specify the Key Performance Indicators (KPI) measuring (or quantifying) the portfolio.
2. Determine for each of the projects its impact on each of the KPIs.
3. In addition, specify the current performance on each of the KPIs, and possibly a desired performance.

How to implement these steps is not discussed in detail here. Evidently, the management of a company should decide which are the important business values and corresponding KPIs. For each of the KPIs, a calculation method must be developed to facilitate step 2.

6.3.2 Requirements

As will be clear from the approach proposed in the previous section, it is preferable to know all projects in advance[127] to allow integral decision making.

In order to characterize a project in a single yield parameter, the following requirements must be met:

1. The relative importance of KPIs must be known;
2. A method needs to be constructed for comparing KPIs of different dimensions, since one cannot simply add the various numbers.

Ad 1. Various methods can and should be used in order to determine relative weights for the KPIs. We mention three methods which ideally should be used in unison, in order to establish a set of ‘robust’ weights:

- *Pairwise comparisons*: let management indicate the relative importance of any two business values on a scale of (for example) 1 to 7. By means of standard methods from literature, a set of weights can be derived (for example, Saaty’s AHP (Saaty and Vargas, 2001)).
- *Expert opinions*: given the full set of business values, let experts quantify the weight per business value.
- *Risk-based approach*: based on estimated probabilities of events, let management indicate how much money they would spend to minimize these.

---

[127] Note that the method allows dynamic modification of the project portfolio as well, since eventually a project is characterised by its yield independently of other projects.

Ad 2. The natural approach to facilitating comparison of items in different dimensions is to make them dimensionless. The method of choice is to relate the impact of a single project, to a *reference value*. This can be the actual (measured), the worst-case (when 'doing nothing') or desired performance, or a combination of these.

6.3.3 Mathematical optimization model

The overall impact of project *j* is calculated as follows:

$$IMP_j = \sum_{i=1}^{m} \left(w_i \times \frac{imp_{ij}}{ref_i} \right) \tag{6.1}$$

I.e., the impact is scaled with regard to the reference value and weighted with regard to the importance of the various KPIs.

In this equation, the following notation is used:

m	number of KPIs
w_i	weight of KPI *i*
ref_i	reference value of KPI *i*
imp_{ij}	impact of project *j* on KPI *i*
IMP_j	overall impact of project *j*

Now we are ready to define the optimization model. We introduce the decision variables x_j:

$$x_j = \begin{cases} 1 & \text{if project } j \text{ is carried out} \\ 0 & \text{otherwise} \end{cases}$$

To optimize the overall yield, the following model must be solved:

$$\begin{aligned} \max \ & \sum_{j=1}^{n} IMP_j x_j \\ \text{s.t.} \ & \sum_{j=1}^{n} c_j x_j \le b \\ & x_j \in \{0,1\} \end{aligned} \tag{6.2}$$

Here *n* is the number of projects. In addition,

cj	*cost/investment associated with project j*
b	*available budget*

The formulation (6.2) is known in the literature as a *knapsack* problem (see e.g. Nemhauser and Wolsey (1988)): the (single) budget constraint is the knapsack-constraint, and one aims to optimize the yield while the portfolio must "fit in the knapsack".

Obviously, other valid formulations can be constructed by considering the yield on each of the KPIs separately. In the remainder of this paper we will build on formulation (6.1).

6.3.4 Algorithm

While the knapsack problem belongs to the class of NP-complete problems, and as such is considered difficult (i.e., its optimal solution requires a running time exponential in the size of the input *n*), in practice it can be solved to near-optimality in polynomial time.

The obvious method is to consider the “yield per Euro” of a project:

$$Y_j = \frac{IMP_j}{c_j} \tag{6.3}$$

Obviously, the higher Y_j, the more attractive it is to carry out project *j*. The solution method for obtaining near-optimal solutions is the following. We call this the first-order heuristic:

1. Compute Y_j for all projects.
2. Sort the projects according to decreasing Y_j.
3. Add the sorted projects to the selected portfolio until the budget limit *b* is reached.

In terms of the decision variables that are introduced above, a solution is obtained of the form:

$$\overline{x} = (1,\ldots,1,z,0,\ldots,0) \tag{6.4}$$

The solution is binary except for at most one variable, indicated by *z*; it is known as the solution to the *linear relaxation* of (1), i.e. the integrality constraints in (1) are relaxed to linear constraints which express that all variables must be between 0 and 1. The objective value that corresponds to $\overline{x}$ is an upper bound to any feasible (binary) solution. By setting *z* to 0, a feasible solution to (1) is obtained.

The solution $\overline{x}$ (if not already binary) can be further improved by:

- Extending it with projects that have lower yield, but fit within in the budget constraints (we call this the *extended* first-order heuristic);
- Applying an exhaustive branch and bound procedure, until the optimal binary solution is found.

In practice, the latter is attractive only, when the difference between the objective values of the fractional solution $\overline{x}$ and the solution obtained by the extended first-order heuristic is substantial. For relatively small portfolios, with large differences in yield and investment per project this may occur frequently; for large portfolios this is unlikely.

6.4 Supporting the decision making process in practice

From a theoretical viewpoint, it may seem attractive that solving the mathematical optimization problem gives a definite answer as to which portfolio is optimal. However, as any decision maker with practical experience will recognize, not all intricacies can be handled sufficiently in the framework described in the previous section (indeed, *any* framework will probably turn out to be insufficient for fully automated decision making). In practice, the decision maker(s) will require to further influence the outcome of the process.

The tool and method that we developed, specifically take into account the wishes of the decision maker and provide various possibilities for fine-tuning the decision making process. We distinguish the following features:

- Allowing a layered input in which projects can be clustered into *programs*.
- Fixing projects and/or programs to be part, or not to be part, of the selected portfolio (irrespective of their yield).
- Providing insight into the overall yield as a function of the budget spent.
- Providing insight into the current (measured) performance, to show the minimum investment required consolidating the current performance.
- Providing insight in the impact on the individual business values as a function of the budget, vs. the impact on the overall yield.
- Providing insight in the maximum yield per program.
- Dividing the total investment over the various business values, in order to show how much money is spent to realize the given yield on the business values.

In Figure 26 and Figure 27 a number of these features is illustrated.

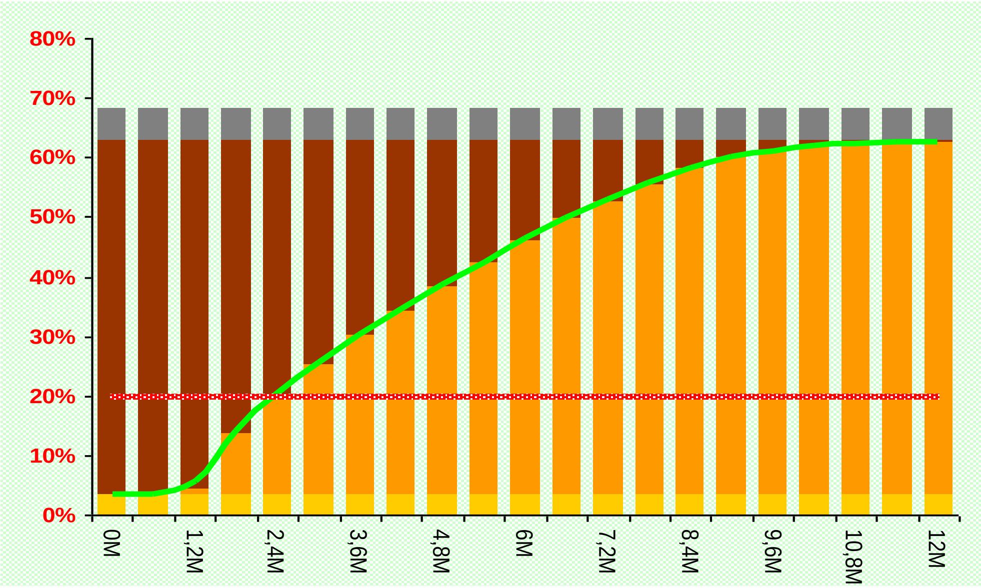


Figure 26 The yield-graph[128]: the green line indicates the theoretically optimal yield based on linear relaxation, the orange bars indicate the optimized yield corresponding to a feasible integer solution. The grey part of the bars correspond to the part of the yield that is ruled out by fixing projects to be not part of the selected portfolio. Analogously, the light orange part indicates the yield that is guaranteed by fixing projects to be part of the portfolio. The red line indicates the current performance. The 0% line is the worst case scenario, i.e. doing nothing more than the compulsory actions.

---

[128] This graph is based on randomly generated inputs.

In practice, it is recommended to use a decision support tool such as this in an iterative manner. The tool itself essentially requires a two-phase approach: in a first calculation phase, the yield graph as a function of budget can be calculated. In the second phase, the budget is fixed, and a more detailed analysis of the corresponding outcome can be made. Based on the outcome of this single run-through, the outcome should be examined using the features available. Based on expert opinions, some projects or programs will be fixed to be part of the portfolio, and possibly the budget will be adjusted. Subsequently, the two phases will be repeated (or possibly only the second phase). This process is repeated until a satisfactory portfolio decision is arrived upon. Using the tool in this way (that is, truly as a decision *support* tool), allows relatively fast decision making.

Note that so far, we have neglected the fact that the inputs are subject to a degree of uncertainty, since obviously we cannot predict the future. This is addressed in the next sections, where we describe some experiences from practice and argue that the outcome of the process is insensitive with respect to the uncertainty associated with the inputs.

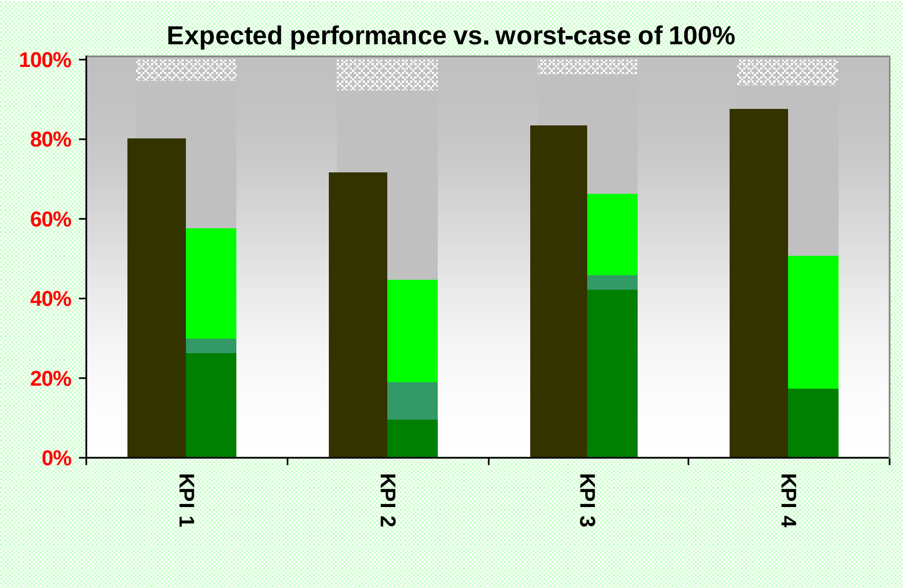


Figure 27: Illustration of the output of the tool: for each of the KPIs the expected performance is presented next to the current performance. The worst-case performance is normalized at 100% in this graph. In addition, the user can derive how much further the performance could be improved on the KPI (the light green bar).

6.5 Applying the portfolio decision

The portfolio decision was implemented in a Dutch energy distribution company. This implementation consisted of two steps. First, a set of weighted performance indicators was developed. The second step was to collect the opportunities to which the portfolio approach was applied. In this section we describe those steps. We end with a short evaluation of the process.

6.5.1 Setting the performance indicators' weights

A multi-criteria portfolio decision requires a set of weighted Key Performance Indicators (KPI). A risk based approach was used to get the values. Top management was asked to envision incidents per business value[129], ranging from negligible to catastrophic. Subsequently, they indicated the amount of money they were willing to spend to prevent one such incident from happening. The total willingness to spend per business value was used as a weight.

6.5.2 Selecting the portfolio

To prepare the decision the engineering staff had been gathering investment opportunities during a full year. The staff was split into sub teams corresponding to different parts of the organization. Some of the opportunities were based on a legal requirement and therefore compulsory[130]. All teams ranked the non-compulsory opportunities concerning their area of attention. Part of this ranking was based on asset performance models to calculate the effects on the business values. The rest of the opportunities were ranked by the teams directly. The total volume of potential expenditure was roughly 140 million euro, divided over about 600 projects. Of this number, about 90 million was fixed in the portfolio, either because of the compulsive nature or because it was selected directly by the experts. The rest was ranked by the yield. This is illustrated in Figure 28, where for the first 90 million euro of budget no effect on the business values is taken into account. After about 140 million of expenditure the performance does not change anymore, indicating that all projects are executed.

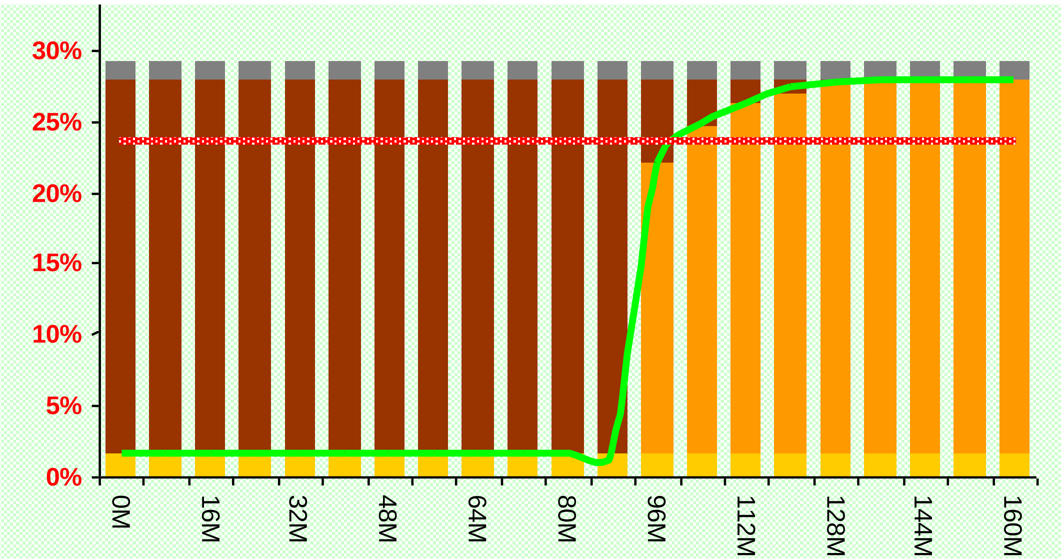


Figure 28: The yield-graph for the practical example.

The red line in the graph indicates the weighted performance level reached during the current year. This means that the “only compulsory expenditure” strategy would result in a 23% deterioration of the performance, compared with the actual performance. Spending about 100 million would keep the performance at the current level (though it could mean a different distribution over the key performance indicators). The maximum achievable performance would be a 7% increase, with an extra expenditure of about 40 million. The green line indicates the theoretical maximum for the

---

[129] The business values that were chosen were: Economics, reliability, safety and compliance.

[130] Strictly speaking those were no part of the ranking, but they were nevertheless included in the project list to arrive at a single budget decision.

budget based on the branch and bound algorithm, the orange columns show the maximum for the first order heuristic. There is almost no difference between the two, which can be explained by the distribution of project size. If one large project would not fit within the budget, it would be replaced by a large number of small projects with more or less the same yield, almost completely filling the gap.

A critical part for the approach was the selection of the combined portfolio. As the teams worked independently, the yields of their project could be unaligned because of differences in the modeling approach. This might result in the selection of all projects from one team and none of the other. Such a result would not be accepted. Therefore special attention was paid to the process of selection by making it a group decision. All team results were combined into one list, and a preliminary ranking was produced. This ranking was discussed in the group. In this first round the value of the this approach became clear. One of the teams found that a major part of their projects, that they considered as vital for the reliability of the grid, was not in the final result. The other teams agreed on the importance of the projects. Adding the corresponding KPI solved the problem. It should be noted that for over 95% of the portfolio no discussions arose.

After this initial round to check for consistency a few rounds followed to set a budget proposal. The budget was set and the teams checked if all their top priority projects were selected. If the team could accept the omission, that would be the final result, otherwise the budget would be increased, and the check process would be repeated. At a budget level of 120 million every team was satisfied. This proposal was sent to head office for approval, accompanied with the budget graph shown below.

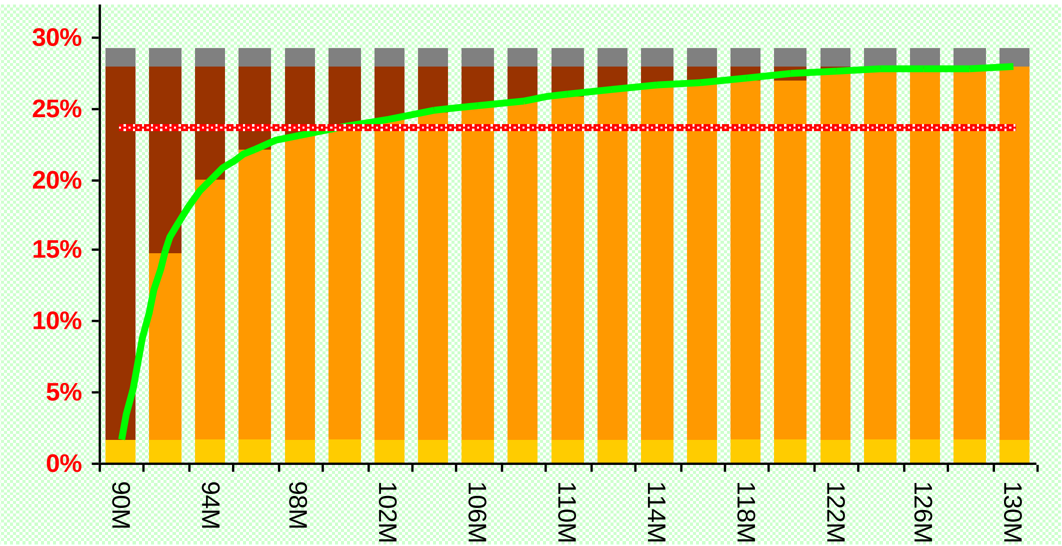


Figure 29: The yield-graph for the practical example, restricted to the relevant trade-off.

The proposal was almost completely accepted. Head office did not question the ranking. However, they did not see the need for such a performance increase, and would be satisfied with the performance reached by a 118 million euro budget. This was sent back to the teams, to allow for a few minor changes, and the selected projects were implemented.

6.5.3 Evaluation

In an evaluation of the approach most participants said they liked the portfolio decision very much, because they felt they were in control all the time. Especially the shown flexibility with KPIs and the possibility to hand-select perceived vital projects was found convincing. However, a few concerns remained.

The first concern was the expected sensitivity to data errors. To address this concern a sensitivity analysis was performed, which showed that the portfolio was almost 100% robust. Even random errors of about 20% in the data did not influence the portfolio performance, and only a few projects would change in selection status. This was explained by the very large spread in the yield of the projects (about 3 orders of magnitude) against the 20% error. That left us with the possibility of systematic errors, but those would not affect the ranking, only the total performance.

The next concern addressed the accuracy of the decision. In the decision we assumed all effects would happen in the next year. However, some of the projects in the portfolio would require several years for completion. The costs of those projects would be spread over the years, but the benefit would only be delivered in the final year. If we would include such a project in the portfolio, we would not use the full budget, nor reach the predicted performance. However, only including the effects for the next year would result in only costs and no benefits, meaning the project would never be selected. As some long term projects clearly had more value than some short term opportunities, we think our assumption resulted in the best decision and accepted the inaccuracy, but the dilemma remains.

A third concern addressed the comparability of the projects. We all agreed that the portfolio decision worked fine for projects within the discretion of the company, but some projects required consent from other parties. This created two risks. If we selected a project that would not be accepted by those parties, the project would not be feasible, requiring a new decision. However, if we would get the consent before the portfolio decision and not select the project, we would lose credibility to those other parties. We resolved this dilemma by calculating different scenarios including or not including these projects.

Finally, some of the projects were mutually exclusive (Wijnia and Herder, 2004). In practice this interdependency occurs only for a few of the projects, so we handled this problem by selecting one of the options manually and comparing the resulting portfolios. This would become quite a challenge if the level of interdependency or the number of projects increased. Therefore, it would be helpful if the interdependency could be indicated in the project properties section.

6.6 Conclusions

The portfolio approach was a successful attempt to shorten the decision making cycle. Even though it required more preparation time, this was recovered in the acceptance and feasibility of the proposed portfolio. The engineers experienced a better opportunity to get their vital projects through the decision making process, whereas the budget holder felt much more assured about the right level of the budget. Even the final budget cuts were accepted by the engineers, because it was very clear the budget holder was willing to bear the extra risk. Finally, the decision proved to be very robust, despite initial doubts.

6.7 Epilogue

The scope of the experiment was procedural optimization, making the best choice out of a set alternatives. The value of this procedural optimization was assessed in terms of the process. The value in terms of content (i.e. cost reduction for the same performance or performance increase for the same costs) was not addressed. There were two reasons for not doing so. First of all, it was simply not a necessity for evaluation of the feasibility, achievable consistency or acceptance of the portfolio decision. Secondly, there was no reference material available against which to benchmark the portfolio decision.

However, such a benchmark could be simulated with the data that resulted from this experiment. This is demonstrated in Figure 30.

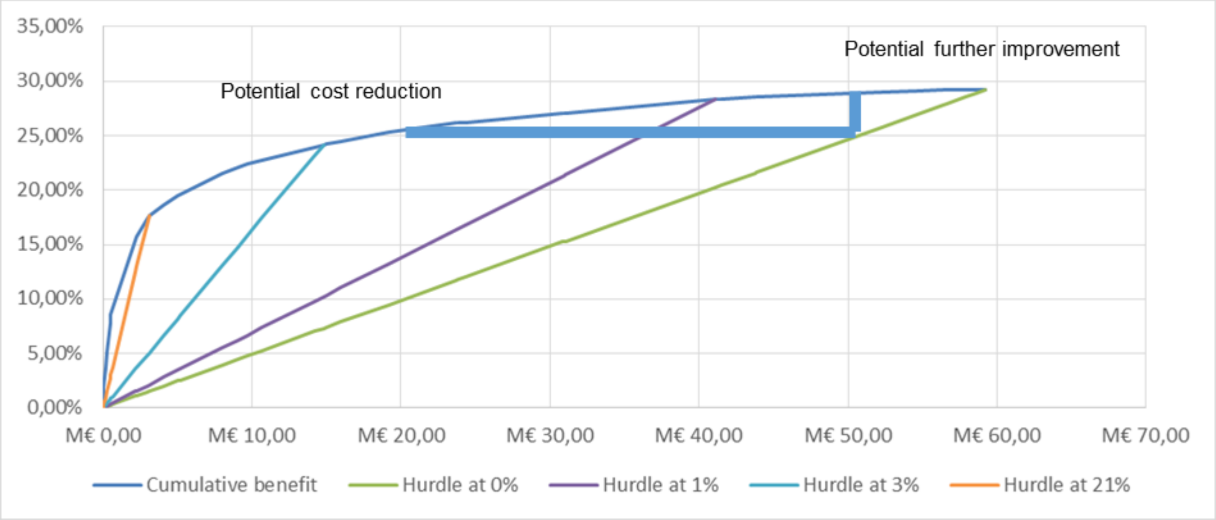


Figure 30: The development of value as a function of the budget of the portfolio decision and several serial assessments for different hurdle rates.

The portfolio optimization is represented by the dark blue curve in the diagram, the cumulative benefit. This is the same curve as that of Figure 29, with the budget scale starting at the ranked projects. Because the projects are sorted on the basis of their yield from high to low, this curve gives the highest achievable performance for any budget.

The alternative approach of decision making would be to assess projects on an individual basis (serial assessment) and approve or reject them based on two aspects:

1. The yield of the project passed the hurdle rate
2. The total costs of the approved projects would not exceed the budget constraint

As there would be no way to guarantee that high yielding projects would be assessed first, projects basically enter decision making in a random order. Assuming that all projects that would pass the hurdle would fit the budget, after assessing all projects some point on the portfolio optimization curve would be reached. But because of the random ordering of the projects, the buildup of value in the approved set of projects would be more or less proportional to the budget. In the diagram, these buildups are represented by the straight lines for different hurdle rates, expressed in % improvement per 10 Million euro.

If the hurdle rate and the budget would be a perfect match in the sense that the budget would be exhausted precisely by all projects passing the hurdle, the portfolio decision would have no value as both selection methods would result in precisely the same set of projects to be executed. But if that match was not perfect, two possibilities are open: 1, the budget was not exhausted, meaning additional benefits could be realized, or 2, during the process some projects that passed the hurdle would be rejected because of budget violations.

Suppose no hurdle rate was specified (the green line), but only a budget limit of 50 million. At 50 million the serial assessment would result in 25% improvement, whereas the (parallel) portfolio decision would result in 29% improvement. The portfolio decision thus could deliver 20% more value than the serial decision. In terms of the budget, the effect is much more dramatic. The same 25% improvement could be reached by the portfolio decision with only 20 million euro budget. For the other hurdle rate lines, a similar improvement can be seen.

In reality, decision making would not be that simplistic, at least, not for prolonged periods. Any decision maker would develop some understanding of the hurdle rate that would match the budget limit, or at least understand that the hurdle rate should be increased with tightening budgets. Furthermore, if relevant projects would be rejected because of budget limitations, it would be possible to reject a lower yielding project that was approved before, though that would not be welcomed by the engineer advocating that project. Both aspects mean that in a stable environment the optimal portfolio selected in a parallel decision on all projects could be approached by serial decision making. But in more dynamic environments, that approximation could be easily 10-20% off. As such an approximation would require the same information as the portfolio decision (i.e. keeping track of the yield per project) that would be very low hanging fruit to be captured.

7 Structuring risk

This chapter was published before as **MODELING INTERDEPENDENCIES IN ELECTRICITY INFRASTRUCTURE RISK (WIJNIA AND HERDER, 2004)**[131]

7.1 Introduction

In all infrastructures one can recognize many potential problems. In the electricity infrastructure, the main problems concern the reliability of the system. The fact that reliability is an issue could be witnessed in the summer of 2003, when large blackouts occurred in the eastern US (Lerner, 2003), Italy and Denmark/Sweden. Due to the hot summer, in Europe a shortage of production reserves almost led to rolling blackouts (like in California 2001 (de Vries, 2004)), which would have been necessary to maintain system stability. As those systems have been in operation for decades, the question arises why those risk were not mitigated properly before. It seems the experience should be long enough to understand these potential problems completely and the knowledge should be sufficient to take adequate measures. At least two answers exist why this is not the case. First, the normal accident theory (Perrow, 1999) states that no matter how safe you make a system, accidents will always occur. This is because harmful coincidences of incidents may occur that are not or less harmful individually. As the number of coincidences increases exponentially with the number of connected elements, for large systems (consisting of millions of elements) it is impossible to identify them all.

A second explanation may be found in the changing demands upon the infrastructure[132], which by nature is very inflexible[133]. Because of the inherent stability, one assumes that risks, once mitigated, will not appear again, but in reality this only holds only when circumstances are stable. If the environment changes, infrastructure managers try to react as well as possible to the new demands, but they may forget about the effect of that reaction upon existing controls. For example, part of the liberalization effort of the electricity markets in the Netherlands is unbundling of production and transport. This results in a more volatile geographical distribution of actual power production in the transmission grid, as the decision to operate a plant is no longer made by the system operator[134] but dictated by the market (de Vries, 2004). In terms of energy balance this is not a problem (energy can be transported over distances larger than the size of the Netherlands), but it is an issue in terms of voltage levels, as reactive power cannot be transported over the distances required. This forced the system operator to penalize reactive power consumption by the network operators (allowed for by the technical codes[135]). The network operators reacted by constructing large capacitator banks. However, besides providing reactive power to the system, the capacitator banks also absorb the tariff control signals. These signals switch the consumers' electricity meters between peak load and

---

[131] Wijnia, Y. C. & Herder, P. M. 2004. Modeling Interdependencies in electricity infrastructure risk. *1st Annual CZAEE International Conference "Critical Infrastructure in the energy sector: Vulnerabilities and protection".* Prague.

[132] Johnson (1980) points out that "Change is the mother of trouble".

[133] Next Generation Infrastructures, Subprogram Flexible Infrastructures.

[134] In the Netherlands TenneT is the independent system operator.

[135] Article 5.5.4.1 NetCode (version July 2004).

base load prices. A dysfunctional system could result in huge deviations between the bills and the actual power consumption of the customers. Blocking filters will resolve the problem, as they prevent the absorption, but forgetting these filters might result in damages much larger than the costs of the banks, filters or even the fines for reactive power consumption.

This example shows that sound decisions on a high system level can create problems on lower levels. Those problems may be worse than the original problem that the decisions were supposed to solve. An approach in which all coupled issues were integrated might have had a different outcome. However, to take such an integrated approach a coherent overview of the risks and the interdependencies between the risks is needed. At this moment no such overview exists. This paper explores the possibilities for creating such a coherent overview. As the different types of risk may number in the thousands, the first step will be to group the identified risks into a manageable number of risk areas. This paper will address the problems encountered in an unstructured approach towards grouping and identify possible benefits using a structured process. The second step would be to identify the interdependencies between the risk areas. The steps will be illustrated with material gathered at a Dutch electricity distribution company.

7.2 Theoretical background

7.2.1 Risk definition

In everyday use, many different meanings are attached to risk. It can range from the object at risk, the threat, the probability, the consequence and any combination of those meanings (Claes, 2001). In the technical sense risk is normally used to express the product of probability and consequence. However, this use of risk is based on the view that a risk has certain objective characteristics, whereas one can observe in everyday life that the valuation of a risk is not coherent with this technical definition (Slovic and Weber, 2002). For example, people protest against nuclear power and smoke at the same time, the latter activity being much more risky to the individual. This lack of consistency troubles the proper characterization of a risk[136], a challenge the National Research Council was asked to address. They concluded that “the purpose of risk characterization is to enhance practical understanding and to illuminate choices” (Stern et al., 1996). In this paper, risk will therefore be defined as “A potential problem that needs to be decided about”.

7.2.2 System boundary

7.2.2.1 Decision making system

According to the National Research Council, risk analysis is a decision driven activity. The first system boundary should therefore address the decision scope of this paper. The assumption is that a single decision maker can decide to combine two different decision problems, presented separately to him, into a single decision problem. If those two problems would be presented to different decision makers they first would have to negotiate about combining the problems, which would probably result in some form of strategic behavior. This would significantly complicate the concept of more-or-less technical interdependencies that this addresses. Therefore this paper only discusses a single actor perspective.

---

[136] Characterization means valuing a risk, for example as high, low or negligible.

7.2.2.2 Technical system

The Dutch electricity grid is operated by two types of actors. The first one, of which only one exists, is the system operator (TenneT), who is responsible for maintaining the energy balance and for operating the transmission grid and the interconnectors with Belgium and Germany. The technical system is relatively simple, consisting of about 20 substations, 50 single line circuits and 10 interconnectors. However, the decision environment is very large and complex, as it involves producers, regional distribution companies and other European system operators. Therefore the system operator level is not an appropriate environment for developing an analytical approach to multi-issue decision making.

The other type of actor is the regional distribution company, of which about 30 exist[137]. The role of such a regional distribution company is essentially to transport the energy from the transmission grid to the consumers. The decision environment of these companies is relatively simple, as they are connected to a single system operator and they do not have the responsibility for managing the energy balance (nor the authority)[138]. On the other hand, the technical system is more interesting, as it consists (for one of the big three) of tens of thousands of substations and tens of thousands kilometers of cable in a multitude of network configurations. This type of actor has a much higher likelihood of encountering interdependent risks that can be decided upon in a single decision. This is the perspective chosen for this paper.

7.2.2.3 Interaction types

The last system boundary is in the interaction type. Although the choice for a distribution company perspective excludes multi-actor issues like ones in which the market design interacts with long-term reliability issues (de Vries, 2004), it still could include a multitude of interactions. These could range from interaction between organization design and risk exposure as described by (Heimann, 1997)[139], the influence of the ageing workforce on network performance or drill down towards the influence of failure of constituting components on the performance of the equipment. This paper is limited to technical system behavior, and does not consider organizational aspects.

7.2.3 Structuring risk

As the total number of risks under the authority of a risk manager may number in the thousands, it is clear that some sort of structure is needed to prioritize the risks that one would like to mitigate (Morgan et al., 2000). Two basic approaches exist to structure the risk. The first one is to work bottom-up, by identifying systemic risk from technical failures and transforming them into decision

---

[137] This was at the time of writing the paper. Enexis for example consisted of 5 independent DNOs. Currently (see chapter 2) only 8 DNOs remain. Stedin, Alliander and Enexis (current names) are the three largest of these 30 and cover about 90% of the market.

[138] Therefore a production plant is just a large customer which has to be served.

[139] With imperfect decision makers and a limited decision making system a tradeoff exists between failing to mitigate a risk and overspending on a risk. Putting decision makers in a series (hierarchical decision making) limits the probability of overspending (each decision maker can reject a mitigation), but increases the probability of failing to mitigate. With decision makers in parallel it is the other way around.

problems. The second one is to work top-down, starting with decision outcomes and identifying the risks that apparently drove them. In this paragraph both approaches will be discussed.

Bottom-up structuring of technical risks usually starts with identification of equipment, analyzing failure modes of the equipment and evaluating the effects of the failure. Interdependencies between failure modes can be modeled using a fault tree, an event tree(Wang and Roush, 2000) or a risk tree(Johnson, 1980). This process is the failure mode and effect (FMEA) analysis. Failure modes are defined by international standards like IEC 60812:1985 (IEC, 1985a). This method has a number of advantages. First of all, it is easy to make a list of all the equipment that one has. Furthermore, the analysis provides hints were to improve the system. However, although the behavior of the components is well understood, systems of components show emergent behavior. Large interconnected systems like electricity networks may fail if components that are a large distance apart fail simultaneously. Therefore, to understand system behavior in terms of component failure one needs to evaluate all possible combinations of failures. Even in relatively small systems this leads to a combinatorial explosion[140]. Normally one assumes that the probability of higher order failures decreases faster than the number of possibilities increases, but this is not necessarily the case(Carreras et al., 2003). Another concern is that this approach may neglect risks that are not entirely technical. An example is an external common cause like a large storm, such as the one that hit France in December 1999, which took out about 6% of the network, requiring weeks for full restoration(Le Du et al., 2002).

The other approach to structuring risk is to work top-down from decision outcomes to the risks that the outcomes were based upon. On an organizational scale, a decision outcome is usually a budget, not an individual project decision. The risks should therefore be viewed as drivers for the budget. In Table 17 some budget categories for a network operator are shown.

---

[140] For example, a system of 100 interconnected components has 100 first order failures, 9900 second-order failures, 970200 third-order failures and so on. Real systems consist of thousands of components.

Table 17: Budget categories and budget drivers

Outcome (budget)	Driver	Outcome (budget)	Driver
New connections	Number of new houses Number of new small businesses Number of large businesses	Fault restoration	Failure rate Failure type Number of assets Accessibility fault location
New infrastructure	Number of new areas Occupation rate Use of new areas	Protective measures	Third party damages Terrorism Vulnerability assets Failure rate
Reinforcements	Load growth Failure rate Critical level (neg) Failure acceptance	Control center	Number of assets Information availability (neg) Activity level
Replacements	Leakage Failure rate Risk level asset location Number of assets	Information availability	Public requests Third party activities near assets
Maintenance	Leakage Failure rate Risk level asset location Legal demands Number of assets		

Each of the drivers can be expanded into lower level risks. This approach has the advantage that the analysis is properly aligned with the decisions that it is supposed to inform. However, organizations are usually divided into departments, where each department has its own budget, corresponding with the table above. It seems reasonable that each department analyses the drivers for budget that they are responsible for. The table shows that some drivers appear in more than one budget (for example failure rate). If the departments work in isolation (and this appears to be the default behavior), drivers may be broken down into different constituting risks, or broken down into the same but differently named risks. This introduces an administrative interdependency besides the physical ones, which makes it almost impossible to determine total risk from the lower-level constituting risks.

The straightforward solution seems to be standardization of the risks at all levels, both in name and definition. However, this view presumes that risks can be defined objectively and that some natural taxonomy of risks exists. Unfortunately, this is most likely not the case. The proper categorization of risks depends on the decision situation and can therefore vary between organizations[141].

Nevertheless, the multitude of risks still needs to be condensed into a manageable number of risk areas for prioritization. (Morgan et al., 2000) argue that one should try a few different structures to find the one that best suits the decision problem at hand. The foundation which they offer for structuring the risks is a process, in which the risk is followed from the originating activity to the undesired effect upon values. It is possible to organize the risks according to steps in the process. Using the phases as the high level entries at least prevents the administrative overlap that may result from a less coordinated top-down approach. Figure 31 shows this analysis for air pollution risks.

---

[141] For example, IBM and Fortis both have the categories credit risk and market risk, but where IBM sees liquidity as a credit risk Fortis regards it as a market risk (Grey and Shi, 2001, Henrard, 2004)

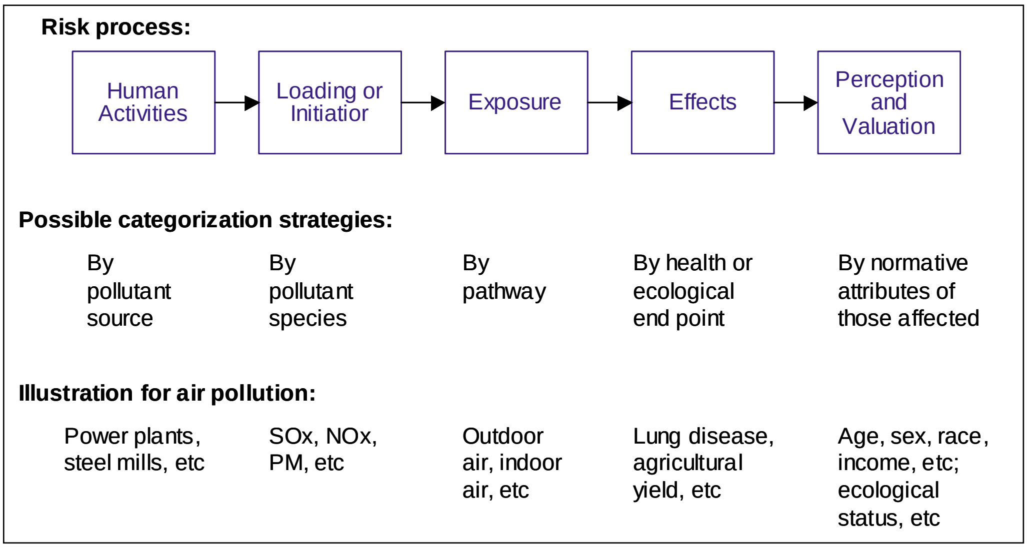


Figure 31: The use of a risk process to structure risk, after Morgan et al. (2000).

In this paper, the risk process concept will be used to model both the overview of the risks and the interdependencies between them.

7.3 Creating an overview

To group the electricity infrastructure risks into a manageable number, the risk process mentioned before will be used. Because only one human activity (transmission of electrical energy) is considered, the first step is omitted. Furthermore, the process steps are renamed in order to fit the risks under investigation. This is shown in Figure 32.

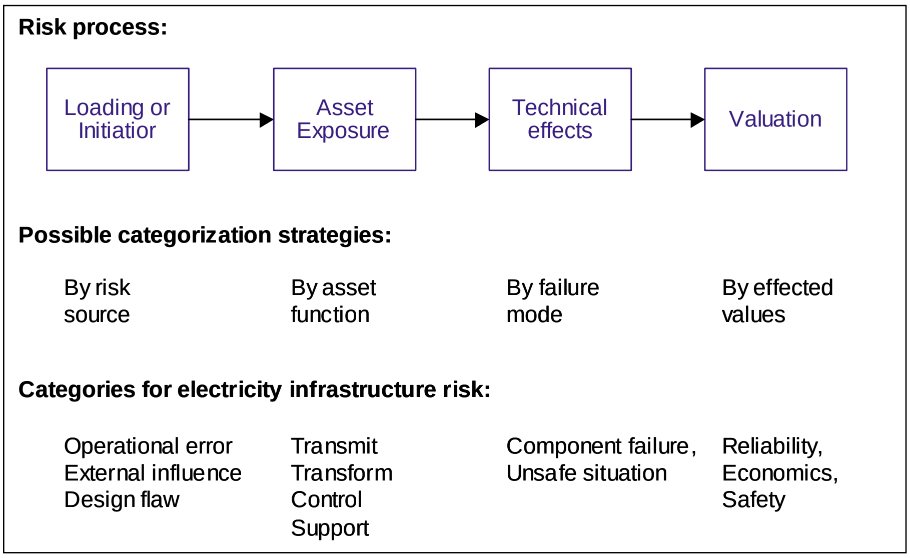


Figure 32: The risk process for an electricity distribution company.

In a risk capture workshop held at an electricity distribution company, several hundreds of issues were found by three different teams (gas, electricity low and medium voltage and electricity high voltage[142]). Those issues varied in different dimensions, of which a small impression is given in Table 18.

Table 18: Dimensions found in risk identification workshop.

Dimension	**Lower limit (example)**	**Upper limit (example)**
Size	100 € (what type of joint for specific customer)	10 M€ (new high voltage substation)
Complexity	Single criterion threshold (load level exceeded)	Multi criteria weighted sum (maintenance concept)
Scope	Technical (settings for protection device)	Ethical (can we continue operating overhead lines if they might increase the likelihood of leukemia)
Time horizon	Tomorrow (new customer application)	25 years (network design for distributed generation)

The issues were grouped into a manageable number, resulting in about 20 risks per team. Although each team worked in the same decision environment and used the same criteria[143] to classify the risks, the resulting top level entries of the risk register showed very little coherence, not only between teams but also within teams. All of the four risk process steps were used as top entry, therefore introducing significant administrative overlap. For example, if one risk is defined as transformer failure (asset exposure) and another one as a long interruption of supply (valuation), the instance that a transformer failure causes a long interruption would be included in both risks. Therefore, using the risk register to evaluate future performance would result in a prediction of two incidents whereas only one would be likely to occur. This over completeness of the risk register could result in overspending on some risks and therefore (due to budget limitations) in missed opportunities for other risks. Another problem caused by the lack of structure was that no simple check on completeness could be carried out, thus increasing the likelihood of neglected risks. Using the risk process as a framework could have provided a solution to both problems, as summing the risk within a step and comparing the result with the long term average performance would give an indication on over completeness (sum larger than average) or incompleteness (sum smaller than average) [144]. Framing the gathered issues afterwards into the risk process model, the items at the bottom of Figure 32 represent most of the items brought up in the workshop.

7.4 Experiences

7.4.1 Types of interdependency

In modeling interdependency between risks one should distinguish between different types of dependency. The risk process provides a tool to visualize these relations.

---

[142] The argument for dividing electricity risks into two groups is historical. However, low and medium voltage is carried by underground cables whereas high voltage is carried by overhead lines. The risks for high voltage therefore differ from those low and medium voltage.

[143] Enexis defined risk classification schemes for Economics, Quality of Supply, Safety, Legal, Environment, Customer and Regulator. Of these 7 values Economics, Quality of supply and Safety dominated.

[144] This is only possible if the categories within one step are mutually exclusive.

The first type of dependency is the *administrative* one, in which all steps of the risk process are used as a top level entry. An administrative interdependent risk register can neither be checked for over completeness nor incompleteness, as is shown in the previous section.

The second type is *causal* interdependency, meaning that a loading or initiator may cause an exposed asset to fail and result in an undesired effect that is perceived and valued by public and staff. The risk process of Figure 33 is described according to this kind of relation.

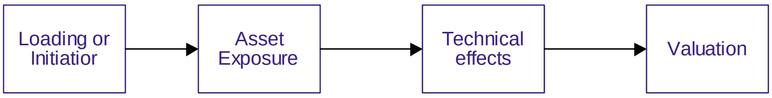


Figure 33: The risk process.

A third type is *parallel* interdependency. As the basic diagram does not allow for multiple processes, the diagram will be expanded to a graph (Newman, 2003) in which the categories will serve as nodes. The connections between the nodes can be regarded as (probabilistic) response functions. For example, external influences like excavation works may contact cables (transmit), with a high likelihood of failure, which may result in an interruption (reliability). This runs parallel to switching on (operation) a transformer on which people are working, therefore creating an unsafe situation which could result in fatalities.

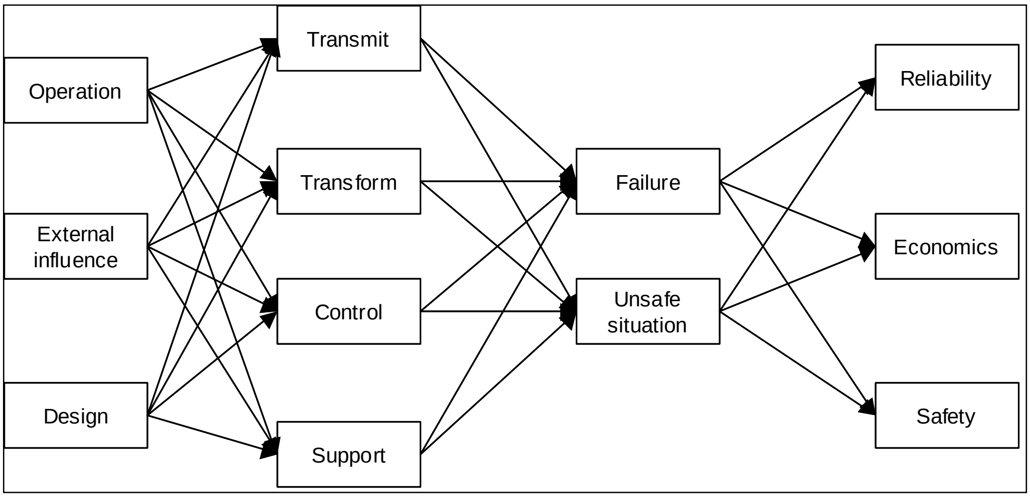


Figure 34: The risk process expanded into a graph.

In Figure 34 one can distinguish 72 different paths (3 times 4 times 2 times 3). These paths can be regarded as risk chains. Risk chains are dependent if they share at least one node. This dependence emerges in the intervention of the risk. In order to find the proper intervention for a risk one should investigate the chain. Continuing the excavation works example, one first could try to reduce the probability that the cable is contacted, for example by making the cable location data available to the excavation contractor. Another intervention might be to reinforce the cable to reduce the probability of failure when contacted. The last alternative might be to redesign the system to create active redundancy, in which case failure of the cable would not result in an interruption. However, any of these alternatives does more than just act upon excavation caused interruptions. For example, increased redundancy reduces the vulnerability of the system to many kinds of failures.

Another form of risk interaction can be observed in *sequential* dependency, which means that one risk is the initiator of another risk. For example, if an external event like lightning causes a transformer to explode, the scattered remains of the transformer may hit other components and cause them to fail. In the risk process this can be represented by a feedback loop from technical effects to external influences and operation, shown in Figure 35.

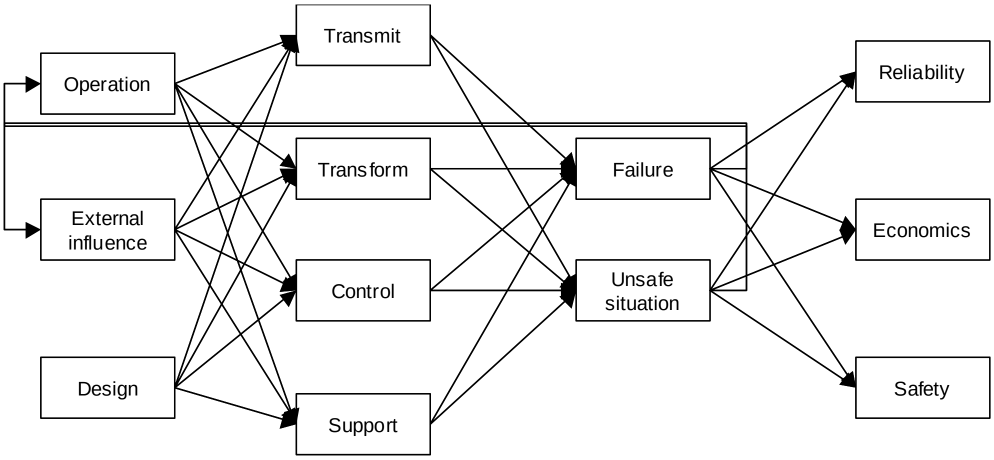


Figure 35: Sequential interdependency in the risk graph.

The last type of interdependency discussed in this paper is the *intervention-coupled* dependency. This means that the intervention for one risk acts as the initiation of another risk chain. The dependency between capacitators and tariff control signals, the example that was discussed in the introduction, is of this type. This kind of dependency can be displayed with a feedback loop from valuation to design, as shown in Figure 36.

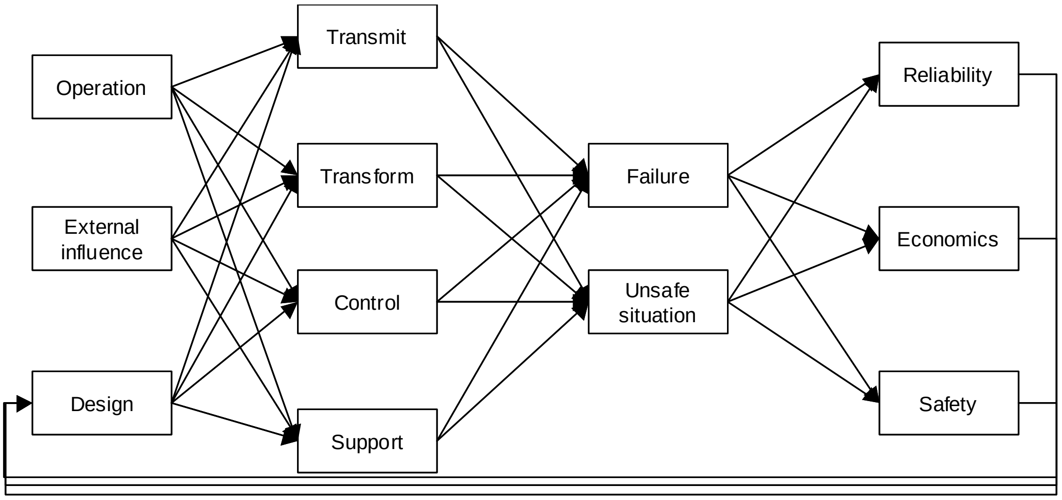


Figure 36: Intervention coupled interdependency in the risk graph.

7.4.2 Expanding the model to fit real networks

In the modeling of interdependency, the scope was restricted to twelve categories divided over four steps. Although in a conceptual sense this accommodated all risks, it was not enough for quantitative analyses of the interdependencies. For example, although a high voltage transformer is conceptually the same as a low voltage transformer, the risk chains containing those transformers normally do not interact. The same reasoning holds for other asset exposures, or for the other risk process steps.

Each node can be split in several other nodes. Although this expansion fits the real networks better, it has the downside that the number of links is proportional to the square of the number of elements[145] and the number of paths is proportional to the fourth power of the number of elements[146]. This means that the modeling effort increases much faster than the number of items.

7.5 Conclusions

In infrastructures, many risks exist that exhibit interdependency. These interdependencies can lead to unexpected decision outcomes. To understand this emergent behavior, an overview of all risks and their interdependencies is needed. As the types of risk number at least in the hundreds, grouping of risks is also needed. However, grouping risks into a manageable number can be treacherous, as an indicative check on over completeness or incompleteness is only possible if risk categories are mutually exclusive. In an unstructured approach it is highly likely that risks are not mutually exclusive. The risk realization process offers a conceptual tool for structuring the risks in order to be summed and tested for completeness. The risk process can also be used for modeling interdependencies between risks. This requires an expansion of the risk process into a graph. In this graph, different kinds of interdependencies can be expressed, ranging from simple causal relations to complex intervention coupled influences. Using the risk process may help decision makers to take the dependencies between the risks into account when deciding what strategy to follow.

7.6 Future work

This paper explored the theoretical possibilities in the grouping of risks and the modeling of interdependencies between the risks. However, as was pointed out in the paper, risk analysis is a decision driven activity. Therefore the only real test for the modeling framework is in its usefulness for decision making. The first step in this test is to use the framework in the risk identification process and see if its use generates more inputs. More inputs indicate a lower likelihood of incompleteness. The second test is in the number of risks that are decided upon in combination. A higher number of those complex decisions would indicate more ex-ante acknowledgement of interdependency. Finally, as the feedback loops create a theoretically infinite number of possible paths, a method to value those feedbacks needs to be developed. This requires further research.

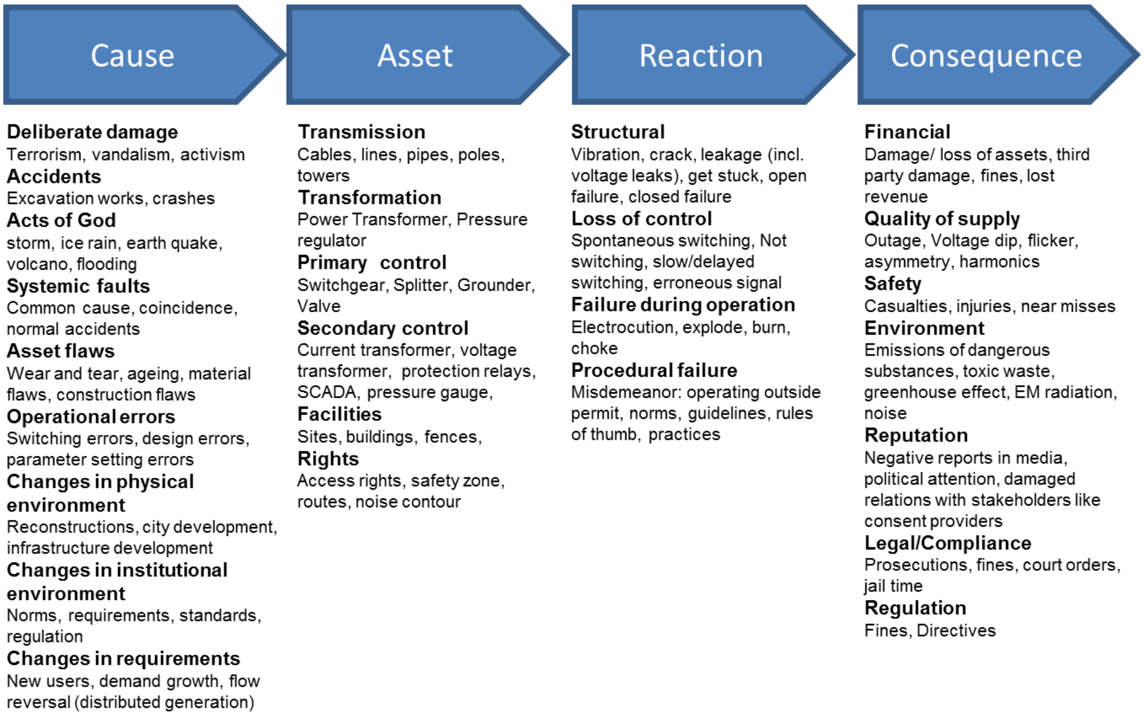
---

[145] If one represents a graph as a matrix with the nodes defining both rows and columns, with the value of the elements determining the strength of the link between the nodes, one can easily see that the number of links is the square of the number of elements.

[146] If one increases the number of nodes per step by a factor two, four steps will result in 2*2*2*2 times as many paths.

7.7 Epilogue

Over the years, in many publications (Masera et al., 2006, Wijnia and Hermkens, 2006, Wijnia et al., 2006, Wijnia and Nikolic, 2007, Wijnia, 2009, Wijnia, 2012, Herder and Wijnia, 2012, Wijnia, 2014) the risk process has been used as a perspective on risk. In virtually every use of the risk process, elements have been added. The current version of the risk overview is shown in Figure 37.



Deliberate damage
Terrorism, vandalism, activism
Accidents
Excavation works, crashes
Acts of God
storm, ice rain, earth quake, volcano, flooding
Systemic faults
Common cause, coincidence, normal accidents
Asset flaws
Wear and tear, ageing, material flaws, construction flaws
Operational errors
Switching errors, design errors, parameter setting errors
Changes in physical environment
Reconstructions, city development, infrastructure development
Changes in institutional environment
Norms, requirements, standards, regulation
Changes in requirements
New users, demand growth, flow reversal (distributed generation)

Transmission
Cables, lines, pipes, poles, towers
Transformation
Power Transformer, Pressure regulator
Primary control
Switchgear, Splitter, Grounder, Valve
Secondary control
Current transformer, voltage transformer, protection relays, SCADA, pressure gauge,
Facilities
Sites, buildings, fences,
Rights
Access rights, safety zone, routes, noise contour

Structural
Vibration, crack, leakage (incl. voltage leaks), get stuck, open failure, closed failure
Loss of control
Spontaneous switching, Not switching, slow/delayed switching, erroneous signal
Failure during operation
Electrocution, explode, burn, choke
Procedural failure
Misdemeanor: operating outside permit, norms, guidelines, rules of thumb, practices

Financial
Damage/ loss of assets, third party damage, fines, lost revenue
Quality of supply
Outage, Voltage dip, flicker, asymmetry, harmonics
Safety
Casualties, injuries, near misses
Environment
Emissions of dangerous substances, toxic waste, greenhouse effect, EM radiation, noise
Reputation
Negative reports in media, political attention, damaged relations with stakeholders like consent providers
Legal/Compliance
Prosecutions, fines, court orders, jail time
Regulation
Fines, Directives

Figure 37: The overview of risk factors for energy distribution assets, a useful tool in quickly identifying relevant risks (= path from cause to asset).

Compared to the original as presented in this chapter, the most significant change was the addition of a deeper level of decomposition, for every element of the diagram. Furthermore, many more types of causes, reactions and consequences are included. A special element, stepping outside the realm of engineering assets, was the asset class “rights”. The diagram above holds the latest version of the risk factors organized along the risk process for the assets in the distribution infrastructure for electricity and gas. This overview of risk factors helps in the identification of risks for organizations that do not yet have a risk register, but in an adapted form it can also help in risk structuring for other asset bases like an IT infrastructure (Wijnia and Nikolic, 2007).

8 Measuring risk

This chapter was published before as **THE SAFETY INDICATOR: MEASURING SAFETY IN GAS DISTRIBUTION NETWORKS (WIJNIA ET AL., 2010)**[147] which is an update of **MEASURING SAFETY IN GAS DISTRIBUTION NETWORKS (WIJNIA AND HERMKENS, 2006)**[148]

8.1 Introduction

The gas distribution industry in the Netherlands claims that the system is safe. To support this claim, they refer to countless numbers of NEN/EN/ISO standards, which are applied in the design and maintenance of the system. Phrased differently, the industry states that gas distribution is highly technical, that experts have thought about the details thoroughly, and that those experts should be trusted. As both the standards and the experts are from within the industry, the basic message is: "You can trust us, we are the experts, we know and we care", a message that worked well for a long time.

In the current society, however, this asked for trust does not come naturally, as Shell, for example, experienced with the Brent Spar. In short, the Brent Spar was an oil storage rig in the North Sea that had to be decommissioned because its function had been replaced by pipelines. A Best Practicable Environmental Option study showed that the best way forward was a deep sea disposal. Greenpeace, however, did not agree with this study. They stated the volumes of toxic materials were much higher than Shell claimed and occupied the rig to prevent the disposal. The public assumed Greenpeace was more trustworthy than Shell and boycotted the Shell gas stations, an act which forced Shell to change course. In the aftermath, Greenpeace had to admit that they were wrong and Shell was right about the toxic content of the Brent Spar, but that did not help Shell anymore. The option of deep sea disposal was definitively out of question.

In the debate on the safety of the gas distribution system, some parallels can be drawn to this Brent Spar example. First of all, the liberalization of the Dutch energy markets in the period 1999-2004 resulted in lots of faulty invoices from the energy companies, destroying the image of a trustworthy industry. The large 3 energy companies ranked high in customer dissatisfaction (top3), according to consumer watchdog television programs like Kassa and Radar. Secondly, because of this bad press, every gas related incident was connected to the liberalization, even if it had nothing to do with the gas distribution system. An example can be found in the explosion in The Hague (June 28, 2003), which according to the first reports was caused by gas leakage from a gas pipeline, where in fact it was caused by a leaking propane tank in the cellar. As a bonus, the public awareness of incidents was further enhanced by the rise of numerous reality TV shows and local human interest programs on the increasing number of nationwide television stations. The odds of an incident drawing attention of one of such shows were further enhanced by new technologies like the internet and cellular phones,

---

[147] Wijnia, Y. C., Hermkens, R. J. M. & Flonk, J. 2010. The Safety Indicator: Measuring Safety in Gas Distribution Networks. *In:* Amadi-Echendu, J. E., Brown, K., Willett, R. & Mathew, J. (eds.) *Definitions, Concepts and Scope of Engineering Asset Management.* Springer London.

[148] Wijnia, Y. C. & Hermkens, R. J. M. Measuring safety in gas distribution networks. Proceedings of the 1st World Congress on Engineering Asset Management, WCEAM 2006, 2006. 1070-1079.

speeding up the propagation of news. Furthermore, the Dutch safety board has published a few reports with some very alarming conclusions. As a result it seems as if the number of incidents is increasing.

In this light, it is not very surprising that the public did not take the industries claim of safety for granted and demanded additional measures. The awareness of incidents increased, and awareness drives the perception by the availability bias (Kahneman and Tversky, 1979). But is there any truth in the perception? Is the gas grid really becoming unsafe? Embarrassingly, to this question no objective answer existed. Beside the statistics on leakage and the standards, the industry did not have any facts supporting the claim of safety. On the other hand, there was no evidence for an increase in the number of incidents either, but the industry did have to prove its claim as perception was the public truth.

To issues revolving about the public perception, the best approach is not to claim expertise and declare safety, as the industry was used to (and Shell did initially with the Brent Spar), but to share knowledge and data (Shells reprise). This created an interesting challenge for the gas industry, as they did not have very much knowledge and data to share. First of all, they did not have an easy accessible database on all incidents and accidents related to the gas network. Secondly, they did not have any methods to translate such a list of incidents into an easy to understand measure of safety.

Fortunately, some experience on quality issues existed. Most gas distribution companies in the Netherlands also distribute electricity, and for the electricity grid, the major quality indicator is Customer Minutes Lost (CML), which had been recorded for some 25 years in the so called “Nestor Enquete”. If such an instrument could be developed for the safety of the gas grid, it might prove helpful in the public debate. However, discussions on safety issues are always difficult, as extensive research on this subject has shown (Slovic and Weber, 2002). Furthermore, even the existence of an objective metric for safety does not guarantee the right perception, as the reliability indicator for the electricity grid shows. The perception is that electricity is interrupted once every year for about 2 hours, where in fact it is only once every 4 years according to the 25 year statistics (Baarsma et al., 2004).

Despite these potential barriers for acceptance, just knowing for sure was tempting enough for the industry. All authors were asked to take part in the development of the Safety Indicator. In this paper, the results of this development process will be presented. First the issues surrounding the concept of safety will be addressed, like actual figures versus disaster potential, adding different types of accidents, the value of a human life and so on. This is the value judgment section. The next step is building a conceptual model for the indicator which translates the measured inputs into a single metric. As will be demonstrated, a direct assessment of the safety consequences is not a reliable indicator for the safety of the gas distribution system. A deeper understanding of the underlying mechanism is needed. This is the conceptual part. This conceptual model will be applied to actual incident data, resulting in a quantitative estimation of the safety risk in the gas distribution system. The next part reviews the implementation of the Safety Indicator in the Netherlands. The paper with conclusions on the usefulness of the indicator.

8.2 Theoretical framework

In this section the issues surrounding the debate on safety will be addressed. As will be demonstrated, much of the issues require some kind of value judgement on a certain aspect of the total safety abstraction. To structure the discussion, first focus will be on the different concepts of safety. Based on the characteristics of the safety risk and the uncertainty involved, a concept will be chosen for the indicator. Next, the issue how to count safety incidents will be addressed. Safety incidents can have quite different consequences, like personal injuries or fatalities, but also property damage. Adding those consequences is not trivial. This non triviality holds within a consequence level (is every fatality equal), between different consequence levels (how much worse is a fatality than a serious injury) and between different values (how should a fatality be compared with financial losses, or Customer Minutes Lost). The section ends with a valuation scheme for the different types of consequences.

8.2.1 The concept of safety

Safety can be used in a myriad of ways. Sometimes it refers to financials (this is a safe investment), sometimes to information (how safe is the internet), but in many cases it refers to personal injuries (road safety figures). Furthermore, safety can be about the actual figures (the annual fatalities on the road) or about the expected figures, the so called safety risk (e.g. nuclear power). Discussions on actual figures revolve around the definitions. For example, does the road safety figure only count the victims at the crash site, or are the ones in hospital afterwards included? Until how many days after the accident should those be counted? Those issues are solved relatively easy.

Discussions on safety risks are much more difficult. This is because different views exist about the essence of risk. In the first conceptual approach, risk is seen as an entity that can be objectively measured or calculated. In this view, risks are typically expressed as a product of consequence and likelihood. This objective risk is used by financial managers, safety engineers, decision analysts and so on.

The second approach challenges the assumption of objectivity, as it is not this objective risk that determines if a risk is acceptable. Some people smoke (high objective health risk) but worry in the same time about electromagnetic radiation from high voltage transmission lines (health risk uncertain). Research has shown that this acceptance level is driven by so called psychometric factors which are shown in Figure 38 (after Fischhoff et al. (1987)) and not only by the objective risk level. Risk is therefore very subjective.

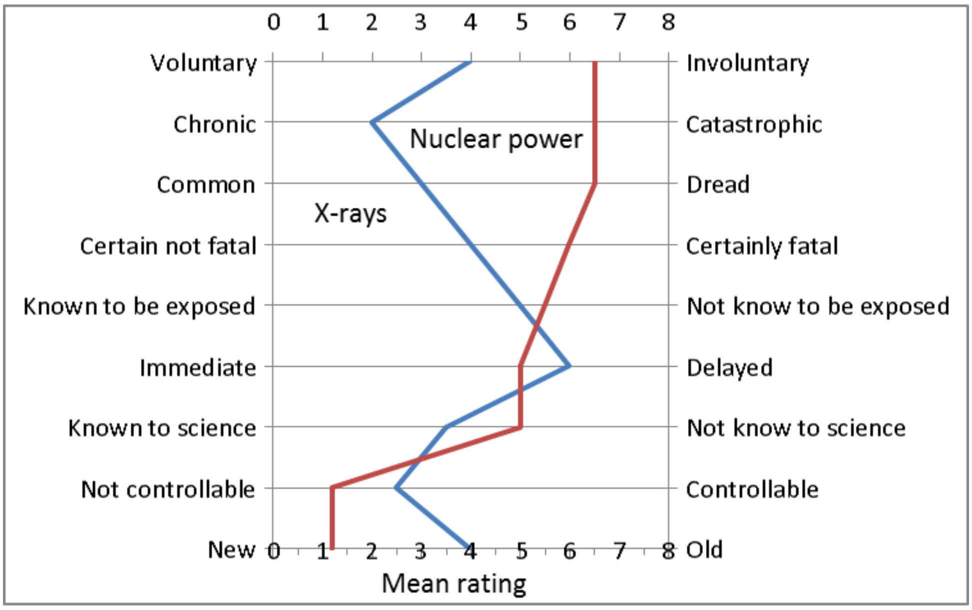


Figure 38: Psychometric factors for risk perception, after Fischhoff (1987).

Despite the usefulness of this subjective risk concept in risk communication, it does pose some problems for risk managers. For example, if the acceptance of a risk is determined solely by its disaster potential (like it was the case for nuclear power), no amount of risk measures can make it acceptable, as the disaster will always be thinkable. However, stopping all activities with a disaster potential might reduce the standard of living by a considerable amount, a sacrifice that few citizens are willing to make.

The constructive approach tries to bridge the gap between the objective and the subjective concepts, by stating that a risk is a social construction that human beings have “invented” to deal with the dangers and uncertainties of life. The dangers and uncertainties are real, as is the response to the risk (Thomas Theorem), but the risk itself does not need to be real. This approach is advocated by the USA’s National Research Council (Stern et al., 1996). A risk can therefore be defined as *a potential problem that needs to be decided about*. This view is further supported by the value judgements needed to arrive at a risk figure (as mentioned in the introduction to this section).

Which of those concepts should be used for the Safety Indicator? For answering this question the classification scheme of Klinke and Renn (2002) could be used. They propose 4 different kinds of decision approaches for different risks, being rule based, risk based, discourse based and the precautionary principle. The drivers in this classification scheme are uncertainty, disaster potential and the social mobilization factor. As the uncertainty is pretty low (over 50 years of experience), the key issue is whether or not the gas distribution system has disaster potential. Natural gas can cause large explosions (e.g. Piper Alpha) but this is only possible at pressures well above the ones used in the gas distribution system (EGIG, 2008). Therefore, a risk based approach could be used, even though some value judgement is needed. To manage the potential for social mobilization it is best to make it a collaborative effort by the industry and the regulating bodies.

8.2.2 Comparing incident types

In order to be able to add different types of accidents into a single number, some kind of a valuation scheme is needed. This valuation scheme should answer three questions: Which are the affected values that will be taken into account? What levels of severity will be distinguished? What are the equivalent values for the different levels of severity?

Those questions are not unique for a Safety Indicator. On the contrary, most companies familiar with risk management use a valuation scheme in their risk matrices. What is needed would be a risk matrix without a probability axis, being a consequence scale (terminology based on ISO 31010 (2009b)). In case of the Safety Indicator, the only question left to answer would be what affected values should be taken into account. The use of personal safety would be trivial, but was any other value needed? The obvious candidate would be property damage, a well-known consequence of gas explosion. Property damage could be assessed objectively. Therefore its inclusion would not threaten the objectivity of the indicator. In Table 19 the severity of different consequences for those values is shown.

Table 19: Severity levels of personal injuries and damages.

Severity level	Personal injury	Damage
6	Multiple fatalities	> 10 million euro
5	Fatality	1-10 million euro
4	Serious injury	100k-1M euro
3	Lost time incident	10k-100k euro
2	Near miss/first aid	1-10 k euro
1	Unsafe situation	< 1k euro

In this matrix the logarithmic character of the severity classes is very clear for the value “property damage”. Quite surprisingly, this logarithmic character also holds for the personal injuries. Research shows (Heinrich, 1931, Whiting, 2001, Saldaña et al., 2003, Perrow, 1984, Körvers, 2004, Nichols, 1973, Clark, 2004) that only a fraction of the near misses develops into a fatal accident. Numbers range from 300 near misses per fatality to 1000 per fatality. Even though summing injuries is quite different from adding damage (financial damage can be transferred, personal injuries cannot), if used with the mentioned scaling factors it makes sense statistically. It allows us to express any combination of safety incidents in the equivalent fatalities. The scaling factor between the values seems to be an industry standard (Shell Global Solutions, 2002). By expressing even the injuries in the monetary equivalent, all incidents can be expressed on a single scale, although this does in no way mean that a human life can be replaced by a certain amount of money. After all, no markets exist in which people sell their own lives.

8.3 Direct assessment of the safety incidents

Now that safety has been defined, counting can start. For this the Kiwa/Gastec accident database will be used, which holds records of all network related accidents (Category I incidents) in the Netherlands dating back until 1993. In addition the reportable incidents (Category II incidents) of the Dutch safety board will be used, although they only kept records for 2004 and 2005. A Category I incident is an event in which someone got hurt or third parties property was severely damaged (> € 500.000) as a direct result of the event, whereas an Category II incident is an event in which this might have happened, or the property damages are below € 500.000. All Category II incidents require the activation of the emergency services, for example for evacuation. The total number of Category I incidents was 39 over the period 1993-2003, which averages 3-4 accidents per year (van Akkeren and Wijnia, 2005). However, the variance is very large, as shown in Figure 39. The number of accidents varies between 2 and 5, but the equivalent value ranges from almost zero in 1996 to over 10 million in 1999. There is virtually no correlation between the number of incidents and the total value of the incidents, the calculated correlation is only 0,25. Besides, it is highly unlikely that the outcome is correlated to the quality of the grid. That does not vary this fast. It means a metric based on a direct assessment of the accidents does not produce a robust outcome if it is to be used for investment decisions.

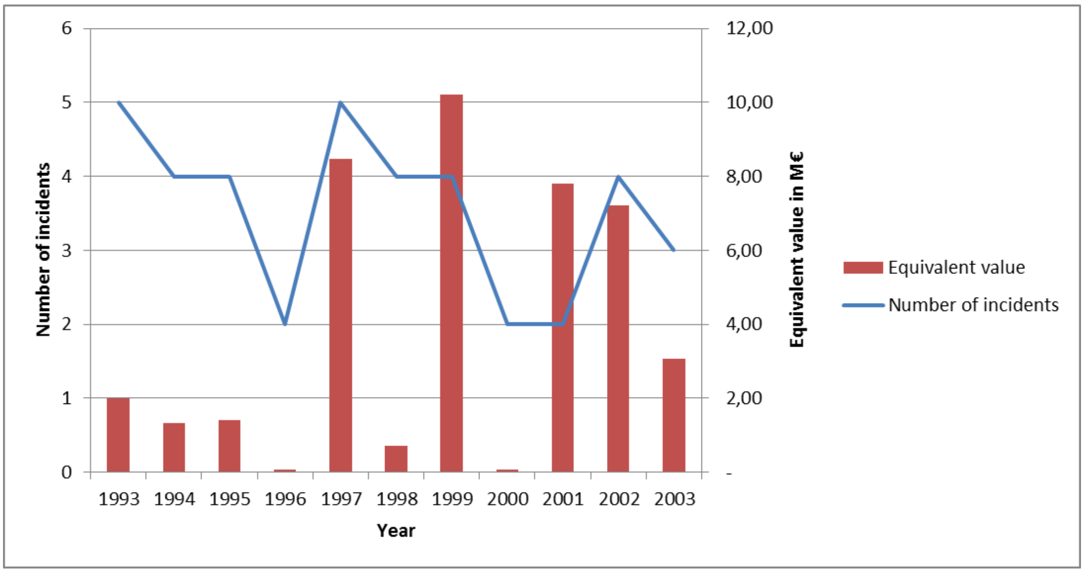


Figure 39: Gas related accidents in the Netherlands in the period 1993-2003, number and equivalent value.

This lack of robustness is a result of the low number of accidents, on average about 3 per year. As the indicator is meant to help in investment decisions, lack of robustness would be a fatal flaw. To overcome this problem, the data used to calculate the safety has to be extended. For this the incident process (Wijnia and Herder, 2004, Körvers, 2004, Morgan et al., 2000) as shown in Figure 40 will be used to get a better understanding what those data should be.

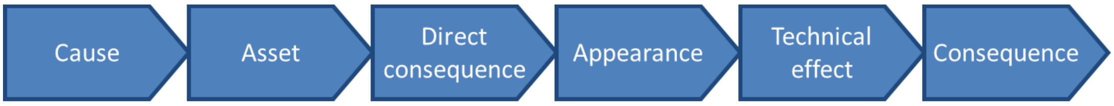


Figure 40: Incident process.

The incident process states that value consequences do not appear out of nothing, but are a result of a chain of cause and effect. For example, excavation works (cause) can damage a service line (asset), resulting in leakage (direct consequence). The escaping gas might accumulate (appearance) in a closed space. If the gas explodes (technical effect), it might destroy property and injure people (value consequence). Barriers might exist between the different phases in the incident process. Not all excavation works damage pipelines, not all damaged pipelines leak and so on. This means that if the processes that led to the mentioned 39 accidents can be reconstructed, safety can be measured based on situations occurring earlier in the process, the so called near misses. As the number of those situations is much higher than that of real accidents, it is more likely to produce a statistically robust figure.

This incident process based approach also matches the concept of safety as a risk, as it addresses not only the real accidents, but also the potential accidents. Applying the incident process to the recorded accidents led to the conclusion that only a few combinations of causes and assets (the so called precursors) of the thousands possible ever led to a safety accident. The causes and assets are shown in the incident triangle of Figure 41.

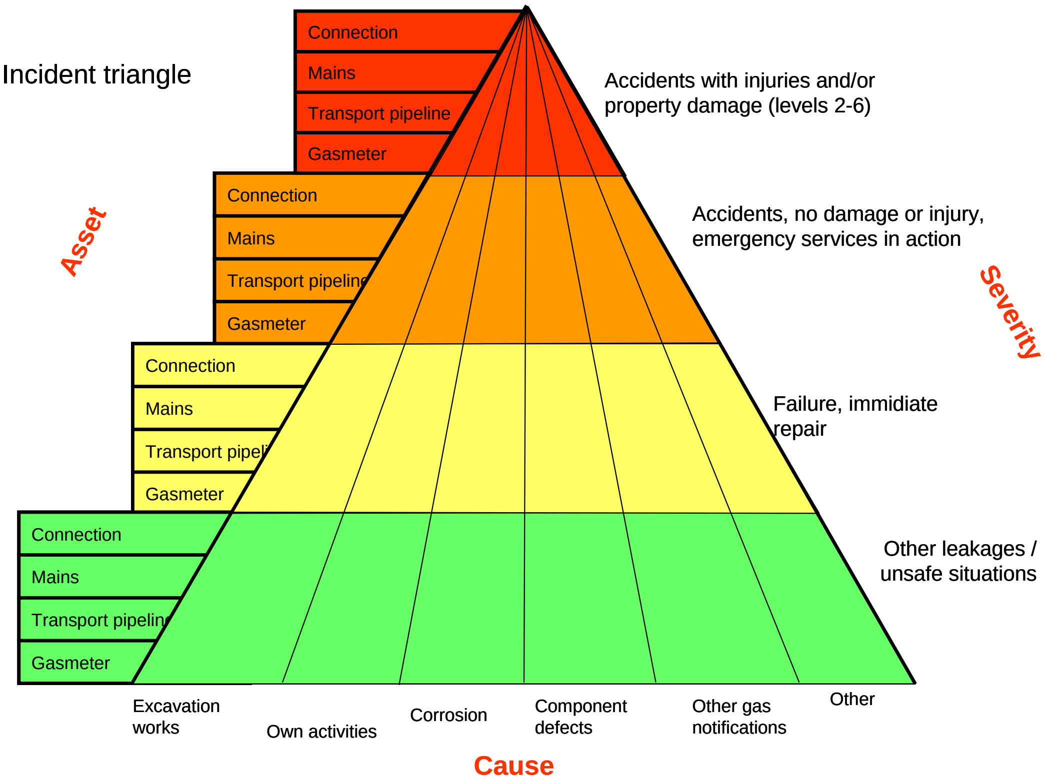


Figure 41: Incident triangle.

A direct link exists between the incident process and the triangle. For each combination of asset and cause, four levels exist on which an incident can end, corresponding to the phases 3 to 6 of the incident process. The top level accidents (red, Category I) are those incidents that completed the incident process and had a significant value consequence. The second level accidents (orange, Category II) represents the incidents that showed technical effects, but did not create heavy value losses. The third level incidents (yellow) are the ones that did have a public appearance, but no immediate dangers. Finally, the bottom level (green) represents those incidents in which a cause had a direct effect on an asset. Situations in which the asset was not damaged by the cause are not considered to be incidents.

In the incident process, a limited likelihood of propagation exists. Only a fraction of the assets will be exposed, only a fraction of causes acting on assets will create damage, only a fraction of the damage will reveal itself by escaping gas, only a fraction of the gas clouds will be within the explosion limits, and finally, not all dangerous situations will actually damage property or hurt people. Therefore, the number of incidents in the green level is much higher than the red level. This is in line with the iceberg model as described by Heinrich (1931). An indicator based on green level incidents should in theory be more reliable than an indicator that measures top level accidents. In practice, however, this does not hold, as the reliability of the indicator also depends on the quality of the incident registration, which is far better at the top level than at the bottom level. Fortunately, in the Netherlands there is Nestor-Gas, a database which holds records of incidents that required immediate action (yellow). This theoretically satisfies both the need for a robust number of incidents as the need for a reliable data source, and therefore will be used as the basis for the indicator.

With the framework presented above it is possible to calculate the total risk of a group of assets. However, absolute figures often do not mean much, as it does not tell about the size of the population from which the figures were derived. For example, road safety is always expressed as the number of fatalities per 1.000 inhabitants, or per million miles. This makes it possible to compare different years within a country, or to compare different countries. Which metric is the correct one, however, depends on the question to be answered. The inhabitant based metric tells about the likelihood of an average person to die from a road accident. If the question is about the safety of road design, it is better to use the mileage based figure. For the Safety Indicator, both an inhabitant related metric and a network length related metric can be used. Because the issue is public safety, it seems logical to use the inhabitant based figure. So, to calculate the Safety Indicator for a group of assets, the total risk for that group will be divided by the number of connections in the asset group.

This concludes the theoretical framework of the Safety Indicator. The concept of the Safety Indicator has been defined and the methodology for adding accidents into a single number has been specified. In short, the Safety Indicator is a metric that is proportional to the likelihood of an individual to die from an accident in the gas distribution network. In the next section this will be quantified.

8.4 The original quantitative model

In the previous section the concept of the Safety Indicator has been established, in this section the concept will be applied on the incident data. First the list of cause-asset combinations (the so called precursors) that led to an accident in the period 1993-2003 will be presented. Next, the calculation of the expected effect for a single precursor will be demonstrated. The section ends with the calculation of the expected average effects of all relevant incidents, that is, incidents that resulted in an accident in the period 1993-2003.

8.4.1 Relevant precursors

In mapping the accidents on the incident process it was found that only a few combinations of assets and causes ever resulted in an accident. These combinations are presented in Table 20.

Table 20: Accidents per precursor for the period 1993-2004.

ID	Asset	Cause	Personal injury	Damage (=total number)
1	Service line	Excavation	4	47
2	Service line	Own activities	2	3
3	Service line	Soil settlement	1	1
4	Service line	Component defects	2	4
5	Service line	Corrosion	2	9
6	Service line	Other leakage	5	1
7	Mains LP[1]	Other	0	25
8	Mains LP[1]	Component defects	5	6
9	Mains LP[1]	Own activities	18	23
10	Mains LP[1]	Excavation	4	164
11	Mains LP[1]	Soil settlement	0	41
12	Mains LP[1]	Other	2	4
13	Mains LP[1]	Other leakage	1	12
14	Mains HP[1]	Component defects	0	1
15	Mains HP[1]	Corrosion	0	8
16	Mains HP[1]	Excavation	2	68
17	Mains HP[1]	Soil settlement	0	9
18	Gas Meter Installation	Component defects	4	2
19	Gas Meter Installation	Own activities	0	8
		Totals	**52**	**436**

[1] LP= Low pressure (30-100 mBar), HP= High Pressure (1-8 Bar)

Note that the total number of accidents is higher than presented in the previous section. This is because the period of registration was longer. There were 11 years of only accidents (used in previous section), and 1½ years of data on accidents and serious incidents. The serious incidents were extrapolated to a 12½ year period to match the accident data.

8.4.2 Calculating the average incident value

The method for calculating the average expected incident data will be demonstrated for the LP mains-excavation precursor in Table 21. This is the combination with the highest number of reported incidents.

Table 21: Average incident value for excavation and LP mains precursor.

	Severity levels						
	6	5	4	3	2	1	**total**
Number of incidents	0	0	2	44	94	28	**164**
Value per incident	100.000.000	10.000.000	1.000.000	100.000	10.000	1.000	
Value per level	0	0	2.000.000	4.400.000	940.000	28.000	**7.368.000**
Average value							**44.927**

In the calculation the first step is to determine the number of reportable incidents per severity class. Then, this number of incidents per class is multiplied with the value of an incident in that class and summed over the levels, to get the total value of this precursor. This total is divided by the number of reportable incidents to get the average value per reportable incident.

However, this is not the total number of incidents on this precursor, as most of the incidents do not qualify as reportable. They are near misses. Unfortunately, for the near misses, the available data spans an even shorter period. The near misses are recorded in Nestor-Gas from Kiwa/Gastec Technology, but that only started in 2003. The near misses incidents relating to the Mains LP – excavation precursor (some filtering within the database) were extrapolated to match a 12½ year period. For this, it was assumed that the quality of the grid and thus the number of incidents did not change significantly. This may sound as a bold assumption, but the involved experts all worked within the business for more than 12½ years and felt it was a reasonable assumption. By dividing the number of reportable incidents by the number of near misses, the probability that the near miss turns into a reportable incident can be calculated. The results are shown in Table 22.

Table 22: Accident probability for incidents.

			Near Misses		**Reportable incidents**	**Probability**
ID	Asset	Cause	2004	1993-2004	12½ year period	
10	Mains LP	Excavation	781	9.763	164	0,0168

The last step was to multiply the average value of a reportable incident with the probability that a near miss turns into a reportable incident. This is the risk per near miss, shown in Table 23.

Table 23: Average risk per near miss.

ID	**Asset**	**Cause**	**Incident value**	**Incident probability**	**Near miss risk**
10	Mains LP	Excavation	44.927	0,0168	755

8.4.3 Results for all combinations

Table 24 shows the results for all precursors.

Table 24: Risk figures for all precursors.

ID	Asset	Cause	Incident value	Incident probability	Near Miss Risk
1	Service line	Excavation	172.660	0,00158	272
2	Service line	Own activities	1.066.667	0,00049	527
3	Service line	Soil settlement	1.010.000	0,00006	62
4	Service line	Component defects	3.052.500	0,00178	5.427
5	Service line	Corrosion	1.334.222	0,00104	1.394
6	Service line	Other leakage	60.000	0,00008	5
7	Mains LP	Other	71.200	0,00741	527
8	Mains LP	Component defects	2.538.333	0,00828	21.007
9	Mains LP	Own activities	701.435	0,01172	8.221
10	Mains LP	Excavation	44927	0,01680	755
11	Mains LP	Soil settlement	8.244	0,00576	48
12	Mains LP	Other	282.750	0,00041	116
13	Mains LP	Other leakage	102.333	0,00173	177
14	Mains HP	Component defects	1.000	0,01600	16
15	Mains HP	Corrosion	1.000	0,02286	23
16	Mains HP	Excavation works	51.721	0,15543	8.039
17	Mains HP	Soil settlement	100.000	0,01469	1.469
18	Gas Meter Installation	Component defects	2.050.500	0,00001	23
19	Gas Meter Installation	Own activities	100.000	0,00137	137

Multiplying the number of incidents with the associated risk of the incidents and summing over all incident types gives the total safety risk. This is shown in Table 25.

Table 25: Total risk, connection risk and individual risk.

ID	Asset	Cause	Near Miss Risk	# Near Misses	Total risk
1	Service line	Excavation	272	2.384	649.200
2	Service line	Own activities	527	486	256.000
3	Service line	Soil settlement	62	1.313	80.800
4	Service line	Component defects	5.427	180	976.800
5	Service line	Corrosion	1.394	689	960.640
6	Service line	Other leakage	5	1.010	4.800
7	Mains LP	Other	527	270	142.400
8	Mains LP	Component defects	21.007	58	1.218.400
9	Mains LP	Own activities	8.221	157	1.290.640
10	Mains LP	Excavation	755	781	589.440
11	Mains LP	Soil settlement	48	569	27.040
12	Mains LP	Other	116	780	90.480
13	Mains LP	Other leakage	177	555	98.240
14	Mains HP	Component defects	16	5	80
15	Mains HP	Corrosion	23	28	640
16	Mains HP	Excavation	8.039	35	281.360
17	Mains HP	Soil settlement	1.469	49	72.000
18	Gas Meter Installation	Component defects	23	14.077	328.080
19	Gas Meter Installation	Own activities	137	467	64.000
	Total				7.131.040
	Per connection	**7.031.000**			**1,014**
	Per inhabitant	**16.000.000**			**0,445**

The risk per connection is about 1 euro per year, or 45 eurocents per inhabitant. In terms of equivalent fatality risk this is about once every 20 million years per inhabitant. Equivalent risk means that any financial damage is translated into a fatality risk. In the 12½ years only one fatality occurred, so the actual fatality risk was about once every 200 million years per inhabitant. Both figures are well below the once per million year limit for individual fatality risks.

8.5 Implementing the Safety Indicator

In the quantitative model significant differences occurred between types of incidents, that were recognized by the practitioners. The Safety Indicator thus had enough appeal for the network operators to adapt it. However, as the dataset was limited in size, the numbers could be wrong. The network operators thus agreed upon a trial period for the Safety Indicator in which it could be further tested and refined (Hermkens, 2005). They agreed to use 2006 as the trial period, after which the Safety Indicator would be updated. In this section, the major differences with the original model will be presented, based on **STATUUT VEILIGHEIDSINDICATOR** (Hermkens and Pulles, 2007). This document describes the method by which the Safety Indicator score should be calculated, supplemented with the rules on how to add or alter parts.

8.5.1 The value system

In the discussion on the Safety Indicator, the network operators mentioned that the social disruption (evacuations due to gas cloud) were not properly valued, as only the financial consequences would be counted. The metric proposed to measure this was evacuation hours. For an incident, this is the number of people evacuated times the duration of the evacuation in hours. The metric is not perfect, as it does not capture all potential consequences (like closing down roads), but it seems to be a good proxy. Furthermore, network operators were not very comfortable in expressing everything in euros. Therefore, consequences were translated into a dimensionless norm value. Table 26 shows the new severity levels.

Table 26: Severity levels of personal injuries and damages revisited.

Severity level	Personal injury	Property damage	Social disturbance	Norm value
6	Multiple fatalities	> 10 million euro	>100.000	1.000.000
5	Fatality	1-10 million euro	10.000-100.000	100.000
4	Serious injury	100k-1M euro	1.000-10.000	10.000
3	Lost time incident	10k-100k euro	100-1.000	1.000
2	Near miss/first aid	1-10 k euro	10-100	100
1	Unsafe situation	< 1k euro	<10	10

However, even though the Safety Indicator uses a dimensionless number, in this paper the monetary equivalent will be used, to keep results comparable to the 2006 values. There is just a scaling factor between them, which is 60 euros per norm point.

8.5.2 Normalization

In the original model, all risk was normalized to the number of inhabitants, based on the rationale that the Safety Indicator (SI) should express external risk. However, this was not very helpful in targeting investments, as it was not related to asset quality for all assets, specifically the gas mains. Network operators working in a low density area would score worse (= higher SI score) with the same grid quality, as they have more mains per inhabitant. That did not seem right, as a gas leakage in open field is virtually risk free, and gas leakage in cities requires evacuation. To correct this flow, mains related precursors would be normalized to the length of mains, whereas service line precursors, gas meter precursors and so on would still be normalized to the number of connections, as described in the formula below.

$$SI_b = \sum_{CRi} R_i \frac{M_{b,i}}{FA_b} + \sum_{MRi} R_i \frac{M_{b,i}}{FH_b} * AML \qquad (8.1)$$

Where:

SI_b: Safety Indicator score for network operator b
CRi: Connection related precursors
Ri: Risk per precursor i
$M_{b,i}$: Annual number of incidents at network operator b for precursor i
FA_b: Number of connections at network operator b
MRi: Mains related precursors
FH_b: Length of mains at network operator b,
AML: Average length of mains per customer in the Netherlands

If all incidents in the Netherlands are reviewed, the number will be completely inhabitant based.

8.5.3 Precursors

The original model held only 19 relevant precursors that produced a real accident in the past period. Due to the expansion of the value system, the network operators felt more precursors were needed.

This new list included two new causes (molestation and historical construction errors), a new asset (stations), and new combinations of existing assets and causes. Table 27 shows the new precursors, their normalization, and the annual risk. At the bottom of the table, the annual risk in the new version and the new precursors is compared to the original model of 2006.

Table 27: New precursors in the 2007 model and their annual risk.

ID	Cause	Asset	Normalization	Annual CAT I risk [€]	Annual CAT II risk [€]	Total Annual Risk [€]
4*	Historical construction error*	Service line	Connections	660	4.200	10.260
5*	Historical construction error*	Service line	Connections	0	4.200	4.200
6*	Molestation*	Service line	Connections	9.000	12.000	2.100
11*	Corrosion/ageing	Mains LP	Mains	0	209.400	209.400
13*	Historical construction error*	Mains LP	Mains	47.580	33.000	80.580
14*	Molestation*	Mains LP	Mains	4.380	2.400	6.780
23*	Historical construction error*	Mains HP	Mains	0	4.800	4.800
24*	Own activities	Mains HP	Mains	8.760	13.200	21.960
27*	Molestation*	Gas meter	Connections	0	26.400	26.400
29*	Component defect	Station*	Connections	0	51.000	51.000
30*	Molestation*	Station*	Connections	0	102.000	102.000
	Total			**3.628.800**	**4.866.000**	**8.494.800**
	Total new precursors			**75.780**	**462.600**	**538.380**
	Total original precursors			**3.553.020**	**4.403.400**	**7.956.420**
	Totals in 2006			**3.847.636**	**3.283.404**	**7.131.040**

As can be seen, the new precursors did not contribute much to the Category I risk, only about 2%. This is no surprise, as if they would have caused an Category I incident, they would have been in the original set. In the Category II area the contribution is more substantial (about 14%) which is enough to justify their inclusion. On the total risk the contribution is 8%, as Category I and II risk are comparable in size. The risk of the old precursors increased slightly (from 7,1M to 8M), though the risk in Category I (the real accidents) slightly decreased. The increase is probably due to the new metric, which was added to address the issue of evacuation in Category II incidents, the decrease is probably just the inherent variation in the annual incident value (see Figure 39).

8.5.4 Actual Implementation

After the revision of the method, the network operators started using the Safety Indicator for measuring their safety performance. However, during this implementation large differences occurred between network operators, that were hard to explain based on the perceived quality of the grid. Grids, constructed with the same materials, operated and inspected on the same code of practice, in a similar area, sometimes scored completely different in terms of the Safety Indicator. This was already mentioned in the first publications on the Safety Indicator (Hermkens, 2005, Wijnia and Hermkens, 2006). Potential explanations are differences in registration culture, or even just random variance. This sheds some doubts on the applicability of the Safety Indicator in the regulation of the distribution business in the Netherlands. If the safety score is essentially a random metric (despite efforts to make it more stable), it is not useful as a steering guide. And if it is differences in registration culture, the floor is open for manipulation. For this reason, the results per network operator have not been released to the public. In a report to the regulator KIWA (Pulles and van Eekelen, 2009) mentions that the Safety Indicator is only useful for regulation if the problems in incident registration are solved, and if the safety culture in organizations is mature enough. The network operators are currently working on a more consistent way of recording incidents, by standardization and stringent auditing. This might result in a certification scheme similar as that applied for registration of electrical outages (NESTOR E).

8.6 Conclusions

It is possible to link the accidents occurring in the grid to a limited number of precursors, that could be used to measure the safety of a system based on the number of occurrences of minor incidents. The benefit is that there are many more small incidents than large ones, and that an increase of small incidents would occur before the increase in accidents, thus providing the network operators time to respond. However, this is only possible if the registration of the minor incidents is reasonably reliable, as the benefit of having more data available might be offset by the diminishing quality of the data. This was considered in the original version of the indicator, and incidents were limited to incidents serious enough to make a phone call about. In the test run of the indicator, some amendments were made, like a new value, other normalization and new precursors, but these additions did not alter the method much, nor the measured risk in the system.

However, despite the considerations on the data quality, in implementing the Safety Indicator this proved to be the showstopper. The Safety Indicator showed large differences between regional network operators with very comparable grids, which are probably due to differences in registration. For this reason, results on the individual network operators have not been published yet, and the Safety indicator is not adopted into regulation. For any progression of the Safety Indicator, the data quality issue has to be sorted. Future steps might be certification of the registration process, as happened in registration of electrical outages in the Netherlands.

9 Whole system optimization[149]

This chapter was published before as **LONG TERM OPTIMIZATION OF ASSET REPLACEMENT IN ENERGY INFRASTRUCTURES (WIJNIA ET AL., 2006)**[150].

9.1 Introduction

In most western societies, energy distribution systems are quite old. Most of them were built in parallel to the booming economies of the 60s and 70s, an era in which economic growth was synonymous to an increase in power consumption. In this period one could see a growth in the power consumption of about 7% annually, which meant the demand doubled roughly every 10 years. The major concern of the energy companies was to expand the capacity of the network fast enough to keep up with demand, so new lines were built with an oversized capacity. Most of the staff currently at work in the industry was hired during these years. However, at the end of the 70s the economic growth came to a halt. By the time the energy companies realized that the construction rate could be slowed down, the network already had a large surplus. Furthermore, the oil crisis of the 70s created awareness of energy consumption, which resulted in energy conservation programs. So when the economic growth resumed, it was for quite a long time not reflected in the asset construction rate. At the moment, the asset construction rate is even lower, due to the tariff cuts introduced by the regulator to emulate a market (Figure 42)

The oldest[151] assets in operation are about 60 years old, the average age of the asset base is about 30 years. However, assets do not have perpetual lives, as normal wear and tear takes its toll. For example, almost all assets are vulnerable to corrosion. In electricity networks almost all parts are vibrating at 100 Hz, resulting in metal fatigue. Those influences build up over time, and nobody knows for certain how fast deterioration will take place. The best estimate for the lifespan is between 40 and 80 years, depending on the type of asset. This means we can expect at least part of the asset base to reach end-of-life within 10 to 20 years. End-of-life is defined in this paper as a non-repairable failure of the asset.

---

[149] The results of the experiment were presented at an asset management conference organized by Enexis to celebrate the certification against PAS 55. At this conference some journalists were present who picked up the story on the potential problems in replacing an ageing asset base. About two weeks later, questions were asked about the presentation in Dutch parliament (Tweede Kamer der Staten Generaal, 2007). Even though this was by no means the first time the need for thinking about the replacement of infrastructure assets was ventilated, apparently the quantitative prediction of performance triggered some political concerns. In the following years, the Dutch regulator became more interested in the problem of asset replacement. Enexis was invited to present the findings of the experiment for a conference of European regulators in 2009. Neither the findings nor the proposed strategies were fundamentally questioned.

[150] Wijnia, Y. C., Korn, M. S., de Jager, S. Y. & Herder, P. M. 2006. Long Term optimization of asset replacement in energy infrastructures. *2006 IEEE Conference on Systems, Man, and Cybernetics.* Taipei, Taiwan.

[151] This means, the oldest in significant asset numbers. Some cables are from the 1920s but these are exceptions.

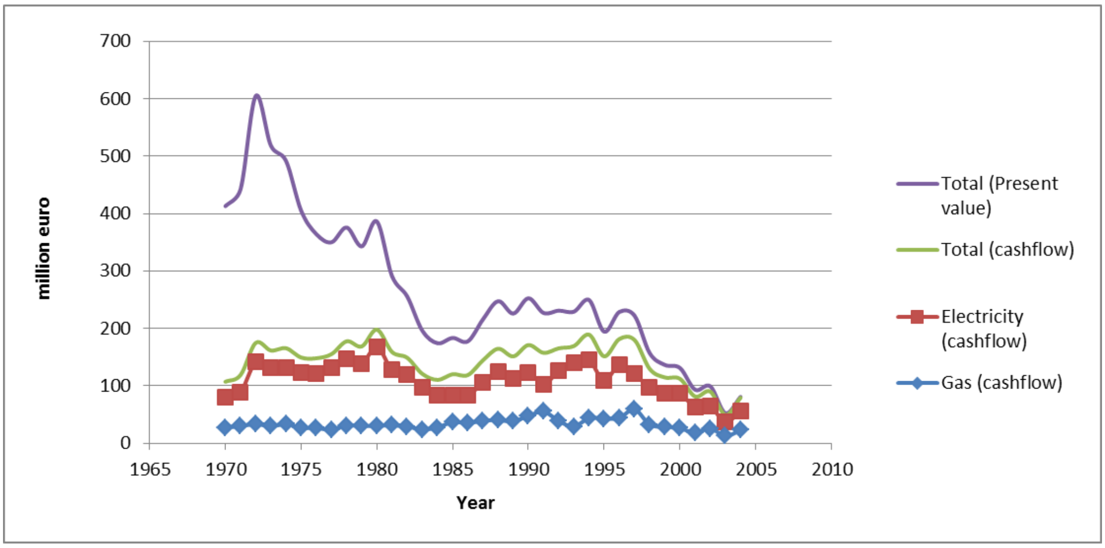


Figure 42: Historic investment volumes for Enexis. In cash flow terms the volume is quite constant, but corrected for inflation (the "present value") the peak in the 70s is clearly visible.

In parallel, engineering staff retires after about 40 years in service. As most of them were employed during the seventies, we can expect a large part of them to retire within 10 years. Figure 43 shows that about half of the work force will retire within 7 years.

Figure 43: Age profile of the employees. The average age is 47 (red line). More than half of the population is over 50, retirement age is about 60 (after 40 years of service).

Unfortunately, this is just before the large scale asset replacement program should start. In other words, first the company work force would be halved, and then the work volume would double, in effect quadrupling the work load. Even though some efficiency improvements could be realized, a quadrupled work load is clearly not feasible.

Although the problem mentioned above is serious enough, it might be even worse. This is because in times of underemployment certain activities will be executed, regardless of the overtime needed (these activities are failure restoration, corrective maintenance, corrective replacements and new connections), because not executing them will most likely put the company out of business. Activities like preventive maintenance and preventive replacements will only be carried out if any time is left.

However, corrective maintenance and corrective replacements take more time than the preventive equivalents. So, once the asset age causes failures in the network, its repair will take time away from preventive actions, thus increasing the probability of failure (**Figure** 44). Those failures will again demand time at the cost of preventive work, and so on. The spiral of decline is born.

Summarizing the situation, we can state that the energy companies face a serious challenge in replacing the assets, simply because the people needed to do the work will be lacking. The shortage will amplify itself, resulting in a spiral of decline. It seems wise to speed up the asset replacement program.

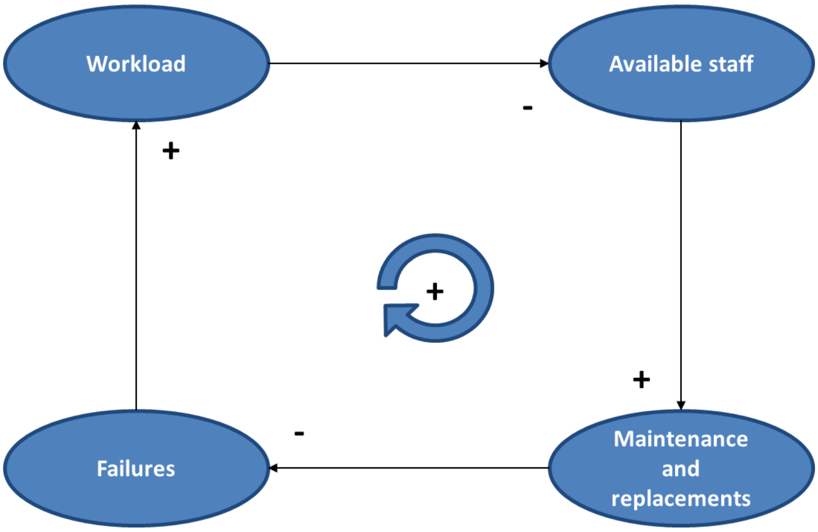


Figure 44: The spiral of decline. Corrective actions take more time than preventive actions. An increasing number of failures will reduce the time spent on maintenance and replacements, therefore increasing the number of failures.

In order to do so, an asset replacement program needs to be developed for an asset base with an uncertain life span, and that addresses feasibility, financial value, reliability, safety, workload or a combination of those items. First, we will describe the objective function of our asset replacement program. Second, we will estimate the number of failures if no preventive action is taken. The next step is to determine the optimal replacement moment for the asset base. Finally, we determine how the constraints in the staffing dimension influence the feasibility of this optimal replacement moment. Only then we are able to see if it makes sense to speed up the replacements. We will finish this paper with a discussion of uncertainties and an outlook to the future.

9.2 Objective function

Most utility companies care about three values: financials, reliability and safety[152]. For manageable optimization purposes, a single-objective is pursued (Wijnia and Warners, 2006). The common way forward is to express safety and reliability in monetary values (Sugden and Williams, 1978). Strictly speaking, there is no monetary equivalent for human lives(Slovic and Weber, 2002), but a value of about 10 M euro per human life is used in the industry (Wijnia and Hermkens, 2006). For reliability,

---

[152] Other values could be legal, environment, brand, but for a monopolistic utility brand is not important, legal is a constraint, and environmental damage potential is limited, so that it becomes a legal constraint.

the Dutch distribution companies use the regulator's incentive on reliability, which should reflect the societal value (Baarsma et al., 2004), being 25 eurocents per customer minute lost (CML).

As asset replacement is an investment decision, the best objective function is the Net Present Value (Brealey and Myers, 2000). As the replacements do not introduce new financial risks to the companies, it seems reasonable to use the weighted average cost of capital (WACC) as a discount factor. The Dutch regulator set the WACC at 6,5% in nominal terms, or 4% in real terms, assuming an inflation of 2,5% annually. For a real discount rate of 4%, it takes about 50 years for a perpetual annuity to reach 90% of the total value, so a time horizon of about 50 years is needed. This gives us our objective function: maximize the discounted weighted sum of the values financials, reliability and safety, or:

$$\max\left(\sum_{n=1}^{50}\frac{return_n - safety_n - CML_n}{(1+r)^n}\right) \tag{9.1}$$

in which *n* is the year, *r* is the real discount rate and *CML* is customer minutes lost. *Return* is the difference between network income and network costs, being the total of maintenance, operation and investments. As the income is assumed independent[153] of the network *costs*, this max function could be replaced with the following min function.

$$\min\left(\sum_{n=1}^{50}\frac{costs_n + safety_n + CML_n}{(1+r)^n}\right) \tag{9.2}$$

This will be the objective function used throughout this chapter.

9.3 Expected failures

In estimating the future number of asset failures per time unit it is worth noting that there is no such thing as an average asset. A typical distribution company might have more than a thousand different types of assets, each with its own failure modes, maintenance requirements, failure effects, failure rates (both age dependent and age independent) and so on. However, it is not feasible to model all those differences with limited modeling resources. On the other hand, modeling them as one type of asset might work for trend forecasting, but it would not allow for easy model validation. To deal with this controversy, some way to categorize the thousand and more types of assets into a manageable number is needed. As asset failures are risks, we will use the risk process proposed by Morgan et al (Morgan et al., 2000). This process was first demonstrated for air quality issues, but we have further developed the process (Wijnia and Herder, 2004) for risks in infrastructures and we proved the concept for gas networks(Wijnia and Hermkens, 2006). The simplified risk process is shown in Figure 45.

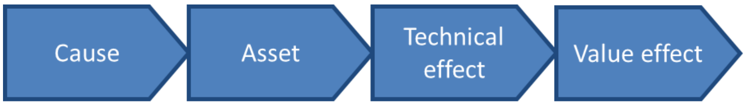


Figure 45: Simplified risk process for asset base.

---

[153] This is the principle of price-cap regulation instead of the cost plus principle.

In a review of the value effects of asset failures, it was shown that the effect of ‘causes’ on reliability was largely determined by the network level[154] at which they occurred, whereas the safety effects were largely determined by the asset type. For example, in electricity the network levels are i) high voltage transmission, ii) high voltage substation, iii) medium voltage transmission, iv) medium voltage substation, v) medium voltage distribution, vi) ring main unit and vii) low voltage distribution. Outages on level i affect about 100.000 customers, on level vii only about 10. For safety effects, the relevant asset types are transformers (fire risk due to oil content), switchgear (explosion risk when breaking short circuit), cable (negligible safety risk) and secondary equipment (no direct value consequences in case of failure). Using this classification, the total asset base of more than a thousand types could be compressed into about 70 different types.

The next step in estimating the expected failures was to determine the failure rates for those asset groups. We used the bathtub curve model. This model assumes components have a high failure rate right after their start up (infant mortality) and after a certain period (wear out), whereas the failure rate between these periods is more or less constant. Because the assets we are looking at are already constructed, we limit the bathtub curve to the wear out phase. This is not an uncontroversial choice, as research has shown that only a limited number (about 11%) suffers from wear out. However, this might be because assets are sometimes replaced for other reasons than failure[155]. Assets in distribution networks will be used until end of life. Besides, it can be shown that the alternatives for the bathtub (*constant failure rate* or *limited life span)* are the extremes of a parametric bathtub curve model:

$$h(t) = h_0 * e^{\frac{-\ln(h_0)}{T_1} * t} \tag{9.3}$$

In this equation, h_0 is the failure rate at age = 0 (assuming no infant mortality), T_1 is the age at which the failure rate is 1, and t is the actual age. With a very high value for T_1, this model will look like a constant failure rate; with a very low h_0 the asset will fail at a very precise age. In the table below some asset types for the medium voltage transmission network and their parametric values are displayed. The value $T_{0,01}$ is the age at which 1% of the population has failed[156], or the trouble free lifespan.

---

[154] This is the so called Netvlak.

[155] Take for example the short life span of a mobile phone.

[156] It was this value and the T_1 value that were assessed on the workshop floor, the h_0 rate was calculated from those two values.

Table 28: Asset Ageing figures.

Asset	h_0	$T_{0,01}$	T_1
1 MST Circuit breaker open	1,00E-04	60	120
2 MST switch closed oil	4,64E-04	40	100
3 MST switch closed air/gas	1,00E-04	40	80
4 MST protection mech	2,15E-05	40	70
5 MST protection electr	9,99E-05	20	40
6 MST protection digital	9,99E-06	15	25
7 MST cable pilc	3,16E-05	50	90
8 MST cable pilc peatland	9,79E-04	40	60
9 MST cable xlpe	1,39E-03	30	100

Next, the asset age profiles are determined. The profiles are shown in Figure 46.

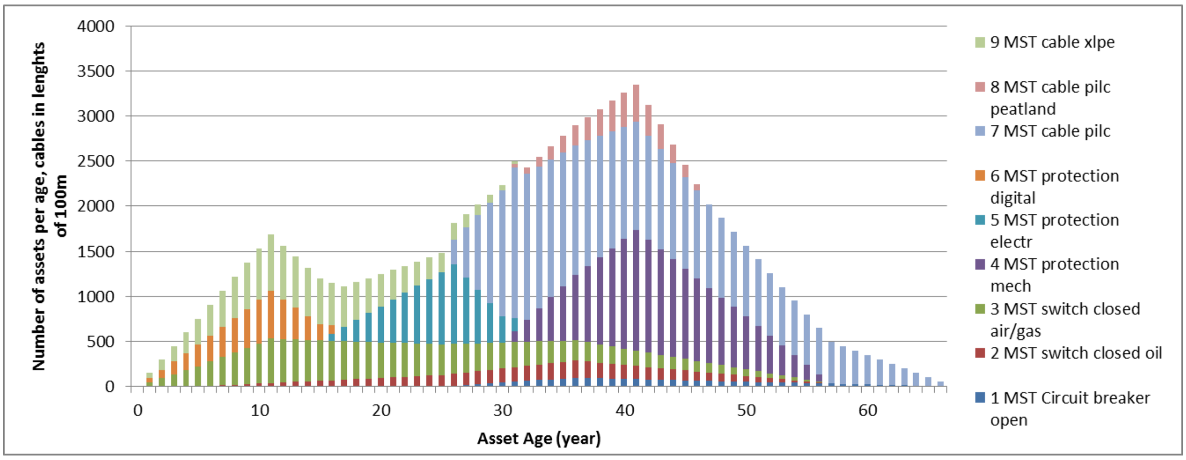


Figure 46: Age profiles assets. Cable assets are counted in lengths of 100m. All asset costs per unit are in the same order of magnitude (1-10 k€).

These data are fed into a Systems Dynamics model (constructed in the tool Powersim) that simulates the asset failures as a function of time. By multiplying the age profile with the failure rates the number of asset failures is calculated. Failed assets are then removed from the population, the age profile is shifted 1 unit to the right (assets are 1 year older), and the number of asset failures is calculated for the next year, and so on. In the diagram below the failure rates for the years 2005-2055 are shown[157]. Those failure rates can be translated into costs, safety incidents and customer minutes lost by multiplying the number of failures per asset type with the average failure consequences per asset type.

---

[157] The CML figure in the peak is about 100 minutes per customer per year, 4 times the current level in the Netherlands. This would not be accepted. In European perspective, however, this is still a good performance (see table 2, chapter 2).

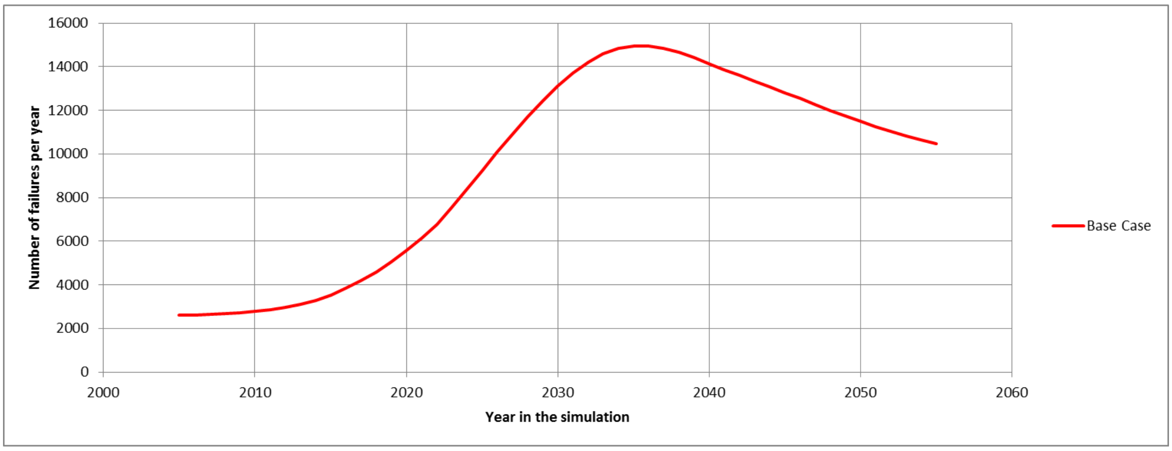


Figure 47: Estimated failure wave in the base case.

With the asset age profiles used and the failure rate calculations it is clear that the number of failures will rise dramatically within 15 to 30 years. The pattern is validated against the actual number of failures, which is about 2.500 failures a year. Any lower life expectancy of the assets would result in a higher number of failures, so we have used conservative lifespan estimates.

9.4 Optimal replacement age

We have assumed that assets that fail cannot be repaired, they would have to be replaced. However, replacement after failure is more costly than preventive replacement. This is because failures occur at inconvenient times, require service restoration action, temporary measures and so on. Furthermore, uncontrolled failure has costs in the form of Customer Minutes Lost and safety incidents, which would not occur if the replacement was planned before failure. As the failure rate rises, the failure risk (= failure probability times consequence) increases. Figure 48 shows that if this increase of the failure risk becomes higher than the decrease of the planned replacement costs[158], it makes more sense to take preventive action than to take the risk. Assuming the consequences do not depend on asset age, the risk is proportional to the failure rate. In Figure 48 the total costs (sum of failure risk and replacement costs) of an asset are plotted against the planned life cycle. This includes the monetary equivalent of the customer minutes lost and the safety incidents. For the total costs the annuity of the present value is used, being the annual amount of money with which the asset could be operated perpetually.

---

[158] The graph shows that it is about the derivatives and not the absolute value. The minimum of the yellow line is clearly at a very different age (about 55) than the one at which failure costs equal the replacement costs (about 40).

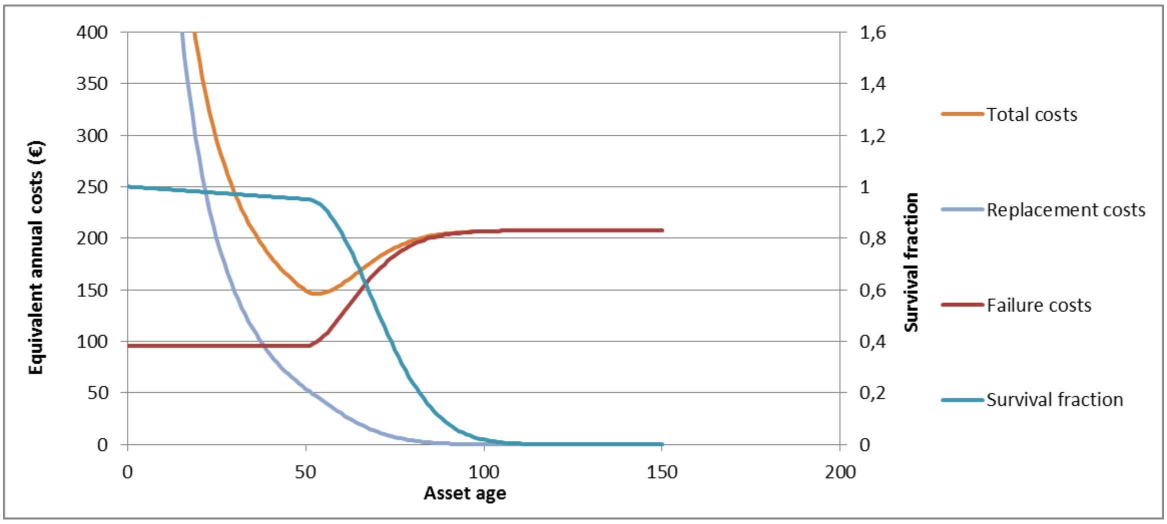


Figure 48: Replacement age optimization.

It is clear this asset has a lowest life cycle cost at a little over 50 years.

Although this approach provided a clear answer for the optimal replacement age, it did not allow for easy sensitivity analysis. If the optimal replacement age would have to be determined from a graph for more than 70 assets for every change in the parameters, making multiple runs would require too much time. Therefore, in the overall model, this replacement age was approximated with the equation[159], shown below. The optimal moment is when additional risk equals the replacement costs:

$$H(t)*(C_c - C_p) = r*C_P \tag{9.4}$$

C_c is the corrective replacement cost, C_p is the cost of preventive replacement, r is the interest rate and *H(t)* is the conditional probability that the asset fails in year *t*, whereas *h(t)* from Eq. (9.3) is number of times the asset would fail per year (the inverse of the mean time between failures). The simplified[160] relation between the two is:

$$H(t) = 1 - e^{-h(t)} \tag{9.5}$$

Substituting (9.3) and (9.5) in (9.4) and solving for *t* gives:

---

[159] This is one of the most important reasons for using a parametric failure rate model: it allows for mathematical solving of the optimal replacement age.

[160] Technically, it would be the integral of h(t) over the year. The simplification holds if the failure rate is constant over the year. For asset ageing, the typical growth rate of the failure rate is in the range of 10% per year, making the simplification acceptable.

$$t = -\frac{T_1}{\ln(h_0)} * \ln\left(-\frac{\ln\left(1 - r * \frac{C_p}{C_c - C_p}\right)}{h_0} \right) \tag{9.6}$$

With this formula, changes in the parameters could be directly translated into different optimal replacement ages. The table below shows the optimal ages for nine assets

Table 29: Optimized asset replacement ages.

	$T_{0,01}$	T_1	T_{opt}
1 MST Circuit breaker open	60	120	94
2 MST switch closed oil	40	100	74
3 MST switch closed air/gas	40	80	64
4 MST protection mech	40	70	53
5 MST protection electr	20	40	27
6 MST protection digital	15	25	18
7 MST cable pilc	50	90	78
8 MST cable pilc peatland	40	60	51
9 MST cable xlpe	30	100	78

9.5 The effect of personnel constraints

Now that we have determined the optimal replacement age, we will evaluate the feasibility of replacement at this optimal age. Feasibility is determined by the availability of staff for preventive replacements. To model this, we used the following assumptions.

1. Employees enter the company and retire at fixed ages (in the model 20 and 60 years respectively)
2. Corrective replacements are always carried out (even at the cost of overtime), preventive replacements only if working hours are available
3. Preventive replacements are prioritized to maximize value for money
4. Recruitment matches the retirement outflow plus the overtime plus the non-executed planned replacements, but is limited by the recruitment capacity (set at a maximum of 100 employees per year, at a company size of 1.500).

Outflow due to retirement is modeled in the same way as asset failures. The number of asset failures determines the corrective replacements. Assets older than the optimal replacement age are put on the “to-do-list”. The assets on this list are ranked in accordance to the ratio of corrective over preventive costs. Only the part of the to do list that fits within the budget is actually replaced, the rest is left in the network for further ageing. In Figure 49 the green line shows what would happen if a preventive replacement program was executed.

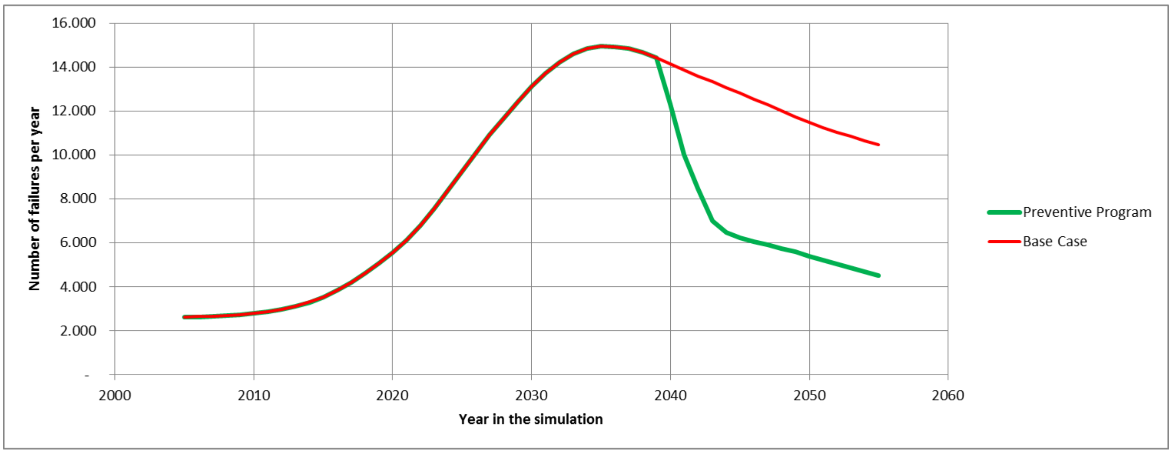


Figure 49: Failure wave in base case (red line) and with a preventive replacement program in place (green line).

Apparently, the preventive program does not have any effect for 35 years. A closer inspection of the model showed that because of the high retirement rates not enough new employees could be recruited to carry out the preventive work. Only by the time the failure wave has passed its peak, recruitment has caught up with the personnel needs, and the remaining old assets are replaced at a very high rate, freeing up even more time. This leads us to the first major conclusion: simply putting a technical replacement program in place does not help, additional measures are needed. Furthermore, the results showed that the lack of personnel during the first 30 years is not the only personnel problem, since there would not be any work for the recruited replacement personnel after the replacements. One could argue that this problem is not a real problem but only the result of the way the recruitment policy is modeled, as the model continues to hire staff even if the peak is near. In reality, the recruitment pace would be slowed down. However, determining the right time to slow down requires looking forward, which is not possible in the used software[161]. In addition, some alternative recruitment policies were tested (a target size for the organization and only hiring for the 'delta' in the to-do-list and not the complete to-do-list). The first one performed better but was not feasible (the optimal organization size could only be determined in retrospect, and a small deviation from the optimal size cancelled the advantage), whilst the second dampened the personnel overshoot, but was too slow in the startup phase. This affects the present value terms negatively.. The original model, though a bit crude, performed best.

9.6 Strategic alternatives

We have seen that a replacement program alone is not enough to manage the failure wave. The personnel constraints do not allow for any preventive action in the first 30 years. Additional measures would have to be developed that influence capacity, workload or timing.

These measure can be formulated along 4 directions.

1. Higher recruitment rates: this would allow for earlier replacements

---

[161] The work around is to run the model, determine the right moment and run the model again with a policy change at that moment. It requires a few runs to determine the optimal timing.

2. Flexible personnel resources: this would allow for earlier replacement, but at a higher cost. The overcapacity would not occur.
3. Speeding up replacements: Would take advantage of current personnel capacity and thus allow for earlier replacement
4. Efficiency improvement: Would decrease the workload, thus freeing time for replacements.

For each of these alternatives, the cost could be minimized. However, in reality any combination of actions could be chosen. In Table 30 we show the outcomes of the simulation. The base case, in which no preventive replacements are carried out, is set at a net present value[162] of zero. The values of the other options are deviations from this base case.

Table 30: Goal function outcomes for alternative strategies.

Strategy	NPV (M€)	Overshoot (Fte)	Overshoot (%)
Base case (doing nothing)	0	157	7%
Only preventive program	110	1.156	53%
Preventive, recruitment (150)	558	703	32%
Preventive, flexibility (200)	574	136	6%
Preventive, Speeding up replacements 3 years	563	1.073	49%
Preventive, 50 % Efficiency improvement over 20 years	893	504	23%
Preventive, recruitment (150), flexible (400)	583	236	11%
Preventive, recruitment (150), flexible (200), efficiency (50%)	988	42	2%

From this table it is clear that a preventive replacement program would only add little value to the company, and cause a large staff abundance after 35 years. Adding a higher recruitment effort, would increase the company value and reduce the overshoot. Recruiting earlier and faster would increase the value. Efficiency improvements deliver the highest total value to the company. Other single solution alternatives result in about the same figures in terms of company value, the flexible solution is the only one reducing the overshoot problem. Efficiency improvements would add most value to the company, but there is some uncertainty whether this improvement is feasible. A smart combination of directions (excluding efficiency effects) is slightly better than single direction alternatives, the combination with efficiency improvements is much better, and has a negligible staffing overshoot.

9.7 Sensitivity analysis

In the discussion about the failure rates we mentioned that these were basically unknown. We know for certain, however, that the assets will not fail earlier, because then we would be in the middle of the failure wave right now, which in fact we are not. However, the life span might be longer than we now assume is reasonable. What would happen to our conclusions if we would, for example, make all life spans twice as long, meaning that the majority of assets would pass the age of 100 years? As one would expect, if life spans are long enough, there hardly is an asset ageing problem. Twice the life span is long enough, but even with a 50% increase of the life span the issue would still persist.

---

[162] This net present value includes the CML and safety incidents.

A second uncertainty is whether the action is needed right now. What would happen if we waited for another 5 years before really starting the replacement program? Running this scenario showed that it was not as good (appr. 200 M€ negative) as starting right now if the failure rates were according to Table 28, but it still was much better than the base case (appr. 400 M€ negative). If the life span would be 50% longer, a 5 year delay would indeed deliver value, showing that the uncertainties in the life span impact the selection of the best way forward to a meaningful extent.

9.8 Conclusions

We have seen that the energy distribution company[163] we investigated faces an asset failure wave. These failures will induce costs, customer minutes lost and safety incidents. To reduce those consequences, a preventive replacement program is needed. However, only a cost-optimized replacement program is not enough, as it would not be feasible due to personnel constraints. Either the recruitment rate should be increased, the staff should be made more flexible, the replacements should be sped up or the efficiency of the company should be increased. It is best to make a smart combination of those options, as it increases the company values the most. This conclusion is not very sensitive for misjudgments in the asset life spans. Even with a 50% increase the conclusions would still hold, only if the life spans are much longer than currently assumed reasonable, the problem disappears, thus making conclusions about the problem invalid. However, values for life spans that would eliminate the asset replacement problem and the failure wave are unrealistically high.

9.9 Future work

The results presented in this paper assumed that the current asset base needs to be replaced by a similar asset base. This is only true when no changes to the current power grids are expected. This assumption, however, can be challenged since the ways of energy production, conversion and tranportation may change rapidly in the years to come. For example, micro-scale Combined Heat and Power (μCHP) units may penetrate our households, creating the need for a grid that can deal with bidirectional flows (in case of electricity delivery into the grid by households) and it may render the current power grid as a mere backup system to the μCHP units. Another example presents itself in developments in the Dutch natural gas grid, which is currently a monodirectional star-like network. When other energy carriers, such as hydrogen or bio-gas are fed into this network at local points, or when the Netherlands become an international gas hub, we may need to reconsider its topology.

Secondly, the grid has developed into its present topology through a haphazard, ad-hoc fashion as new extension were added to the existing grid when opportunity arose. If we were to redesign and construct the grid from scratch, an optimal network would look quite different from the currently existing grid. So, when we are considering our replacement strategy, should we not also rethink our network topology? Or, in view of the uncertain future developments such as μCHP and new energy carriers, should we replace, design and build robust infrastructures that can deal with many future uncertainties or should we aim for flexible solutions, which may render our current asset base obsolete? We will address these questions in our future work, eventually leading to recommendations for an optimal and flexible asset management strategy.

---

[163] We think that it applies to all energy distribution companies, as their asset bases are comparable.

9.10 Epilogue

One element of the model was to prioritize replacements in time of scarcity. The mechanism used for this ranking was highly similar to that of the procedural optimization of the portfolio decision. The proposed projects (i.e. nominated asset replacements) were ranked according to their yield, the ratio of expected corrective cost over preventive costs. Additionally, the optimization used some form of a hurdle rate, as assets only would be nominated if they were above the replacement age. This part of the model allowed to test the value of prioritization, by comparing the ranked results with unranked results (i.e. random selection of nominated assets). As mentioned in the epilogue to the portfolio decision, this more or less reflects a decision approach where decisions are made one by one. Each of the nominated assets had a positive business case, as they were only nominated for replacement at or beyond the optimal replacement age. This comparison is shown in the graphs below. The green lines represent optimization of the portfolio by ranking the assets, the red lines represent unranked serial decision making.

In terms of the number of incidents (the top graph) the benefit of prioritization is relatively limited, about 20%-25% at the most. But for the SAIDI (middle graph) and the number of safety incidents (bottom graph), the results are much more dramatic. The improvement prioritization brings is about 50% in peak situations. The graphs also demonstrate that prioritization only adds value in certain periods and not continuously. This occurs in times of partial scarcity as argued before. If there are enough resources to replace all nominated assets prioritization does not matter. Interestingly, a portfolio decision also does not add value if resources are very scarce and none of the nominated assets would be replaced. This for example was the case in the test of Figure 49, when obligatory corrective actions limited the capability of replacing assets to zero.

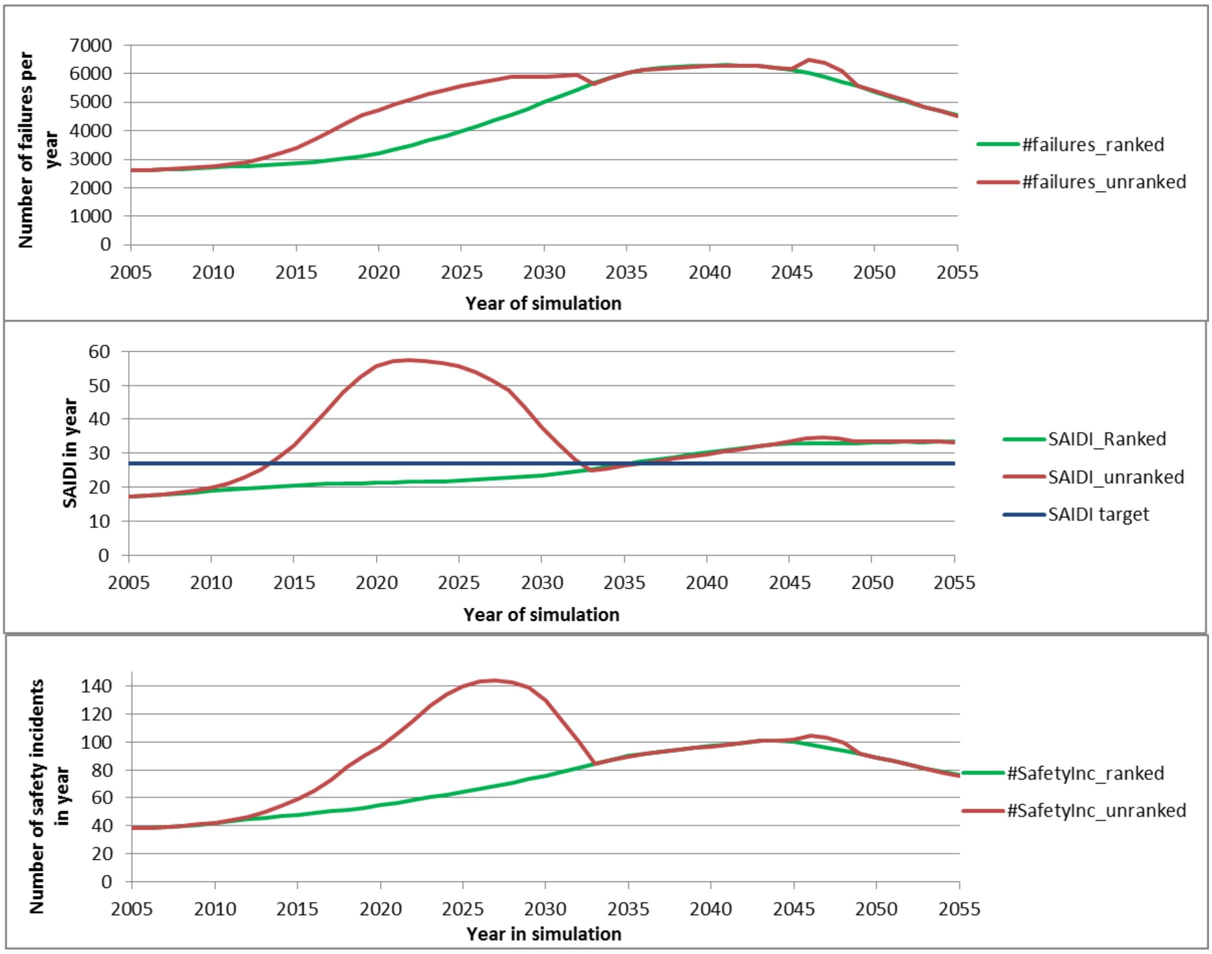


Figure 50: Comparing ranked with unranked decision making. The green lines show the achieved performance when decisions are ranked in a portfolio decision, the red lines the performance when budget is allocated on a FIFO principle representing serial decision making. The top graph is the number of incidents, the middle graph the SAIDI (average time per year a customer experiences an outage) and the bottom graph the number of safety incidents.

Perhaps an unexpected aspect of the diagrams above is that even in the ranked solution there is a considerable rise in the number of failures, outages and safety incidents. This may raise the question whether the constraints are still too tight, or that the optimal decision was wrong, or even that the model was entirely wrong. Many of such questions were asked in discussing the results of the experiment within the organization and to outside stakeholders (e.g. those mentioned in the footnote to the title of this chapter). Neither of these proposed hypothesis proved to be right, which could be demonstrated immediately in the discussion by running the simulation with alternative parameters. The rise of the failures even with an optimized replacement program in place is an unavoidable side effect of risk based optimization of replacements. The optimal moment is when the risk for next year is equal to the cost of advancing the replacement one year. But as the risk slowly develops, it is higher in the years right before replacement than it is for a new asset. The aging of an asset base thus inevitably increases the number of failures. Only when the replacement strategy would result in a decreasing age of the asset base, the number of failures would go down again. This happens in the diagrams after 2045. Demonstrating this to the skeptical audience in general resolved the issue.

10 Capturing the risk position

10.1 Introduction

One of the key enablers of asset management is risk management (Woodhouse, 2014). In order to manage risks, it is vital to know them and to understand them. In literature it is debated whether that is fully possible. There does not seem to be agreement on what a risk precisely is, given the numerous definitions that are around (Beer and Ziolkowski, 1995, Aven, 2012, Slovic and Weber, 2002). Additionally, some regard risk as inherently subjective, not existing out there waiting to be measured (Slovic and Weber, 2002) Further indicators for the impossibility of fully understanding risk can be found in the lack of generally accepted lists or taxonomies of risks to be managed in various industries. Given the relevance of risk management in and the maturity of many industry sectors one would expect such lists to be standardized if possible, but no evidence for this was encountered. It is more the other way round. Within the same industry risks may pop up in different branches of the taxonomy. There even is little agreement on how to capture risks or how to build a register. Standards like ISO 31000 (ISO, 2009a) and the accompanying guideline 31010 (ISO, 2009b) provide little more than a list of potential techniques and the requirement to use the tools and techniques fit for purpose. On the other hand, the need for a full list of risks is little debated.

In this chapter, the construction of such a full list of risks is investigated. First, literature with regard to risk identification will be reviewed. After this, the “experiment” conducted at Enexis in the period from 2002-2007 will be described. This consisted of several rounds of identification and analysis, until a stable process and a good overview of the risks was achieved. As the predicted total consequence of the risks on that list was more than what was actually observable, some more thought is spent on the reason why that was the case and how the mismatch could be fixed. This is reflected back on the characteristics of the ideal risk register.

10.2 Literature review

10.2.1 Risk identification

In risk management literature it is often emphasized that risk identification should be comprehensive. (The Institute of Risk Management (IRM) et al., 2002) for example states in **A RISK MANAGEMENT STANDARD** that:

> *Risk identification should be approached in a methodical way to ensure that all significant activities within the organization have been identified and all the risks flowing from these activities defined. All associated volatility related to these activities should be identified and categorized.*

A similar opinion is advocated in **CATEGORIZING RISK FOR RISK RANKING** (Morgan et al., 2000), stating that “Risk categories should be exhaustive, encompassing all relevant risks.”[164]

This view has not lost any strength over the years. The standard on risk management ISO 31000 (ISO, 2009a) states that “Comprehensive identification is critical, because a risk that is not identified at this stage will not be included in further analysis.”

---

[164] See also Table 31 of this thesis.

Unfortunately, what method of identification gives such a comprehensive overview is not very clear. In the ISO standard (ISO, 2009a) itself only hints are given on how to conduct risk identification:

> *The organization should apply risk identification tools and techniques that are suited to its objectives and capabilities, and to the risks faced. Relevant and up-to-date information is important in identifying risks. This should include appropriate background information where possible. People with appropriate knowledge should be involved in identifying risks.*

This hint even contains a form of self-reference, as the tools for risk identification should be suited to the risks faced. But how to know which risks are faced before identifying them?

In the accompanying document **ISO/IEC 31010 RISK MANAGEMENT: RISK ASSESSMENT TECHNIQUES** (ISO, 2009b) little further guidance is given:

> *Risk identification methods can include:*
>
> - *evidence based methods, examples of which are check-lists and reviews of historical data;*
> - *systematic team approaches where a team of experts follow a systematic process to identify risks by means of a structured set of prompts or questions;*
> - *inductive reasoning techniques such as HAZOP.*
>
> *Various supporting techniques can be used to improve accuracy and completeness in risk identification, including brainstorming, and Delphi methodology.*

In the appendix to the guideline some techniques for risk assessment are listed, including a recommendation for the phase in the process. Some methods that are strongly applicable in the risk identification phase like brainstorming, interviews and Delphi can be regarded as a very broad scan for risk, while others like FMEA and RCM seem to be more restricted in the scope of risks they can capture. On the other hand, they are regarded as strongly applicable for the whole process, a clear benefit over the mentioned broad identification methods.

In a review of the quality of national risk assessments (Vlek, 2013), the use of scenario analysis as a means of risk identification was evaluated. The concluding comment with regard to scenario development was:

> *Thus, on the one hand, scenario developers may be overzealous and come up with incident warnings that may be exaggerated. On the other hand, as a constructive process, the timely identification and elaboration of hazard and threat scenarios may be too latent or sloppy, such that real dangers are denied or neglected, and may become manifest as unfortunate surprises.*

The solution advocated is not a specific method but a social one in the form of a well-organized and carefully conducted process.

> *Here, there seems to be no other solution than a well-organized, carefully conducted process in which the possible false-positives and false-negatives are creatively considered and their more or less likely costs and benefits are explicitly weighed.*

Construction of the risk position apparently is still regarded as a procedural approach.

10.2.2 Characteristics of the ideal risk register

In **CATEGORIZING RISK FOR RISK RANKING** (Morgan et al., 2000), a list of desirable attributes for a risk categorization system is presented (see Table 31). It has to be recognized that risk categorization is not the same as risk identification. Typically, the identified risks (or hazards, which are used interchangeably in the introduction) will be clustered into categories so that a manageable number remains. Risk categorization thus comes after the identification of risks. Yet, many attributes presented seem desirable for any list of risks, whether categorized or not. The attributes will be used as a reference point for evaluating the list of identified risks.

Table 31: Desirable attributes of an ideal risk categorization system for risk ranking, after Morgan et al. (2000).

Categories for risk ranking should be:

1. **Logically consistent**
    - 1.1. Exhaustive so that no relevant risks are overlooked.
    - 1.2. Mutually exclusive so that risks are not double-counted.
    - 1.3. Homogenous so that all risk categories can be evaluated on the same set of attributes.
2. **Administratively compatible**
    - 2.1. Compatible with existing organizational structures and legislative mandates so that lines of authority are clear and management actions at cross purposes are avoided.
    - 2.2. Relevant to management so that risk priorities can be mapped into risk management actions.
    - 2.3. Large enough in number so that regulatory attention can be finely targeted, with a minimum of interpretation by agency staff.
    - 2.4. Compatible with existing databases, to make best use of available information in any analysis leading to ranking.
3. **Equitable**
    - 3.1. Fairly drawn so that the interests of various stakeholders, including the general public, are balanced.
4. **Compatible with cognitive constraints and biases**
    - 4.1. Chosen with an awareness of inevitable framing biases.
    - 4.2. Simple and compatible with people’s existing mental models so that risk categories are easy to communicate.
    - 4.3. Few enough in number so that the ranking task is tractable.
    - 4.4. Free of the “lamp-post” effect, in which better understood risks are categorized more finely than less understood risks.

Unfortunately, this ideal system is not presented. As noted in the paper, some attributes are potentially conflicting, like 2.3 (large enough in number) and 4.3 (few enough in number). If these domains overlap there is a solution, but as the paper points out, that is often not the case in reality. Large enough is typically in the 100s range, whereas few enough is limited to a few dozen at most. The suggestion the paper makes with regard to finding the right categorization is to define the applicable chain of risk processes as a starting point. Categorization strategies could then be developed from the risk processes down. In a way, that is creating the taxonomy per phase of the risk process and fitting the identified hazards or threats into that taxonomy.

10.2.3 Taxonomy based identification

Risk taxonomies are in use at least since the early 70’s when Johnson published his work on Management Overview and Risk Tree (MORT) (Johnson, 1973, Johnson, 1975). The focus was on the investigation of accidents. The original document is almost 600 pages. A concise manual (80 pages) has been prepared in 1992 (Knox and Eicher, 1992). In the Netherlands, a guideline on MORT is maintained by the Noordwijk Risk Initiative, latest release in 2009 (Noordwijk Risk Initiative Foundation, 2009).

Risk taxonomies are also used for the identification of risk. The Software Engineering institute of Carnegie Mellon University published a report on containing the taxonomy of software development risk (Carr et al., 1993). However, they explicitly state that

> *The TBQ [Taxonomy Based Questionnaire YCW] application is semi-structured. The questions and their sequence are used as a defining but not as a limiting instrument. [Page 8]*

TBQ is thus regarded as a starting point for a semi structured brainstorm, not a definitive checklist. The TBQ contains 194 questions on 3 classes, 13 elements and 64 attributes. Whether all the attributes are instances of the same concept can be debated, but given the fact that the taxonomy is used as a starting point that may not be very important. It is not to be regarded as an absolute reference point.

Other efforts have been made in the professional world. Deloitte for example has several Risk Intelligence maps, containing all kinds of risk factors(Deloitte, 2010). Unfortunately, reviewing the risk factors in these maps reveals that most risks are about missing interventions and not about the risk events the interventions are supposed to control. From a practitioners solution driven viewpoint this may be valuable, as it gives a reference for what needs to be in place to be in control. But in assessing the risk position it is not very helpful as it can only be used to assess the existence of controls but not for assessing the risk themselves.

Even though several taxonomies could be found, little evidence (if any at all) exists that they are used in practice to establish the risk position. That seems to be more the result of coordinated application of a risk management process.

10.3 Identifying risk in practice

In this section, the identification of risk at Enexis during the period 2002 to 2007 is described in two phases. The first phase considers the batch process like risk identification with which Enexis started asset risk management. This ran from 2002 to early 2005. In this period several risk identification workshops were conducted. Each of these workshops basically started from scratch. This choice was made deliberately to encourage the participants to empty their minds and to get as much buy in as possible. The idea was that the new risks identified in these sessions could be added to the existing risk register. Unfortunately this proved to be incorrect. Each and every list of risks was differently structured and virtually impossible to integrate with previous results. Furthermore, most of the identified new risks were somehow part of the existing (though differently structured) registers. This resulted in the new list replacing the old list, and not using it as an addition. Of this batch process period only the initial effort, which triggered the search for structuring risk, is described. Other identifications were not fundamentally different.

The second phase started with the decision of Enexis to certify itself against PAS 55 in 2005. The asset management process was redesigned to comply with those requirements. One of the major changes was the transformation of the risk management process from a batch oriented one towards a continuous process (Korn and Veldman, 2008). In this continuous process, newly identified risks had to be added to the existing register. This allowed the old registers to be integrated as well. The endpoint of this phase was reached in early 2007 when all the entries of the risk register were assessed and the risk register could be tested for comprehensiveness.

10.3.1 Risk identification by means of a workshop (batch process)

As mentioned in chapter 7, to capture risk, three risk identification workshops were organized to get the risk management process started: one for gas distribution, one for medium and low voltage and one for high voltage. The reason for this split was organizational: budgets were divided along these lines. The groups in the workshops consisted of the asset engineers working on the maintenance and investment plans, amended with field engineers, control engineers and so on to capture the operational risks. The participants were given a short introduction to the asset management process, the value system[165] and the rules in the workshop, and then were asked to write down the risks they perceived as most relevant on a post-it. The number of risk thus identified was astonishing. Each of the workshops easily produced well over a 50 risks in a 15 minute session. Precise numbers were: Gas 72 risks, MV/LV 168 risks, HV 122 risks. Entries in the workshop were not all of the same type. Specific events (truck crash into station) as well as conditions (deterioration component), stakeholders (pressure groups), materials (asbestos), designs, operational errors, decisions, lacking solutions, asset management principles, interventions and many others were mentioned. Some of the risks mentioned (marked with an !) were about the need-for and risks-of asset management. Apparently, the asset management initiative was not welcomed by everybody (this unfortunately will not be a surprise to any practitioner). In Table 32 an excerpt of the risks is given in alphabetical order (of the original Dutch, translation by author):

Table 32: Sample from risk workshop entries.

High voltage risks	Medium and low voltage risks	Gas Risks
< N-1, loss of redundancy, planned outage HV line for maintenance not allowed, spiral of decline	Truck crashes into station	Liability for damage as a consequence of interruptions
10 kV potential transformers explode	1 cable in street instead of both sides of the street	Asbestos cement piping
Defect in HV line grounding	Connecting 2 new customers where only 1 could be served	Maintenance backlog
Pressure groups	Shareholders do not provide budget	Internal invoicing between asset manager and service provider (benefit versus costs) (!)
Snapping earth wire in substation	Contractor ruptures cable	Valve inoperable (not closing, broken)
Deterioration of component	Noncompliance with 50V / 5s rule	Valves at station overgrown
Deviation in uniformity	HSE report disappears	Operational errors in MP grid
Archiving (decisions, agreements)	Asset manager and service provider communicate conflicting messages (!)	Lightning strike in station
Collapse HV tower	Operational safety grounder, voltage tester	Doors of station overgrown
Limiting maintenance (!)	Temperature monitoring MV/LV stations; sign of overload	1st generation PE (4 bar) in rural areas

The original purpose of the workshops was to assess the risks as well. Because the number of risks was much larger than expected, this was not feasible. Just discussing all of them, removing the clear duplicates and mapping them to the most relevant value at risk consumed all the time left (workshops were planned for 3 hours). The assessment of the risks was postponed to a new workshop. In preparation of this next workshop, the teams were asked to group the original entries into a manageable number.

---

[165] Referred to as the business values, the aspects of importance to the organization

Both the MV/LV team and the HV team grouped the risks on their gut feeling, to arrive at a list that in their perception facilitated them in addressing the most important issues. The gas team tried a more analytical approach. They tagged the entries either as Event, Direct Cause, Condition or Information Issue. This resulted in 63 items, out of the original 72. Apparently some more near duplicates could be removed. The events were used as the risks, and for each event the related Causes, conDitions and Information issues were identified. See Table 33 for an excerpt.

Table 33: Mapping of workshop entries to risk list after Essent Netwerk Noord (2002a). Column G holds the id of the risk, whereas columns C, D and I hold the number of the workshop entry mapped onto the new risks. It is clear some overlaps exist, as the new risks could be mapped to several entries, and some entries were linked to more than one risk.

G	Risk events	C1	C2	C3	D1	D2	D3	I1	I2
1	Overpressure	5	6		2	7	8	3	
2	Underpressure	2	6	10	3			3	
3	Explosion of station / fire	6	9		5	9		1	
4	Leakage by pipe rupture Low pressure, free outflow of gas	1	2		1	4		1	

For all teams records were kept on the entries that were combined into the resulting categories, which allowed them to demonstrate to the workshop participants that all inputs were used, building trust. In the assessment workshops, for all risks the likelihood and consequences were established, resulting in a risk level. The top ten of risks per group (i.e. the risks with the highest exposure[166]) were presented to top management in the form of the risk management statement. This resulting list is shown in Table 34.

Table 34: Resulting risk register approved by management, after Essent Netwerk Noord (2002b).

Rank	Gas	MV LV	HV
1	Interruption of supply	Explosion components	Touch voltage too high
2	Overpressure in the grid	Contact with live wires at works	Sag too large
3	Leakage by Medium Pressure pipe rupture, release of gas cloud	Noncompliance with environmenatl permit	Accident by switching error
4	Leakage by Low Pressure pipe, unrestricted outflow	Large interruption of supply (many customers)	Snapping of overhead line
5	Free outflow of gas into house	Long interruption of supply	Explosion of component
6	Accident at work	Noncompliance with regulations	Accident by induction current / voltage
7	Long response time at unsafe situation	Too low return on investment	Interruption of supply
8	Damaged relation with community (e.g. graafrust)	Damaged relation with community and customers	Reactive power shortage
9	Conflict with regulations	Overvoltage	Non-compliance with regulations
10	Leakage/fire after explosion of electrical joint	Undervoltage	Overloading components

---

[166] During the workshops, it became apparent that risk was a confusing concept. Many regarded it as probability times consequences, which is the expected outcome. The term risk level was used to categorize this (high, medium, low), later developing into exposure (see H11 capturing the value system)

For these risks it was mandatory to further analyze them to support the estimated exposure by means of a formal risk analysis document. A relevant part of analyzing the risk was creating a risk diagram with a group of people, in which the logic of the risk could be documented. The risk diagram was not a very formal approach, it borrowed elements from fault trees, event trees and cause and effect analysis to name a few. The purpose was to generate more insight into the risk, by relating the risk event to underlying causes, potential consequences and the conditions for both causes and consequences to materialize[167]. The risk diagram allowed many of the original entries of the risk workshop to be included into the analysis. A beneficial side effect was that the diagram almost naturally provoked questions on what else could cause the same effect, thus resulting in a much richer understanding of the risk.

An example diagram is shown in **Figure** 51. This is for risk 2 of the HV list, Sag[168] too large. This “risk” is marked by a red circle in the diagram. Starting from this point, the causes for this to occur were identified. For example, the sag could result from an extended conductor, caused by thermal expansion due to loads higher than designed for. This immediately invited other causes to be mentioned, as it could also be the result of higher ambient temperatures than designed for as a result of climate change. The discussion made apparent that it was not the sag in itself that was the problem, but the resulting ground clearance[169]. But an inadequate ground clearance could also result from changes at the ground level, like new constructions under the HV line, or perhaps by soil deposition. It could also be the result of changes in the norms regarding ground clearance (over the years, the distances regarded as safe have been increased) or a change of the activities below the line (not all use of the ground requires the same clearance). Another viewpoint that was expressed during the drawing of the diagram was that for most incidents with sagging conductors it would not have mattered if the sag would have been at the right level. The contact with the conductor would have been there anyway because the “vehicle” was way too high (trucks with an upright crane, ships with high masts), or people behaved dangerously (raising a 12 meter fishing rod whilst standing below a line, flying a kite in the vicinity of a line). This raised awareness that it was not the sag that was the problem (it actually is a state, not an event), but the contact with the conductor. The sag by itself was still a real problem if new buildings were (unjustly) allowed beneath the line so that in a later instance the clearance would have to be increased at the cost of the HV operator. The sag could also be of importance in liability claims, but those would in general result from some incident that occurred and thus would be a secondary risk. Besides, if raising the conductor truly would not reduce the number of incidents, raising them purely for liability reasons would be a bit cynical.

---

[167] This is in essence the same as applied in BOWTIE analysis, but that concept was not commonly known in the organization by then.

[168] Sag is the height difference between the lowest point of the overhead line and the mounting points.

[169] The distance between the lowest point of the line and the ground surface.

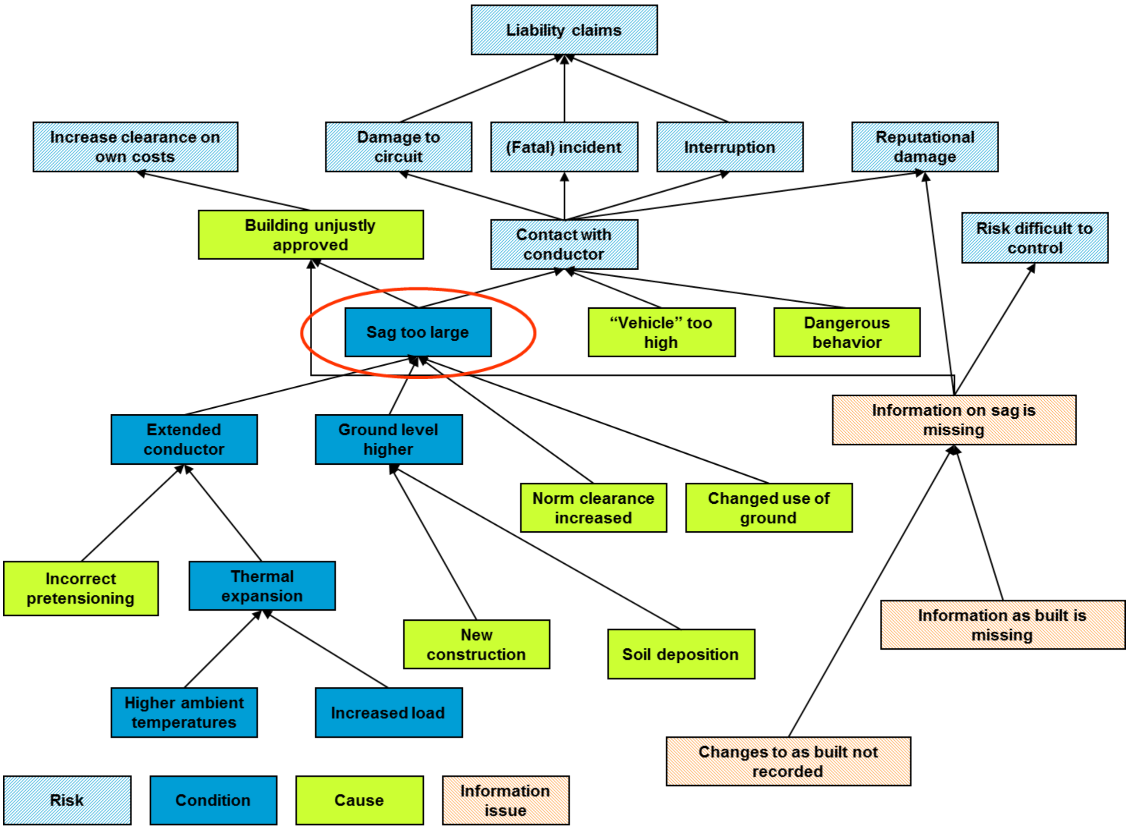


Figure 51: Example risk diagram for sagging high voltage lines.

The example clearly demonstrates how the process of drawing the diagram increases the insight about the risk and the options to mitigate it. Which type of diagram is used probably does not matter much, though more formalized approaches may facilitate communication of the results. On the other hand, the formalism might also hinder participants in sharing their knowledge on the risk because they do not understand how to frame their knowledge in the formalism.

However, that the risk diagram workshop should be aimed at getting most information out of the participants does not mean the participants are always right. Common errors for example were:

- the inclusion of interventions in the diagram (e.g. overdue maintenance)
- reversal of cause and effect relations (e.g. asset unsafe because not compliant)
- the inclusion of unspecified elements (e.g. climate change)

Each of these errors would result in a focus on a specific solution, which would not necessarily be the best. If the risk is overdue maintenance, the only solution is to maintain, whereas from a lifecycle perspective corrective replacement may be the better option. Framing safety as a result of complying with standards removes the awareness that even compliant assets can fail and that standards can be wrong. Using unspecified elements hinders quantification and thus a proper optimization of cost and risks. It has to be recognized that the label "error" is perhaps too harsh, as most errors contained information that increased the insight in the risk. But it required constant attention to reframe those errors into useful information, it was not something that happened naturally.

10.3.2 From batch to continuous

One of the disadvantages of running risk identification as a batch process is that newly encountered risks may have to wait to the next round to be processed. If there is a short cycle, this is not very hindering, but with longer cycles it really becomes a nuisance. As risk identification formed the starting point for intervention development, which provided the input for the budget decision for the following year, it also ran at an annual cycle. If a new risk was encountered during the year, it would have to wait until the next session. In the start-up phase that did not matter much, as the number of risks was much higher than could be dealt with (Wijnia, 2007). But with the growing number of risks that were analyzed, this question popped up more than once, especially regarding risks for which the intervention clearly was efficient (i.e. added value according to the value system much higher than the costs). As a first measure, a workaround was developed. For risks below the attention level[170] a shortcut of combining risk analysis, intervention development and planning was allowed. But even high risks could pop up during the year which would require immediate attention. A continuous process would be much more fitting to the situation. As the asset management processes would have to be redesigned in 2005 anyway to comply with PAS 55, this was used to make the change from batch to continuous.

Part of this continuous process would be establishing an “antenna” for permanently detecting risks[171]. The risks collected this way were entered manually into a Commercial Of The Shelf (COTS) application for maintaining a risk register for Enterprise Risk Management which allowed the use of several taxonomies in parallel for recording the risks[172]. The most important taxonomy from an asset risk management perspective proved to be the asset hierarchy[173]. The asset hierarchy is the second phase of the risk process developed to model interdependencies between risks (Wijnia and Herder, 2004). However, as the COTS application originated in project risk management and only later

---

[170] In Enexis, risks of the level High and above would require a separate analysis and formal decision of top management about the mitigation (which still could be accepting the risk). Risks below this attention level (i.e. Medium and below) could be accepted without approval of top management. The organization was willing to bear the risks in order to pursue its overall goal of efficient distribution of gas and electricity. In that sense the concept is quite close to the risk appetite often used in an investment context.

[171] The first attempt to this was to open a mailbox to which employees could post new risks. The team of risks analysts could then review them one by one. Implementation of this idea was almost overnight, and risks started flowing in, at a rate of a few per week. Managing this stream proved quite difficult. Most entries were not entirely clear to the analysts, and some more information or explanation was needed. As mailboxes go, requesting extra information is done by replying to the mail. But this meant multiple mails would be in the system about the same risk, requiring some structuring, like directories in the mailbox. Some new risks would be very much like previously mentioned risks, creating the dilemma whether to create a new directory or to place it in the existing one. Some mails would contain multiple risks, and then the choice had to be made where to place the risk, or perhaps to multiply the mail. The problems earlier encountered in building a risk taxonomy in the register now were replicated in the risk mailbox. But the biggest problem was that track was lost on the risks, resulting in some mails being unanswered for a very long time. In short, this was a risk managers worst nightmare. The mailbox system was therefore quickly abandoned and replaced by a more robust system with unique identifiers per notification, tracking, reporting and archiving. This was Infonet, an in-house developed service for information requests. In January 2009 this system was still running and, according to Korn and Veldman (2008), in 2008 it collected on average about 3 new risks per week.

[172] This is in line with the suggestion of Morgan.

[173] This is in line with the findings of In **MEASURING RISK** (chapter 8) and **WHOLE SYSTEM OPTIMIZATION** (chapter 9).

developed into ERM the use of the asset hierarchy was not allowed as a primary taxonomy. Therefore, the organizational structure was added as the primary key to indicate responsibility for managing the risks. The values, previously used as the top level of the taxonomy, now only appeared in the impact assessment, though it was possible to report on them over all the risks.

To ensure newly identified risks could be compared with already existing knowledge on risks, the decision was made to enter the full history into the application. At that time, early 2006, the volume was about 200 risks, of which over 100 full analyses. The risk identification antenna provided roughly one new risk per working day, adding up to 200 per year. Old risks had to be re-evaluated against the (then) actual risk matrix, as it changed in between[174]. It took until September 2006 until the full history was in the register. Risks that had a full analysis behind them were quite easy to re-evaluate, though it still was labor intensive because the needed information (the quantified impact and likelihood) was hidden inside a document. But if the risk only had its levels recorded without the underlying rationale, re-evaluating it against a new matrix essentially meant evaluating it as if it was a new risk. Additionally, the newly collected risk would have to be evaluated for the first time as well. Originally the plan was to evaluate all those risks with a team of experts in a risk evaluation meeting once every three weeks, but that quickly proved to be impracticable. In the evaluation meeting roughly 3 risks could be discussed (about 1 hour per risk) to establish the level. But in the same three week period 15 new risks would be collected. The process design thus created a backlog of 4 risks per week. To solve this problem the process was redesigned. Instead of evaluating the risk with the experts, the risk management team would provide a first estimate, which then would be validated with the experts[175]. The discussion then could focus on risk in the attention zone and risks for which the estimate was uncertain. Risks well below the attention level were not that important after all. The discussed risks were generally further analyzed. This redesign proved to be successful. The rate at which risks were estimated went up to about 7 per week, above the rate at which new risks were identified. That meant the backlog would in the end disappear. Estimation still required about 1 hour per risk, but now by a single person instead of a group of 10 people. Furthermore, the approach recognized that the accuracy of and the effort put into risk assessment are inversely proportional. All risks were provided with a rough estimate at a low cost per risk (1 person hour), to filter out those that clearly were irrelevant. The relevant risks would be discussed by a larger group, requiring more effort (about 10 person hours per risk) for a smaller set. The results were used for further filtering: only the most important of those discussed would be analyzed in full detail at a high cost per risk (about 100 hours per risk). By early 2007 all risks (some 450 at that time) had been quantified at least by an estimation.

10.4 Estimating the total risk

With all risks in the register being quantified, for the first time a quantitative estimate for the total risk in the system could be given. A straightforward way would be to add the expected values of the risks, though that would only give one number. Instead the risk register was used to provide a distribution of expected value, based on a Monte Carlo simulation with different combinations of risk materializing in the next year. The value of interest for this simulation was reliability, measured in

---

[174] The matrix only reached full maturity at the end of 2005, see the epilogue of chapter 11.

[175] This redesigned process is described in Korn and Veldman.

Customer Minutes Lost (CML)[176], the sum over outages of the number of minutes the customer were without service. To check the validity of the simulation, it was compared with the actual performance of Enexis in the past 6 years. This comparison is shown in Figure 52.

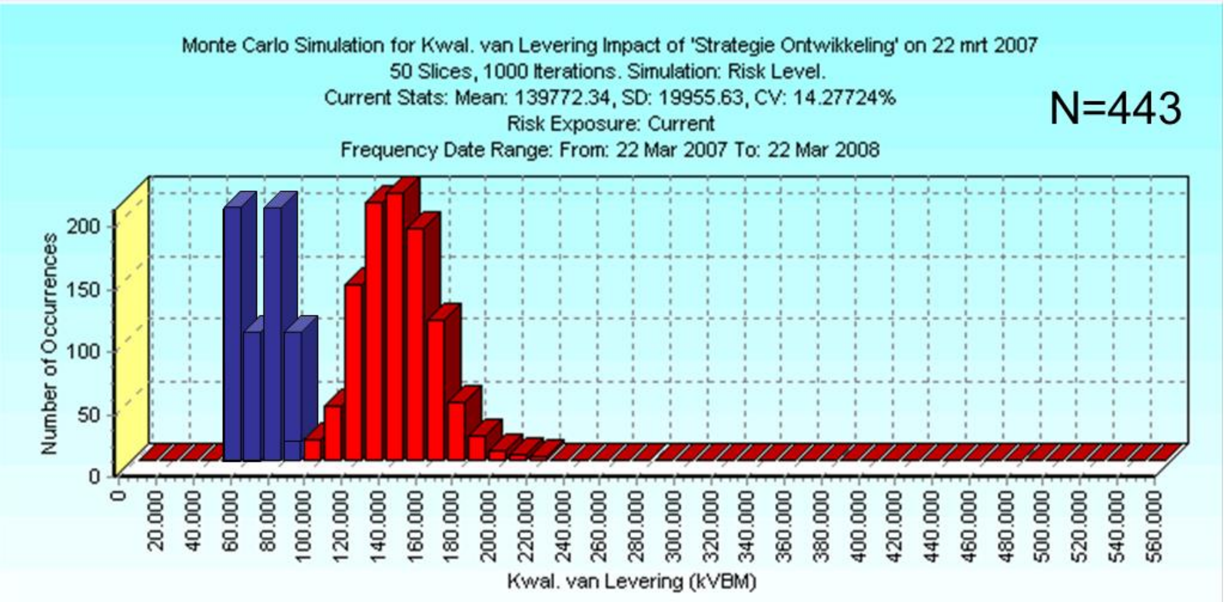


Figure 52: Actual total unavailability (purple, after Essent Netwerk (2007)) against predicted total unavailability for the 443 risks in the register (red, after Wijnia and Korn (2007)). The distributions do only slightly overlap.

Even though the spread in the predicted performance was reasonable compared with the spread in the actual performance, the average was way (more than 2 times) too high (also reported by Korn and Veldman (2008)). Apparently, the risk register contained more misery than reality.

10.5 Understanding the over completeness

According to the register the exposure was more than twice the amount of what was actually observed with regard to availability. Two potential explanations for this phenomenon popped up:

1. Estimating errors for impact and likelihood of risks in the register
2. Overlap between the risks, so that part of the exposure was double counted

Prime suspect was a faulty estimation process. Risk estimation is an order of magnitude activity, as the risk matrix scales by factors of ten. Being off by a factor 2 is still within the right order of magnitude and thus not very harmful in determining the risk level. However, that applies to individual estimates. The presented total is the sum of the exposures for some 450 risks. Whereas some estimates are too high, others will be too low. Random errors in the estimates thus cancel each other out, and the uncertainty in the sum is much smaller than the individual uncertainties. It is therefore virtually impossible that the mismatch is the result of random estimating errors. The only way errors could explain the result was a systematic overestimation of the exposure. It has to be admitted that in risk estimation workshops overestimation is the preferred option. It is better to

---

[176] A customer without service due to a failure for one hour is 60 customer minutes lost. If that same failure would impact 10 customers, the volume of customer minutes lost would be 600.

reduce the risk level after a proper analysis, than to be surprised by a risk that that was wrongfully neglected. But given the effort that was put in estimating the exposure that was not very likely either. Only a few risks had true estimates attached as the risk was really new in the sense that it had not yet materialized. But most were about events that did occur in the past. The estimate then was derived from the observed incidents, though sometimes assumptions had to be made on the fraction of the observed incidents that could be attributed to the risk at hand. So this better safe than sorry rule was not applied, at least not explicitly.

With regard to overlaps between risks the stance was less certain. Clear overlaps (two people mentioning the same risk in a workshop) had been removed in the process, but overlap resulting from using different aggregation levels still was present. The MV/LV risk list of Table 34 for example contains both “Non-compliance with regulations” as the item “Overvoltage”. But overvoltage is against a norm, and thus a specific form of non-compliance with regulations. However, given that only exceptions would be framed at a deeper level of the hierarchy, the effect on the total was relatively mild.

The true cause of the overestimation of the total misery was much more hidden. Risks can be framed in any phase of the process. The table of identified risks in this section clearly demonstrate that it is not only a possibility but that it also happened in reality[177]. For example, the HV list of Table 34 contains both interruption of supply as it does explosion of component, whereas the explosion may be the cause of the interruption. There thus may be a significant overlap between those risks, though that is not immediately apparent from the risk description itself.

10.6 Repairing the overlap

A potential solution for removing the overlap is framing all risks in the same phase of the risk process. This was for example done in determining the long term replacement strategy (Wijnia et al., 2006) or in measuring risk(Wijnia and Hermkens, 2006). However, even though the frames in both studies were centered around the asset, they were not precisely the same. Asset replacement used the asset phase, whereas measuring risk used a combination of cause and asset. Neither of the studies would have worked well using the frame of the other study. From a risk analysis perspective, the enforcement of a single frame is therefore not recommended.

But there is a more fundamental reason why such an enforcement should not be tried. This is because the framing of a risk determines its intervention[178]. An intervention in this view is an interruption of the risk process, as the materialization of misery requires propagation of the event along the full risk process. These interruptions can be placed anywhere in the risk process, both on the phases as between phases. Figure 53 contains some typical interruptions.

---

[177] This is not a problem unique for Enexis. The National risk assessement (Vlek, 2013) suffers from the same problem. The scenarios from DNRA 2011 contain Heavy Storm and Maritime Incident, whereas the heavy storm can be the cause for the maritime incident

[178] In practice it is often the other way around, as demonstrated. Risks are identified as missing mitigations, which are reformulated to become a real risk.

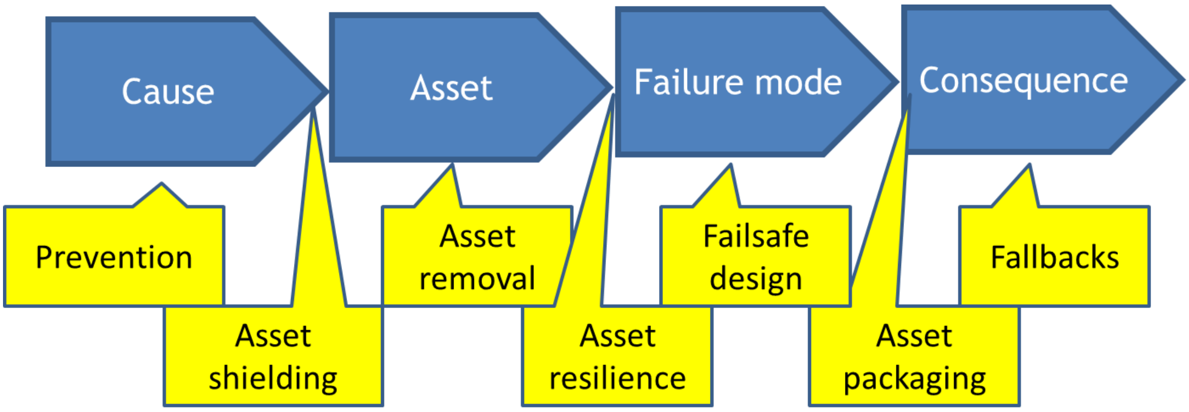


Figure 53: Some common interruptions of the risk process for assets.

If prevention is regarded as the best solution, risks typically will be framed at the start of the risk process. A typical example is the risk of excavation works damaging underground assets. The most It is possible to shield assets, or to strengthen them so that that could withstand the impact, but it seems better to prevent the damage in the first place. This is even more so true because there are many different assets in the underground and prevention would work for all of them, having the additional benefit of sharing the cost of the intervention over many asset owners. This is the reason that KLIC[179] (providing information to parties conducting excavations works) was a combined effort of all managers of underground assets. If the risk would have been framed in a different phase of the process (e.g. as the failure mode rupture) it is unlikely that a combined effort of asset owners would have resulted. But other preventions are possible as well, for example frequent inspection of the routes by helicopter as used for high pressure gas pipelines. If prevention is not possible the cables could be shielded, for example by placing a concrete slab on top of them. Given the cost of the slab and the inconvenience it causes for maintaining the underground, that would be only used for very important assets on individual routes, like some HV cables. But shielding an asset may not only protect against excavation works, it may also function against other potential causes. An option often overlooked is removal of the asset, but some assets (generally only a small fraction of the asset base) may cause more problems than they are worth. The overload risk of chapter 5 applies to redundant circuits, which are a form of asset resilience (though the circuit is then the asset, not the cable itself), combined with failsafe design (the fault is isolated before any further outages and damages occur). Underground cables are also packaged, as the violent behavior of failures (arcs, evaporating materials and cable movement) are confined by the ground around them. The policy to repair faults in redundant circuits within a certain period is about restoring resilience. But in case everything else fails, it may be convenient to have a fallback in the form of Emergency Power Supply.

If all risks in the register would be framed in the same phase of the risk register, only one type of solution would be possible as well. In limited studies that may be acceptable, but regarding the whole diversity and complexity of managing an infrastructure that clearly is absurd.

---

[179] KLIC = Kabels en Leidingen Informatie Centrum, translated as Cable and Pipeline Information Centre. The voluntary provision of information by means of KLIC was succeeded by a legal obligation to request the information, enforced by the Law for Information Exchange of Underground Networks (translated from Wet Informatie- uitwisseling ondergrondse netten (Rijksoverheid, 2008)).

The solution that was chosen instead was linking all risks including those that were not framed in the asset phase of the risk process to the assets that they impacted. The excavation works risk for example was linked to all underground assets in the hierarchy. To detect overlap, all risks per asset in the hierarchy were then reviewed. Risks linked at deeper levels in the hierarchy were rolled up to the examined level. This way overlap between risks defined at different aggregation levels became apparent. But overlaps between different phases in the risk process are much easier to detect in a risk list of some 20 items (the typical number of risks associated with the asset type level[180]) than in a list of almost 450 risks. As a first measure, risks that were fully covered by other risks to the asset were deactivated. This resulted in some 125 risks being deactivated.

Running the Monte Carlo simulation again over the reduced set (the "acknowledged risks") resulted in a prediction that was much more in line with the observations, as shown in Figure 54.

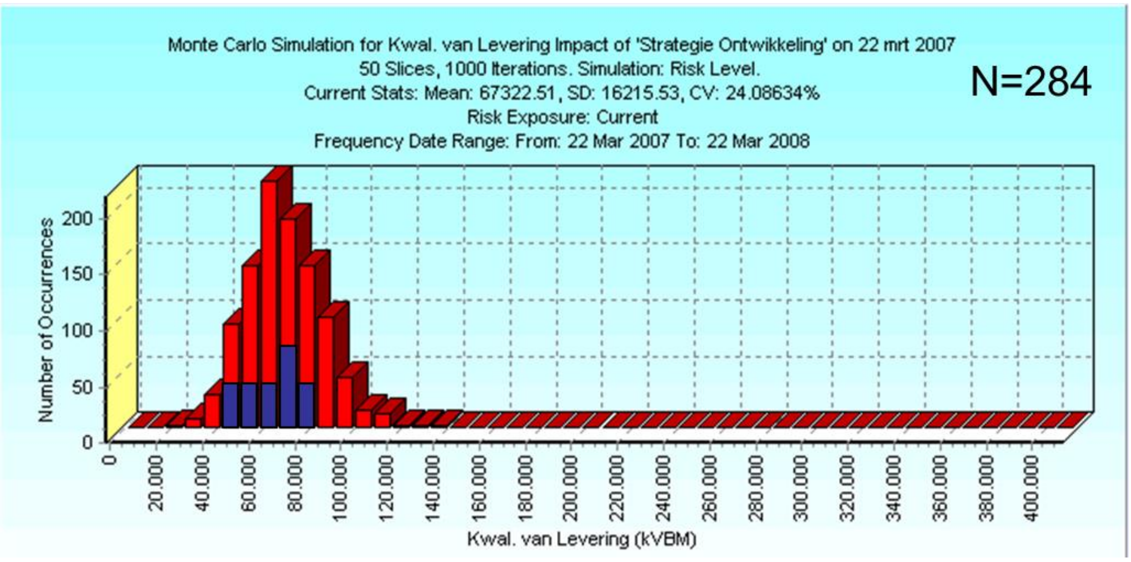


Figure 54: The actual versus predicted unavailability for the reduced set. The distributions clearly overlap, with peak values very close(Wijnia and Korn, 2007).

However, this is only a first order correction. The reduced set is not the definitive list of risks. As the number of risks in the register tends to be growing over time (Korn and Veldman (2008) reported over 800 risks), a more sophisticated method for repairing the overlap is needed, like adjusting the estimates for both overlapping risks to ensure mutual exclusiveness.

In a more fundamental approach, all risks could be defined at a deeper level of the asset hierarchy, as could be the causes, reactions and consequences. This would allow for resolving overlap at a more fundamental level. Moving down one level in the hierarchy typically results in 5-10 times more items, or several thousand for the Enexis case. In Reliability Centered Maintenance (RCM), those numbers are not exceptional, as every type of asset[181] needs its specific maintenance schedule. However, in

---

[180] As mentioned in the long term strategy, there are thousands of asset types in use. In the asset hierarchy only significant types (representing a population of more than e.g. 100 specimen or representing a replacement value of more than e.g. 1 million euro) would be included.

[181] Enexis has several thousand types of asset in operation, see long term optimization.

RCM all risks are framed into the same phase of the risk process: asset plus failure mode. Assuming independence between assets and failure modes, processing time only increases linearly with the number of items. But if that assumption does not hold, even RCM runs quickly into the limits of manageability. With interdependent assets, the processing time increases at least quadratic with the number of risks (that is even when only pairwise interactions are considered). Furthermore, the methods are aimed in general at maintenance, which is only one type of intervention. Given those limitations it is difficult to see how that would work in the practice of risk based asset management in an infrastructure. Assets are networked and thus can be highly interdependent, and the risk register is used for many more interventions than only maintenance. Given that the 450 or so risks in the register took almost a year to process, doing so for a much larger number really would stress the organization. It clearly violates the "Few enough" criterion for a list of risks.

10.7 Conclusions

Risk identification is a subject that receives little attention in literature. This may suggest that it is just a simple activity not worth any conscious thought. In practice, however, a risk register that is both comprehensive and mutually exclusive proved very difficult to develop. This is problem is not unique for Enexis. The problem is also not caused by a lack of competence in risk identification. Instead, the problem is caused by the multi-facetted nature of risk. Risk can be defined in any phase of the risk process. Only if risks are all defined in the same phase it is possible to create a comprehensive and mutually exclusive list. But from a risk management perspective that would not be desirable. The phase in which the risks are framed limits efficient solutions to the same phase. In order to have rich set of interventions to choose from it is vital to have risks defined along the whole chain. Using a single hierarchy (in this case the asset hierarchy) as a reference system helps identifying overlap between risks. By eliminating full overlaps along the risk process, and adjusting/reframing partially overlapping risks a set of more or less independent risks can be created. For maintaining this list it is vital that both the resolved and residual overlap is made explicit, so that in case the addition of a new risk would require the adjustment of an existing risk, the potentially other impacted risks can be identified. Given the experience at Enexis maintaining the risk register is best done in a continuous process. Every new risk has to be checked against the existing entries by itself. If risks would be collected in a batch like setting, it only would result in a stack of items waiting to be processed individually. As the processing time per new risk increases with the number of items in the register, it is advisable to keep the number as low as reasonably possible. In practice, a few hundred risks proved to be manageable. However, even at this number, risks are still defined at a relatively aggregated level. Putting more detail in the risks by moving one level down in the hierarchy easily results in the number increasing by a tenfold. That would almost certainly be beyond any organizations capability. Only by framing all risks into the same phase such numbers are manageable, as then independence between risks is relatively easy to assure. But the result of such an exercise would only be useful in a specific context, and not as a generic base from which all policies and projects would be derived.

A final remark must be made with regard to comprehensiveness and mutual exclusiveness. Whether a risk register is complete can be measured by the number of surprises[182], which are major incidents

---

[182] This measure was advocated by John Woodhouse in the infrastructure exchange 2005.

whose potential of occurrence is not recognized in the risk register[183]. Determining whether the risks are mutually exclusive is much harder, especially at large numbers spread out over the risk process. The only quantitative test can be made against actual performance, but that is an uncertain number by itself. Given the choice between comprehensiveness and mutual exclusivity, a risk manager should probably opt for comprehensiveness. Risk management is more about preventing surprises than about getting the numbers right with high accuracy. In many cases order of magnitude accuracy is good enough. Furthermore, overlap between risks can also be detected when interventions are considered, e.g. if two interventions include the same asset. This opportunity is not present if a risk is missed in the identification in the first place.

---

[183] This is not about the estimate for the likelihood being wrong, but about the whole event missing.

11 Capturing the value system

This chapter was published before as **ASSET RISK MANAGEMENT: ISSUES IN THE DESIGN AND USE OF THE RISK MATRIX (WIJNIA, 2012)**[184].

11.1 Introduction

The management of risk within the profession of asset management has gained in importance over the years as it has been advocated by numerous relevant parties and sources. The Infrastructure Management Manual (IAM, 2002), mentions risk management for example in its definition of advanced asset management (suggesting at least that it is the way forward), PAS 55 (Institute of Asset Management, 2008) includes the risks of assets in its definition of optimal management. In parallel to this growing significance within asset management, general risk management (or Enterprise Risk Management) has gained in importance as well, no doubt influenced by the Sarbanes Oxley act in the USA. Within risk management, an essential stage is assessing the risk level. As risk is a two dimensional entity (often referred to as risk equals impact times probability), assessing the risk level means assessing both impact and probability and judging their combination against some risk criteria. A tool often used in this exercise is the risk matrix. This is a table which has several categories for probability on one axis and several categories for impact on the other axis, and a level indication per cell, often both in description (low, medium or high) as in color (green or red risks). As such a risk matrix is an easy to use tool, many practitioners have adopted it and there are few asset managers around that do not have a risk matrix.

However, despite its widespread use, there is surprisingly little standardization of risk matrices across industries, nor is there a solid framework for constructing a risk matrix. For example, the ISO 31000 standard (ISO, 2009a) on risk management does not even mention risk matrices. The accompanying document on risk assessment techniques (ISO, 2009b) mentions risk matrices in its appendices but does not really go beyond mentioning that the format depends on the context and that it is important that an appropriate design is used. A few design options are considered, and a list of limitations with using risk matrices is given, but again little guidance on how to do it correctly. This lack of guidance seems to be typical for any publication on risk matrices. As a result, practitioners are left to themselves in developing the risk matrix they are using.

From one perspective, this is a good thing. A risk matrix is an expression of the value system (including risk attitude) of the organization that uses it, and thus is it at least specific to the organization. However, it is a bad thing in a different perspective. First of all, even though organizations have different value systems, are they really that different? Organizations operate in the same society, and their attitude with regard to financial losses for example may be more aligned than many think as they may have (in the end) the same shareholders. Additionally, there is a market for insurance, which to an extend is a standardized product and thus presumes comparable risk attitudes. Furthermore, as all kinds of regulations limit the risks organizations may impose on others (like regulations on external safety, operational health and safety, emissions) one might expect (or

---

[184] Wijnia, Y. 2012. Asset Risk Management: Issues in the Design and Use of the Risk Matrix. *In:* Mathew, J., Ma, L., Tan, A., Weijnen, M. & Lee, J. (eds.) *Engineering Asset Management and Infrastructure Sustainability*. Springer London.

hope?) that their attitude with regard to those values is not that different from what society thinks is acceptable.

Secondly, the lack of well-defined procedures for establishing the risk matrix may mean that the practitioners get it wrong. That is, the resulting risk matrix does not capture the value system of the organization, and decisions based on the matrix are not perceived as good decisions.

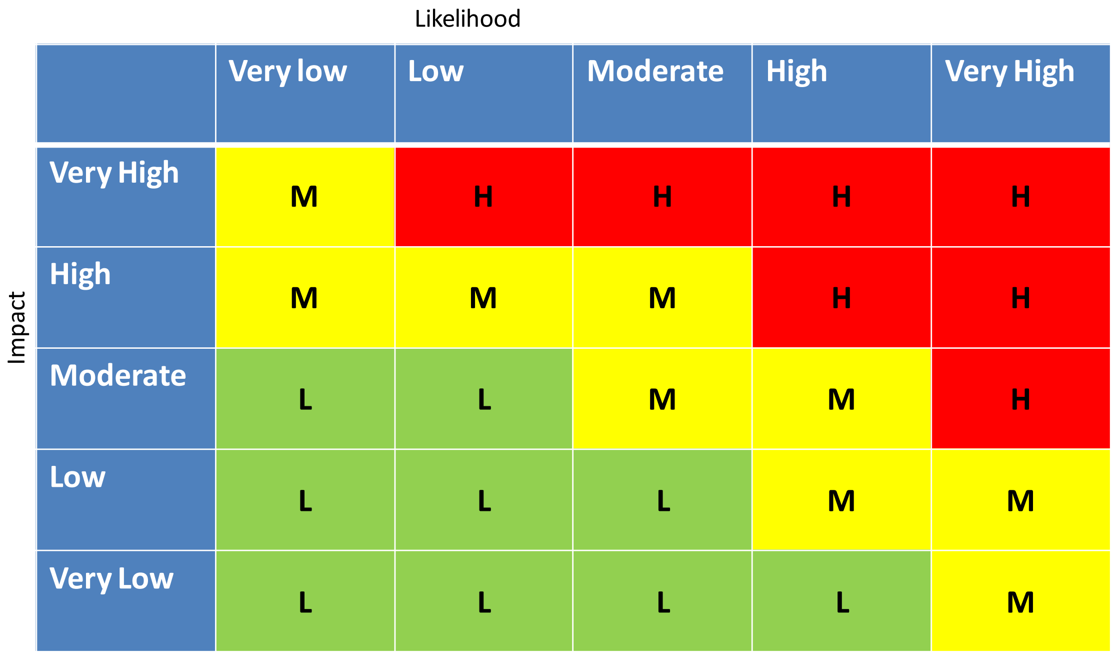
With regard to the first point one might wonder why it is a problem. Some may even argue that going through the process of defining the risk matrix is vital for creating ownership of the matrix and thus building a risk management culture, even though the end result might be achieved by copying a risk matrix from another organization. Besides, the effort required for creating a risk matrix is not that big that it has a catastrophic (or even noticeable) impact on the organization. However, it is serious because of the second point. If there are no standard procedures, and no template to start with, it may be a stroke of good fortune to get it right. As a result, one might expect many decisions based on those risk matrices to be wrong.

This paper deals with the issues in designing and using a risk matrix in asset risk management. First, the basic concepts of risk management will be defined, including the definition of risk, the risk matrix and the risk management process. This is followed by a section on the design decisions to be made when developing a risk matrix. Several options will be discussed, illustrated with some practical examples and real life experiences. The next section is on problems in and criticisms on the use of risk matrices, including a discussion on their validity. This results in a review of the risk management process. The paper ends with conclusions on the value of a risk matrix in asset management decision making.

11.2 Basic concepts

Risk has many meanings in everyday use. For example, Beer and Ziolkowski (1995) mention 13 different definitions. Most significant difference amongst these definitions is that between risk as the entity (the threat, danger or event with a potential for undesired effects) and risk as a measure of bad fortune, often defined as risk = impact *probability. To prevent ambiguity, within the context of this paper **risk** refers to the entity “event” and **exposure** to the expected bad fortune. As mentioned, the exposure is related to the probability. However, in a strict mathematical approach probabilities cannot be larger than 1, whereas within asset risk management risks may materialize more than once. In such an operational risk management environment, it is better to make use of the term frequency, like in “this asset fails twice a year”. Frequencies, on the other hand, lose meaning for unique or very rare events and it is better to use the term probability, even though numerically frequency and probability are virtually equal for small numbers. To avoid referring to probability/frequency the more general term of likelihood will be used in this paper, referring to both concepts. The formula then becomes exposure = impact*likelihood.

The risk matrix is a table used to characterize risks (according to the definitions above, is to characterize the exposure of risks). A risk matrix is a table with impacts on one axis and the likelihood on the other. Figure 55 is a typical example. The color coding and characters represents the risk level, with red for high risks (H), yellow for medium risks (M) and green for low risks (L). This classification (including the names to specify categories) is the expression of the value system of the decision maker and thus subjective. In the example, risk levels more or less comply with the exposure, but this does not need to be the case. Risk levels may depend solely on the impact (resulting in horizontal bands) or solely on the likelihood (resulting in vertical bands).



Likelihood

Impact

	Very low	Low	Moderate	High	Very High
Very High	M	H	H	H	H
High	M	M	M	H	H
Moderate	L	L	M	M	H
Low	L	L	L	M	M
Very Low	L	L	L	L	M

Figure 55: Example risk matrix.

Within the field of risk management, different views exist with regard to the objectivity of risk assessment (Klinke and Renn, 2002). Some claim the exposure can be quantified objectively, as if risks are out there to be measured. On the other hand, many risks to be assessed have not yet materialized (like what is the exposure of the risk Nuclear War?) and quantification is based on models that are inherently assumption laden and thus subjective. Furthermore, there is ample of evidence that risks are not always judged on the exposure (Slovic, 1987). This is for example visible in the example risk matrix of Figure 55, where Very High impact Low likelihood risks are judged as High/red and Low impact Very High likelihood risks as Medium/yellow. Within this paper the objective approach is followed.

With regard to the risk management process, there is no standard terminology. For example, the Risk management standard (The Institute of Risk Management (IRM) et al., 2002) follows the ISO terminology and uses risk assessment as the bundling of Risk identification, Risk analysis and Risk evaluation, whereas the COSO framework (COSO, 2004a) separates event identification from risk assessment (a bundling of risk analysis and risk evaluation). The steps are thus very similar, even though the terminology differs. Within this paper, the ISO terminology is used, as it recognizes risk evaluation (where the risk matrix is used) as a separate step. Figure 56 shows the risk management process (after The Institute of Risk Management (IRM) et al. (2002)).

Figure 56: Risk management process.

Within this process, the risk matrix is defined in step 1, and applied in step 4. It should be stressed that the risk management process does not require the use of a risk matrix, it only requires that some criteria are defined against which the risk tolerance can be judged.

The final part of this section is on the purpose of risk management and the risk matrix. In short, the purpose of risk management is to limit the exposure caused by risks that could have been efficiently mitigated. As there are many potential threats, the risk matrix is used to prioritize attention (Cox, 2008). Within this perspective, an incorrect risk matrix is a matrix that does not prioritize attention correctly.

11.3 Developing the risk matrix

The risk matrix, as shown in Figure 55, is a conceptually simple tool, just a color coded table with likelihood on one axis and impact on the other. However, there are surprisingly many decisions that have to be made to make a matrix. In this section these will be treated.

11.3.1 Orientation of the matrix

There are three decisions to be made with regard to the orientation of the matrix.

1. What axis is used for likelihood and what axis is used for impact?
2. What is the direction of the horizontal axis?
3. What is the direction of the vertical axis?

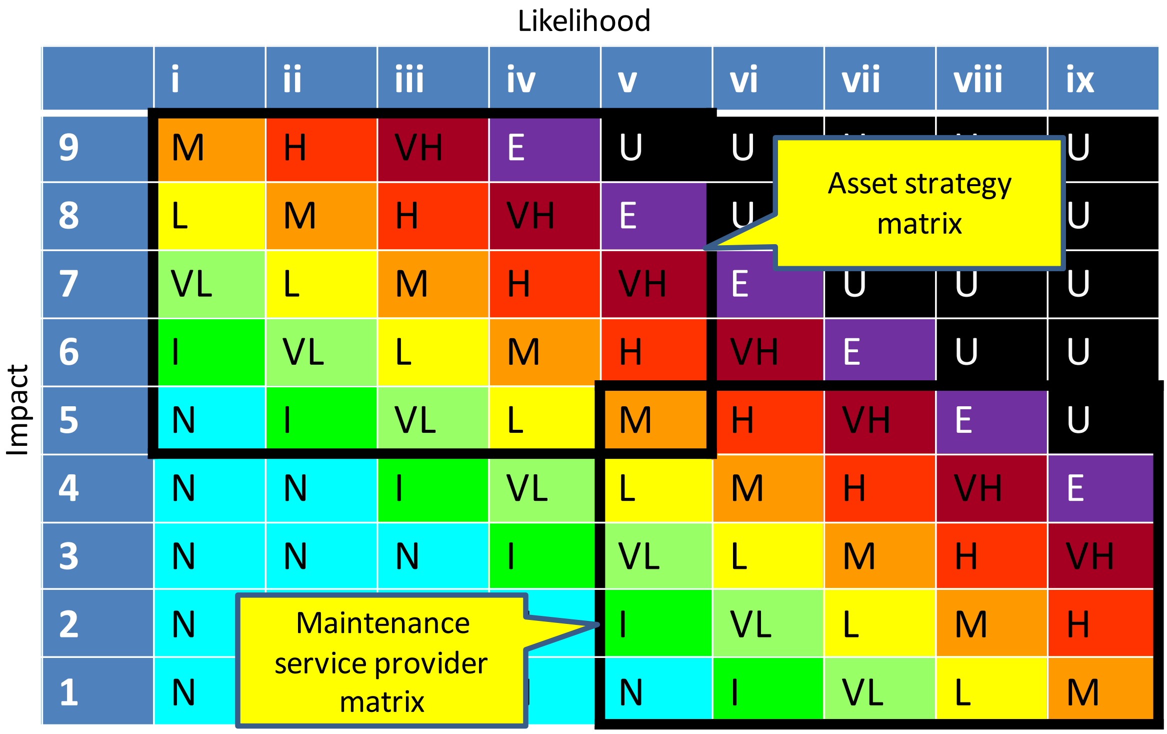
Each of these issues has two options, resulting in a total of 8 different orientations. In many sources (e.g. (Klinke and Renn, 2002)) impacts are on the horizontal axis and likelihood on the vertical, with the most important risks in the top right corner, but it is no standard. The example risk matrix of Figure 55 has the axis reversed, and Cox (2008) uses both orientations in his paper. Furthermore, if risks are mapped together with opportunities, it makes sense to use likelihood on the vertical axis, have the risks on the left hand side (impacts increase from right to left), and opportunities on the right hand side of the vertical axis (potential benefits increase from left to right). Most important risks would then be in the top left corner. If the likelihood is on the horizontal, one could make a case for the impact scale increasing top to bottom having the most severe (negative) impacts at the lower end of the matrix, though this does not seem to be common practice.

With regard to prioritization, the orientation is not relevant. High risks are more important than low risks, no matter where they are plotted in the matrix. In terms of communication it could be important, though. If risks are usually plotted in the example matrix (high risks in the top right corner), then reversing the directions of the axis would result in the most important risks being in the bottom left corner and decision makers could get a wrong impression of the importance of a risk. Key with the orientation of risk matrices is therefore consistency.

However, this is not all than can be said about the orientation of risk matrices. First of all, plotting risks and opportunities in the same diagram is a bad idea. People judge opportunities completely different from risks, to the extent that a risk formulated as an opportunity could result in radically different decisions (see the Asian disease problem (Tversky and Kahneman, 1981)). A risk matrix should thus really be a risk only matrix, and the most important risks could be plotted top right (apparently the most natural position). The second issue regards the choice of the axis. Although it is very common to plot the impacts on the horizontal axis, there are practical arguments against it. In general, there are multiple values that could be affected, whereas there is only one likelihood scale. Thus, if impacts are on the vertical and likelihood on the horizontal, one could plot the effect category description per value besides the matrix on a single landscape sheet of paper. This is easier to read than a portrait orientation, with impact descriptions above the matrix, and thus more likely to be used. Multiple page risk matrix documents (for example, the matrix, impact descriptions and likelihood descriptions all in separate tables on different pages) are a nuisance in practice and not likely to be used.

11.3.2 Dimension and resolution of the matrix

The dimension of the matrix is the number of impact bands, likelihood bands and risk levels to be used. The number of risk levels is often limited to 3: low (L), medium (M) and high (H). Low risks are then acceptable without further notice, high risks require action and medium risks should be considered for action. These three levels could be combined with any numbers of impact and likelihood bands, though 3 by 3, 4 by 4 and 5 by 5 seem the most common. There is no necessity to have equal numbers in rows and columns. However, thinking about the resolution, how much sense does it make to distinguish 9 different likelihood bands (extreme example) if it results in only 3 different risk levels? The same holds for impact bands. One could expect the risk level to increase if the impact moves along the categorizations for the same likelihood level, or vice versa. The key question therefore is what the resolution is with regard to the potential impacts. If there is a meaningful distinction in the severity of impacts, it should be reflected in the risk levels. This will be illustrated with the impact categorization of safety.



	i	ii	iii	iv	v	vi	vii	viii	ix
9	M	H	VH	E	U	U			U
8	L	M	H	VH	E	U			U
7	VL	L	M	H	VH	E	U	U	U
6	I	VL	L	M	H	VH	E	U	U
5	N	I	VL	L	M	H	VH	E	U
4	N	N	I	VL	L	M	H	VH	E
3	N	N	N	I	VL	L	M	H	VH
2	N				I	VL	L	M	H
1	N				N	I	VL	L	M

Figure 57: Overall risk matrix split into 2 sub matrices for different departments. Coding: Unacceptable, Extreme, Very High, High, Medium, Low, Very Low, Insignificant and Negligible.

Suppose an aircraft operator has several types of aircrafts, ranging from 747s to cargo carriers. Probably the worst event that could happen would be the crash of a plane. Would it, with regard to the categorization of the event, matter how many people would die in the crash? Most likely it would. Two 747s crashing into each other (some 1000 fatalities) would be worse than the crash of a regional plane (100 fatalities), which would be worse than the crash of a cargo plane (10 fatalities). But would it matter if a plane was only used at half the capacity? That would most likely be more difficult to judge. The scale continues downward. The next level would for example be a mechanic slipping during maintenance and breaking his neck (single fatality), followed by serious injury (lost time incidents with extended absence), lost time incidents, near misses or first aid, and dangerous situations. With regard to the top event, anything without fatalities could be considered insignificant, but in safety management the key control is the number of unsafe situations (iceberg theory of Heinrich (Heinrich, 1931)). Thus, proper risk management would require the impact bands to span all levels, which would mean 8 levels. A similar exercise could be held with regard to the likelihood axis. Unsafe situations might occur 1.000s of times per year, but aircraft crashes are generally below once a year but probably above once every 100 years (few airline operators will be without any crash). Two planes crashing into each other would be extremely rare, perhaps below once every 1.000 or even 10.000 years (the latter figure means that less than 1% of the operators would experience such a crash in 100 years). This would mean 8 or 9 orders of magnitude in the likelihood range, which is quite similar to the number of impact bands (presuming the orders of magnitude would be used as likelihood bands). However, using a 8 by 8 matrix may be not really necessary. Risk characterization is essentially a decision driven activity (Stern et al., 1996), and it is not likely decisions regarding small impact high likelihood (bottom right) would end up on the same table as the high impact low probability risks. The risk matrix could be split into two aligned sub matrices, one for the (internal) maintenance service provider and one for the asset strategy department. This is shown in Figure 57. However, care should be taken that the sub matrix encloses all relevant risks. Otherwise mischaracterizations could occur. For example, if the service provider matrix was used to classify an Impact 9/ Likelihood v risk it would result in a Medium (M) risk, whereas the full matrix would characterize it as Unacceptable (U).

A key lesson in of the example above is the extreme spread (many orders of magnitude) in impact and likelihood that can occur within risk assessment. This represents a “challenge” in visual representation, as it requires a logarithmic scale to be meaningful. Using linear scales would mean that only a fraction of the risks could be distinguished from each other. For example, if the impacts of the aircraft safety would be plotted on a linear scale, the most extreme event (1000 fatalities) would be plotted on axis’ end (either up or right), and the regional crash (100 fatalities) would already be at 10% of the origin. Anything below 10 fatalities would be compressed into 1% of the scale and thus would by no means be distinguishable. The same holds for the likelihood. Using logarithmic scales is not exactly the discovery of a lifetime (reason why challenge is put between quotation marks), yet some matrices neglect this knowledge. Cox (Cox, 2008) for example bases his arguments why risk matrices are wrong on a 5 by 5 matrix with linear scales for impact and likelihood. Others are more or less logarithmic but not constant in their scaling factor between categories.

Additionally, linear scales have a problem in the margins of uncertainty within the cells. Suppose a 5 by 5 matrix with linear scales. The top cell ranges from 0,8 to 1. A risk put in the middle of this category would have an accuracy of about 10%: 0,9 ±0,1. But a risk put in the lowest category would have an uncertainty of 100%: 0,1 ±0,1. As uncertainty is a central element of risk it would be surprising if an assessment could be made within a 10% margin of error. Logarithmic scales are constant in their margin of error. If a risk is in the middle of any band, it could be threefold higher or lower before it would be categorized differently.

Summarizing, the resolution of a risk matrix should be on the order of magnitude level (=a logarithmic scale), and the dimension should fit the orders of magnitude between the most and least serious impact. If the scaling factors are constant, the top left bottom right diagonals are levels with equal exposure, and should be used for the risk classification, as has been done in figure 3.

11.3.3 Scaling multiple values

Risk matrices often express the risk tolerance on multiple values. Risk matrices require the values to be translated into Key Risk Indicators (KRIs). For some values, like financials, the KRI is quantitative. For other values, like reputation, the KRI may be qualitative or descriptive. Some KRIs mix qualitative and quantitative elements, like safety in the airline example. For fatalities the KRI could be made quantitative (number of fatalities, though the events are still described qualitatively), whereas the non-fatal incidents are qualitative (though there is a trade-off between the number of people and the severity of the accident).

Developing an impact scale for a single quantitative KRI is not that difficult. The orders of magnitude approach should be used, so that the exposure in each cell with the same risk level is equal. Thus, if the likelihood categories are once every 100 years, once every 10 years, once per year, corresponding financial impact categories would be 1 million damage, 100k damage, 10k damage (or any other value scaled similarly). For practical purposes, quantities in a risk matrix are often not represented by single numbers but by bands, like 1 to 10 times per year, damage between 10k and 100k.

However, for qualitative scales things are more complicated. The problem is not in defining categories, but to assure they are an order of magnitude apart. For the safety example, one could distinguish the loss of a left hand, right hand, arm, leg, eye, ear, burns on the body, hands, face, loss of eyesight, hearing, and so on. Insurance companies may do this in order to determine the compensation fee. However, are they distinctive in terms of a risk matrix? Is the loss of a right leg comparable in severity as ten incidents where people lose their left hand? This probably is not the case, dismemberments (or amputations) are comparable in severity, though it is a difficult subject. Fortunately, the severity and likelihood of incidents are negatively correlated. For every 10 first aid incidents there is one lost time incident (Heinrich, 1931, Wijnia and Hermkens, 2006). Therefore, categorizing the impact according to this iceberg theory would result in equal risk levels at every impact level of the iceberg.

After developing impact schemes per value, the next step is alignment. Theoretically it is possible to develop a risk matrix for every value separately (possibly with different orientations, dimensions, resolutions and classifications), but that would be very impracticable and error prone, as risk matrices could get mixed up. Key question with regard to alignment is the comparability of the impact levels: is the most serious impact on finance really as serious as the comparable impact on safety? In other words, if the organization has two choose between two incidents on the same impact level, would it really be indifferent? Another approach is that of substitution, though it generally only works when comparing to the financial value. The question then is whether the organization is willing to spend the equivalent financial amount to prevent the impact. As precise valuation is difficult, this can be replaced by two questions: is the organization willing to spend (without a doubt) the financial amount one impact category lower? And would the organisation accept the impact if the costs would be one impact category higher? However, in this comparison one has to be aware that valuation is highly imprecise, and that the willingness to pay is different from the willingness to accept (the amount that has to be spend to prevent the incident versus the amount that has to be gained in order to accept the incident). Nevertheless, it provides reasonable guidance.

11.4 Problems and criticisms

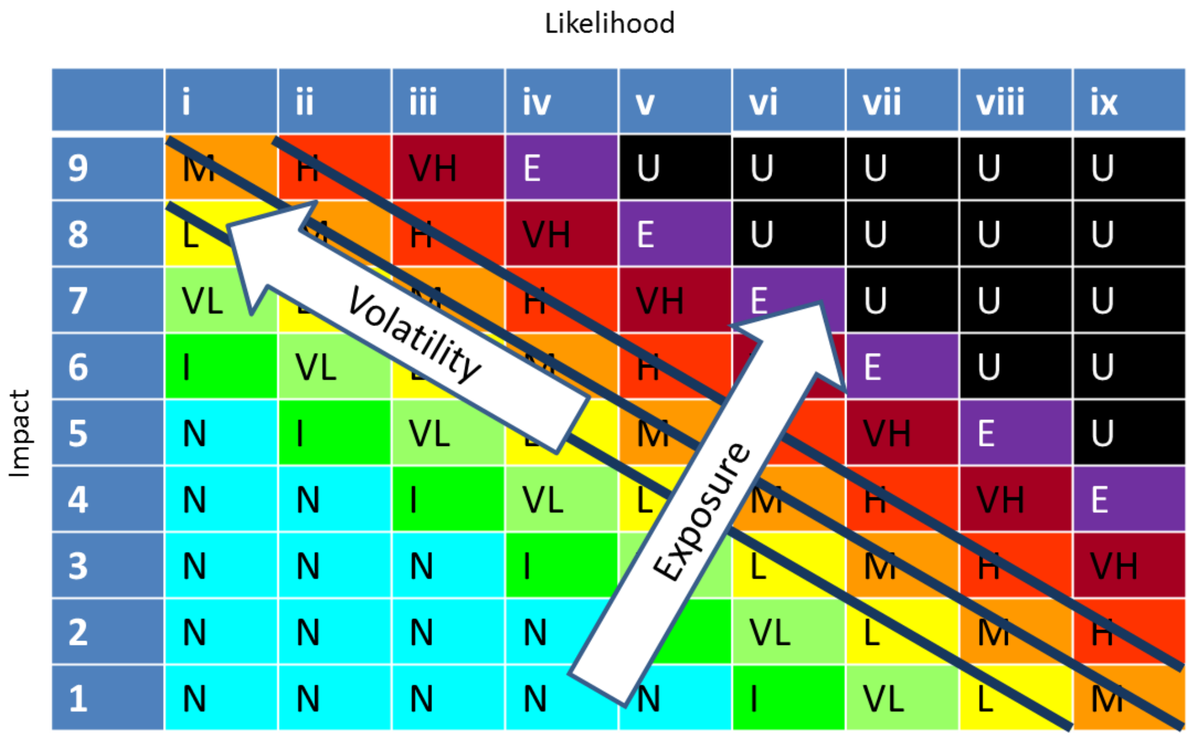
There are several criticisms on and problems in the use of risk matrices. They can be organized along three lines:

1. the risk matrix does not prioritize at all
2. the risk matrix does not prioritize risks correctly
3. the risk matrix results in incomprehensible decisions

11.4.1 The risk matrix fails to prioritize the risks

Typically, the distribution over the risk levels is not uniform: it is very rare to find unacceptable risk, high risks are more frequent, medium risks tend to be the most numerous and the numbers start declining below medium again. The reason unacceptable risks are rare to find is that companies that do not manage them are more likely to go out of business. The same holds for high risks: a company does not necessarily need to have formal risk management in place to understand it has a problem that is worth mitigating. For risks below medium another mechanism is active. People may feel the risk is not that important, and simply do not bother to mention it. This is especially true in risk identification workshop settings, where the aim is to capture the most relevant risks in a limited timeframe.

As a result, many risks may end up on the same level or even in the same cell. The 9 by 9 matrix of Figure 57 and Figure 58 only contains 81 cells and risk registers easily contain 1000s of risks (Morgan et al., 2000, Wijnia and Herder, 2004, Korn and Veldman, 2008). It is therefore inevitable that some levels and even cells contain many risks. This could be regarded as a problem for prioritization, but that is only the case if the risk matrix is applied very rigidly. First of all, if the exposure level is equal, the risks in cells with higher impact/lower probability deserve more attention. This is simply because unforeseen high impacts cause more disturbance, i.e. those risks are more volatile. It is like comparing losing a euro a day (e.g. an unused subscription) to a 1000 euro damage (e.g. broken washing machine) once every 3 years. Furthermore, the logic of prioritization between risks in different cells can also be used for prioritizing risks within a cell, provided that sufficient information is available to make such differentiation. Risks with a high exposure (level) are more important than risks with a low exposure, and for risks with the same exposure (level) the more volatile risk is more important.



	i	ii	iii	iv	v	vi	vii	viii	ix
9	M	H	VH	E	U	U	U	U	U
8	L	M	H	VH	E	U	U	U	U
7	VL	L	M	H	VH	E	U	U	U
6	I	VL	L	M	H	[illegible]	E	U	U
5	N	I	VL	L	M	[illegible]	VH	E	U
4	N	N	I	VL	L	M	H	VH	E
3	N	N	N	I	[illegible]	L	M	H	VH
2	N	N	N	N	[illegible]	VL	L	M	H
1	N	N	N	N	N	I	VL	L	M

Figure 58: Risk matrix with equal-exposure lines (blue lines) within an exposure level (medium), direction of increasing exposure and direction of increasing volatility for the exposure level. Exposure is the primary priority, but within a class of equal exposure the volatility can be regarded as the secondary priority. These priorities apply both for comparing different cells as for comparing risks within a cell. However, an intra cell comparison requires much accurate information. Coding: Unacceptable, Extreme, Very High, High, Medium, Low, Very Low, Insignificant and Negligible.

Another cause for failing to prioritize clearly distinguishable risks is that some risks may impact only one value, whereas others impact more values. Compare the risk of a large customer bankruptcy (generally only financial impact) with the risk of in-situ explosions, resulting in injuries, repair costs, environmental impacts and reputational damage. Assuming all impacts and likelihoods are comparable, the risks would end up in the same cell. In human judgment though the explosion risk would be considered more serious because the total impact is larger, but the matrix is not able to differentiate single value impact risks from multi value impacts, and there are no rules to add risk levels. The only way around it is to express all impacts in a single unit, multiply them by their likelihood and add the expected values to arrive at the total annually expected amount of misery. This is a continuous scale that thus has a much better resolution than the risk matrix, which is discrete. The unit in which impacts are expressed could be dimensionless, but as the impacts scales have been aligned with the financial scale, expressing everything in monetary equivalents is a practical solution. Besides, if the alignment is based on the willingness to pay, the monetary equivalent is also an indication of the maximum amount that could be spent on mitigating the risk.

The final cause for the risk matrix failing to prioritize risks is that it is used for the wrong type of risk. An example is the system performance risk, like the number of outages that occurs per year, the number of delays, the total financial loss and so on. But these are variations around some mean. Even though this variation around a mean corresponds to risk in a different context (the Capital Asset Pricing Method (Brealey and Myers, 2000)), that is not the context where risk matrices are used. Risk matrices exist to characterize event risks, which can be expressed and measured in terms of impact and likelihood. System level performances cannot be expressed in likelihood, as they have a fixed timeframe (in many cases a year). The risk matrix is thus compressed into a single column. Furthermore, system level performances typically have variances measured on a percentage scale, not on an order of magnitude scale. Thus, if a system performance is plotted in the matrix, it will always end up in the same cell, year after year. Besides, one should not be surprised to find an high level system performance in the intolerable area.

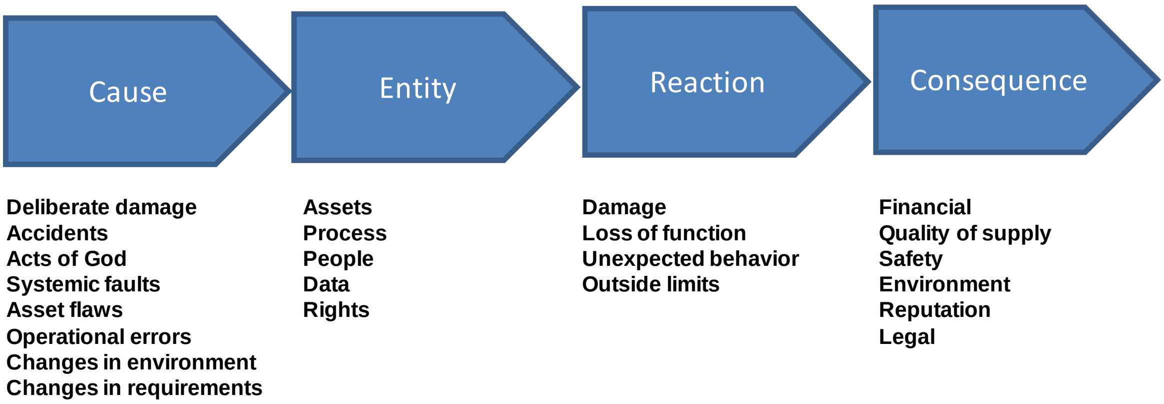


Figure 59: The risk process, based on Wijnia and Herder (2004).

A typical example is the total costs which could be easily 80% of turnover. If an event would occur that would cost that much it would no doubt be in the highest impact category, and if it would occur every year it would be intolerable (in fact, the company would go bankrupt). But that is an additional cost, not the normal cost, as the normal cost has been included in the business plan. Finally, unweighted performance indicators (like the number of delays) lose information on the severity of impacts in the indicator. A 30 minute delay is completely different from a multiple day delay, yet they could be counted as equal in the indicator. A good way to prevent risks being defined wrong is to think of risks as a chain from cause to consequence, as shown in Figure 59. A cause happens, impacting an entity, which shows a reaction which has value consequences. For assets the reaction typically is the failure mode (IEC, 1985a).

11.4.2 The risk matrix does not prioritize correctly

Even if the risks are defined properly so that the risk matrix can be applied, occasionally the risk matrix may produce incorrect prioritizations. In this context, incorrect prioritization is that a risk with a higher exposure is assigned a lower risk level (Cox, 2008). This is in a sense a more serious form of failing to prioritize, as it can result in important risks (in term of the exposure) not getting the attention they deserve. Suppose one risk has a likelihood of 0,99 and an impact of 9,99, whereas another risk has a likelihood of 1,01 and an impact of also 1,01. If the borders are 1, 10, 100 etcetera the first impact would be a lower level risk than the second (same impact band, lower likelihood band), yet the exposure is 10 versus 1. This is a clear mismatch between exposure and risk level. Unfortunately, this potential for incorrect prioritization is inevitable. Both likelihood and impact are an order of magnitude wide, resulting in the risk level being two orders of magnitude wide (1*1 versus 10*10). Yet, risk levels are only one order of magnitude apart, as moving up one category in either direction should result in a higher risk level. It does not matter whether scales are linear or logarithmic, nor if the risk levels are two orders of magnitude apart. At the edges the risk levels overlap. This is not a mere theoretical fact, it really happens in practice, as shown in Figure 60.

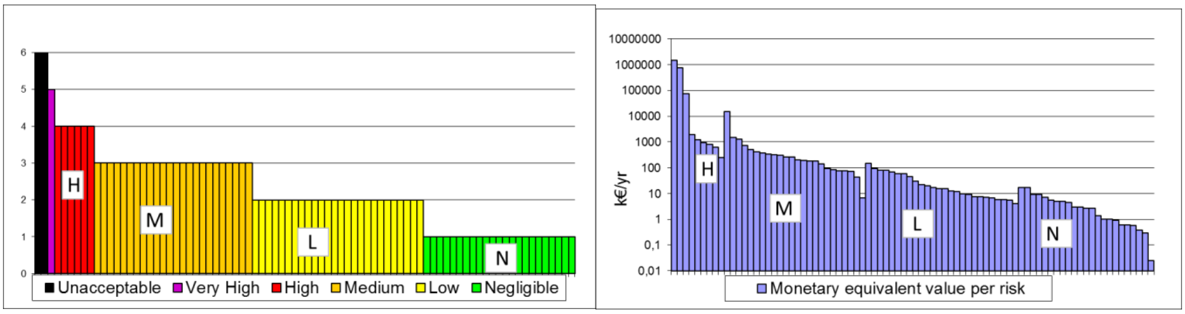


Figure 60: Risk level (left) and exposure (right) for a real collection of risks (both graphs have same order of risks).

Even though there is a mild correlation between exposure and risk level, it is by no means perfect. At the edges of the risk levels, the exposure suddenly goes up, sometimes more than an order of magnitude. On a closer inspection, it becomes apparent that about half the Medium risks are higher in exposure than the lowest High risk. Only Low risks are certainly lower in exposure than High risks. But does this mean that risk matrices are fatally flawed? No it does not. As can be seen in Figure 58, the Low risk band is precisely between very Low and Medium, but there is a significant overlap with these categories. However, there is no overlap with the High risk band, it is just contacted in the edges. Secondly, if risks are placed into an impact or likelihood band, they are probably in the middle, and then the distinction seems reasonable. Only if with reasonable accuracy the risk is near the edge of a band, the risks are prioritized incorrectly. But why would one use a rough method as a risk matrix if more detailed information is available?

Another problem in incorrect prioritization can occur in the lowest bands of the matrix, as they generally are limited on the lower side by 0. Yet, in orders of magnitude that is an infinite number of bands. Thus, in the example of Figure 57 with the border between impact 1 and 2 being 1000 euros, a 1 eurocent loss occurring more than 1001 times a year would be a medium risk, whereas a 999 euro loss occurring 99 out of 100 years would be negligible. The exposure is 100 times higher, the risk level 3 categories lower. But this is simply an incorrect design of the matrix. The bands are orders of magnitude, so even the lowest band has a lower limit. Anything below that limit cannot be accurately measured by the risk matrix. If the risk is relevant, the number of bands should be expanded until the risk fits into the matrix.

11.4.3 The risk matrix results in incorrect decisions

Risk matrices are often advocated as tool for decision making in terms of risk intervention. The risk level determines whether the risk should be mitigated, and the position in the matrix determines what type of intervention is preferable. For high impact low probability risks **transfer** (insurance) is recommended, for more likely high impact risks the impact should be **reduced** (like emergency response plans), high likelihood risks should be **prevented** and low impact low likelihood risks should be **accepted**. This is shown in Figure 61.

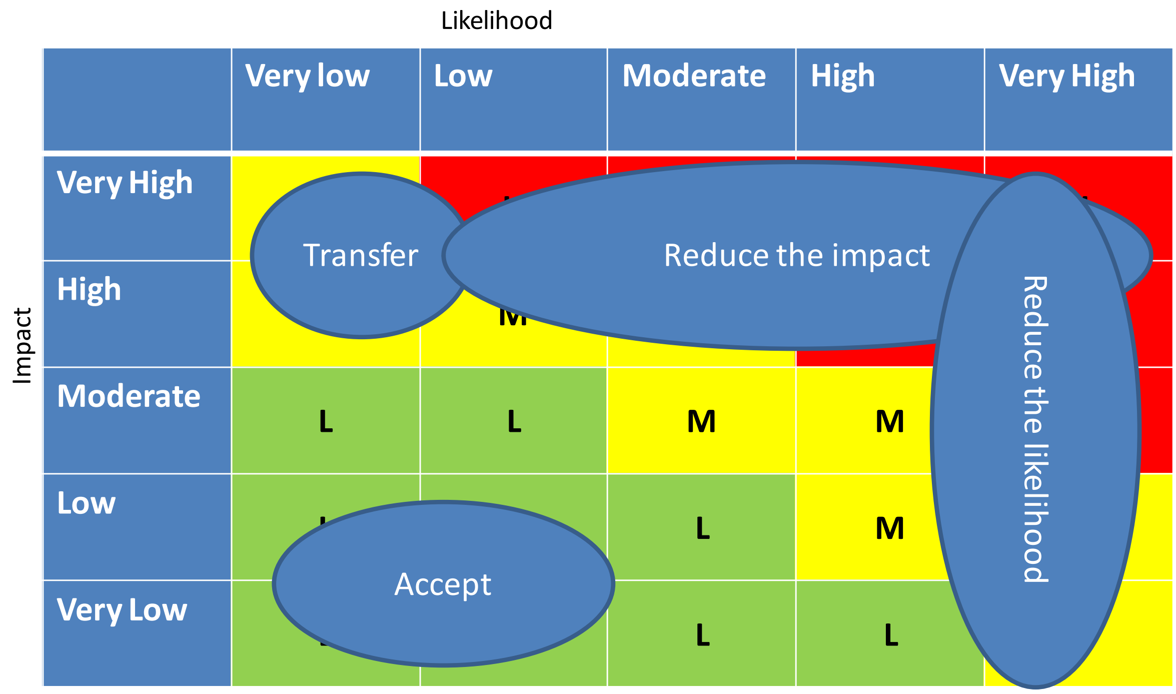


Figure 61: Risk intervention options in the matrix .

A well-known criticism on the intervention scheme is that is does not appreciate the costs nor the availability of a measure. It is not all high risks that have to be eliminated. If interventions are very costly, one could be better off by accepting the risk. Only intolerable risks have to be mitigated (hence the name intolerable), though even this is relative. Some risk may be branded intolerable but prove impossible to mitigate. They have to be accepted and thus are de facto not intolerable (an illustration that care should be exercised in determining what is intolerable). Risk acceptance depends as much on options to mitigate the risk as on the risk level (Stern et al., 1996). Yet, the rule of thumb of Figure 61 has perseverance, presumable because the rule works reasonably well in situations where the costs of interventions are comparable. This condition is not realistic in asset risk management, some interventions may be relatively cheap (e.g. extra inspections) and others very expensive (e.g. requiring complete redesign). But the condition seems to apply in deciding what interventions to develop (but not yet execute). The analysis costs of a risk are not very dependent on the risk level, as are the design costs. And if no other information is available than estimates what order of magnitude impact and likelihood are, it may be the best there is. However, the most important task of risk analysis is resolving the uncertainties around the impact and likelihood, or at least make them explicit. Once the information is available, there seems to be no reason to base decisions on the risk matrix, which is (given the previous to criticisms) a rough tool. There are much better approaches available, based on rationales like cost benefit theory, utility theory or social choice theory (Merkhofer, 1987). Expressing the exposure in terms of monetary equivalents is already a first step towards cost benefit analysis, which has (from a pragmatic viewpoint) the advantage that it results in a clear yes or no for the intervention based on the equivalent net present value. If there are budget constraints, the best set of interventions could be determined with a portfolio decision (Wijnia and Warners, 2006). To do the standard ISO Standard (ISO, 2009a) full honor, it should be noted that it prescribes balancing the costs and efforts against the benefits.

Additional to this point with regard to the accuracy of the matrix as a decision tool, there is the problem of insufficient interventions. Insufficient in this sense is that the risk level before and after the intervention are the same. This could happen for example if an intervention "only" halves the exposure. In terms of the risk matrix it would be an insignificant intervention (no effect) and thus be rejected, whereas it could be a very good intervention in terms of the cost benefit ratio. This problem only applies to risk matrices with logarithmic scales, but it has been argued why risk matrices need to be constructed this way.

Summarizing, the criticism on the risk matrix as a decision tool for implementing interventions is completely correct. Risk matrices are not useful for decision making if more than an order of magnitude accuracy is required. This is generally the case in decisions on actual interventions. Nevertheless, the risk matrix is a useful tool in determining what risks to analyse and what interventions to develop. As the costs of these actions are comparable for all risks, the categorization into risk levels is often good enough. This makes one wonder whether the risk management process introduced in section 2 should be revised. The stage of risk evaluation (judging exposure against risk tolerability criteria) suggests the use of a risk matrix, whereas it became clear that the characterization of a risk in that stage (as preparation for risk treatment) is perhaps only useful in a communicative perspective. Because of the analysis, much better information on the exposure is available than order of magnitude classification, with the additional benefit of being directly useful in decision making over interventions. The only place in which order of magnitude information is good enough is in judging which risks need to be analyzed. That, is between risk identification and risk analysis. This would require the stage of risk analysis and risk evaluation to be reversed, as shown in Figure 62. Such a reversal shows large similarities with the process described for example by Korn and Veldman (2008), where it was introduced to filter the enormous amount of identified risk, which seems to be typical for the asset risk management context.

Figure 62: Improved Risk management process with steps 3 and 4 reversed compared to ISO 31000.

11.5 Conclusion

The risk matrix is commonly used in (asset) risk management. Despite its widespread use, problems do occur and the concept of the matrix itself as a useful tool has been criticised. The problems and criticisms concentrate on 3 points: (i) the risk matrix does not prioritize distinguishable risks, (ii) the risk matrix prioritizes risks incorrectly, and (iii) the risk matrix results in incorrect decisions on interventions. In an analysis of the principles in risk matrix development and of the criticisms on the risk matrix, there prove to be two major causes for these problems, which are design errors and incorrect use of the matrix

Design errors

First of all, the matrix format can be incorrect. The risks a company faces may spread orders of magnitude in both impact and likelihood, combined with (at least initially) a significant uncertainty in what the actual values are. The scales of the risk matrix have to appreciate this, which requires the use of logarithmic scales. Linear scales fail in distinguishing risks in the lower bands, that still may be orders of magnitude apart. Furthermore, linear scales are constant in absolute uncertainty, whereas logarithmic scales are constant in relative uncertainty, which is more applicable.

Secondly, risks matrices are only useful if the different values are properly aligned. This is when the decision maker is indifferent between impacts materializing on different values within the same impact band (i.e. a serious impact on safety is comparable to a serious impact on the financial scale). Ideally, this indifference should be a willingness to pay (the decision maker is willing to spend an amount to prevent an equally classified impact on another value), as that helps in developing risk interventions. If this is not the case, a high risk (even in the same cell) may not have the same meaning for all values.

Use errors

The risk matrix is an instrument to evaluate risks: that is events that have an impact and a likelihood. It is not designed to evaluate (system level) performances, as those are no events. Using it for evaluating the system may result in “intolerable” risks which are just part of the normal operation, or indiscrimination between (from a management perspective) very different levels of performance. Reason is that performances are measured in percentages, risks in orders of magnitude. Squeezing a performance criterion into the impact scale would result in a serious deformation of the matrix and misalignment of the values.

Furthermore, the risk matrix gives an order of magnitude result, which can be one order off. Using it beyond this resolution may result in incorrect outcomes. This almost certainly occurs in decision making on the intervention (with the exception of interventions being orders of magnitude cheaper than the risk) and also when prioritizing risks within the same risk level. If a high resolution is needed the actual exposure should be used, though this only makes sense if the information has better than order of magnitude accuracy. But there is no reason to use order of magnitude tools when better information is available. The order of magnitude information is good enough in determining which risks to analyze.

Summarizing, given that the guidelines for developing a risk matrix are followed, a risk matrix is a useful tool in prioritizing attention with regard to analysis of the risk. However, it fails in decisions requiring more than order of magnitude accuracy like decisions on interventions. Yet, in many risk management process descriptions risk evaluation only takes place after analysis in preparation of risk treatment. It might be better to reverse the steps risk evaluation and risk analysis in the process descriptions, and make perfectly clear at what level of uncertainty an order of magnitude is useful. Once the analysis has been performed, much better information is available and it would be a waste of effort not to use that.

11.6 Epilogue

This addition to the chapter is based upon the internal Enexis document on the history of the risk matrix (Wijnia, 2010).

The development of the risk matrix at Enexis started in 2002, by defining the business values, identifying consequences on those values and determining a willingness to pay to prevent those consequences. This resulted in seven values: finance, safety, quality of supply, compliance, reputation, environment and regulatory. Each of these values had a weight factor (based on the sum of the willingness to pay to prevent the identified consequences on the value). In essence, each value also had a separate matrix, though several values had similar matrices. This resulted in 3 distinct matrices. The risk level in these matrices only had a mild correlation with the expected value. In practice it did not prove very useful. This was not only because the complicated multi matrix system was difficult to apply, but also because it did not fit asset management. Common incidents fell outside the matrix, both because of the magnitude of consequences (far below the lowest impact) and the frequency (occurrence much higher than the highest likelihood). Furthermore, users doubted the consistency across the values. In hindsight, the devised value system violated basically every design rule mentioned in this chapter.

Several repairs were conducted. First of all, the scope of the matrix was increased to cover smaller and more frequent risk events. The consequence descriptions per impact level were adjusted to match the asset management practice. But the axiom of a risk matrix per value was not abandoned in the first revision, nor were the initial 7 values. The resulting system was still regarded as way to complex by the intended users. In the second revision therefore all matrices were compressed into a single one, and only the 4 values that were used regularly (finance, safety, reliability, compliance) were maintained in the matrix. Furthermore, risk levels were aligned with the expected value, by simplifying the consequence descriptions and using a logarithmic scale for the consequences (likelihood had been logarithmic from the start). Inspiration for this second revision came from a visit to Shell Global Solutions (as mentioned in section 8.2.2). Because the financial column still represented a willingness to pay, this way a constant willingness to pay per unit of consequence was achieved over the whole matrix.

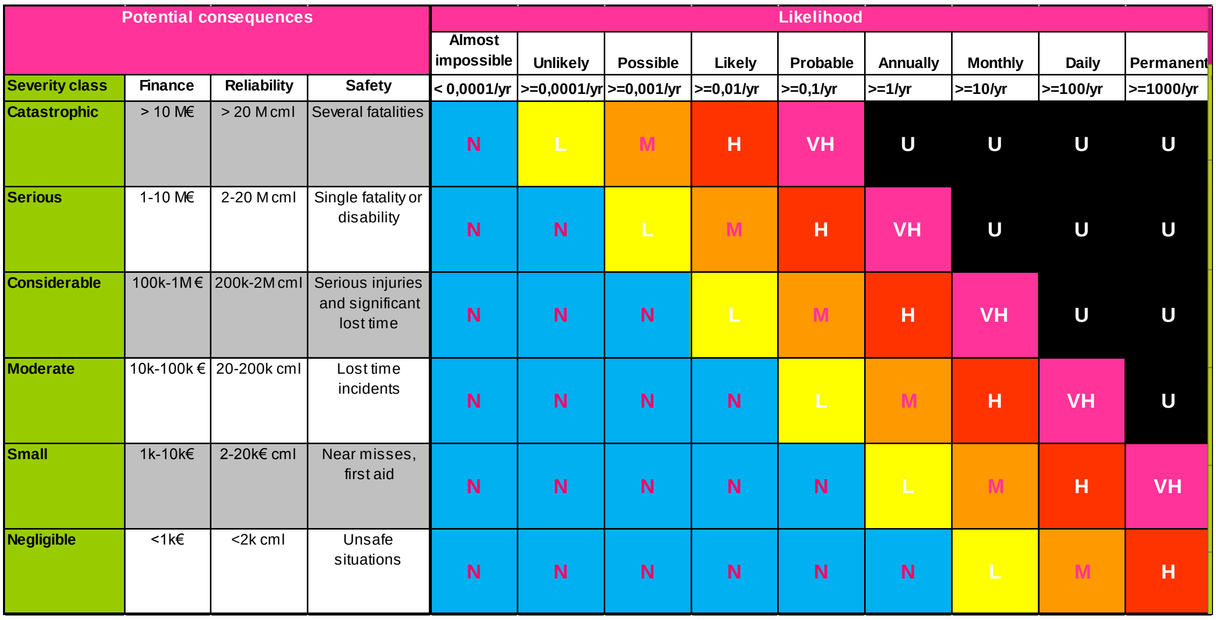
Over time, minor adjustments have been made. After a large interruption (Haaksbergen 2005) it became clear what the characteristics a top level incident on reliability were, and the impacts on that value were rescaled to accommodate that insight. Furthermore, Essent Holding (the owner of the network operator) required reputation to be included in the matrix. Both changes were approved by the board of commissioners of Enexis at the end of 2005, a milestone in the risk matrix development. As proof of its validity, that result has been included in the Dutch standard on asset management, NTA 8120 (NEN, 2009).

In 2008 an effort was made to include the value sustainability again in the matrix (Wijnia and Peters, 2008), but now following the design rules as mentioned in this chapter. The resulting matrix has been in use for a number of years (Enexis, 2011), and with some modifications[185] in 2013 it still was in use

---

[185] Relabeling some business values, plus rescaling of safety (single fatality moved one category up), customer satisfaction (number of small customer complaints) and sustainability (rescaling plus use of CO2 emission as KPI).

during the finalization of this thesis (Enexis, 2015). In Figure 63, an excerpt of the (translated) matrix of 2009 with regard to the values used in this thesis (finance, reliability, safety) is shown. The original matrices are publicly available by means of the Quality and Capacity plans to which reference was made.



Potential consequences				Likelihood								
				Almost impossible	Unlikely	Possible	Likely	Probable	Annually	Monthly	Daily	Permanent
Severity class	Finance	Reliability	Safety	< 0,0001/yr	>=0,0001/yr	>=0,001/yr	>=0,01/yr	>=0,1/yr	>=1/yr	>=10/yr	>=100/yr	>=1000/yr
Catastrophic	> 10 M€	> 20 M cml	Several fatalities	N	L	M	H	VH	U	U	U	U
Serious	1-10 M€	2-20 M cml	Single fatality or disability	N	N	L	M	H	VH	U	U	U
Considerable	100k-1M €	200k-2M cml	Serious injuries and significant lost time	N	N	N	L	M	H	VH	U	U
Moderate	10k-100k €	20-200k cml	Lost time incidents	N	N	N	N	L	M	H	VH	U
Small	1k-10k€	2-20k€ cml	Near misses, first aid	N	N	N	N	N	L	M	H	VH
Negligible	<1k€	<2k cml	Unsafe situations	N	N	N	N	N	N	L	M	H

Figure 63: Translated excerpt of the 2009 risk matrix for the values finance, reliability and safety (Enexis, 2011).

PART 3: FINDINGS, DISCUSSION, CONCLUSION, RECOMMENDATIONS AND REFLECTION

12 Findings with regard to optimization

In this chapter, the key findings per experiment with regard to optimality are reported and discussed. The findings of the experiments will be presented in the order in which they were conducted. The next chapter puts these findings in the context of the knowledge gaps. The research questions will be answered in the conclusion section.

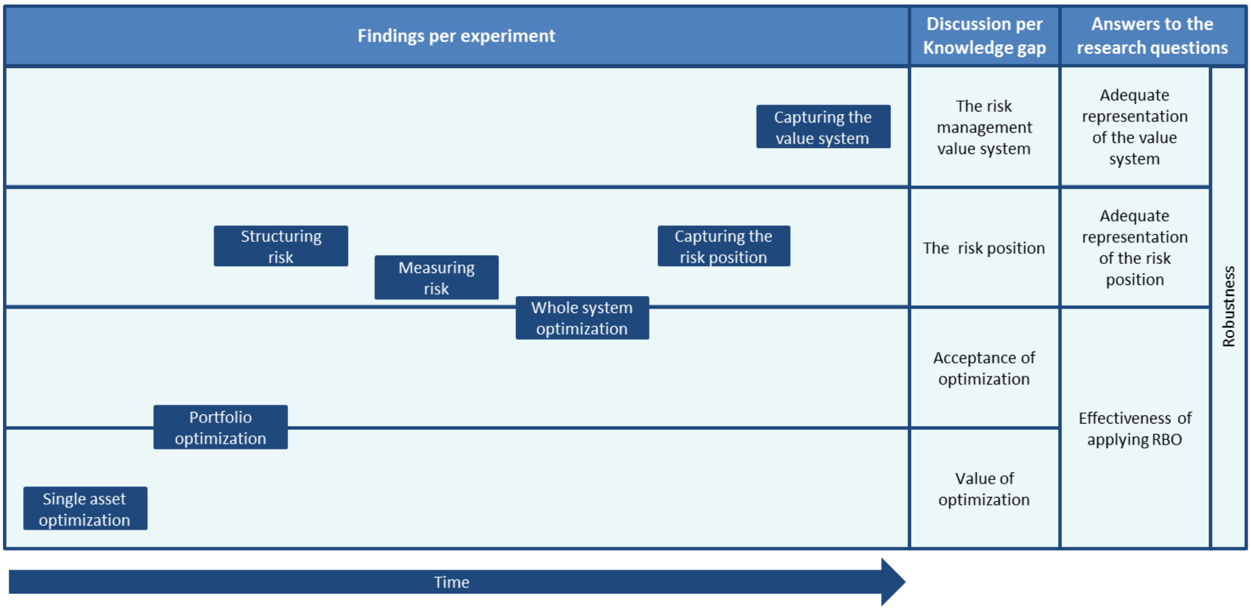


Figure 64: The research timeline.

12.1 Single asset optimization

The first experiment was about the implementation of risk based optimization for single asset decisions. The research goal was to see if there was enough potential value in optimization to justify a wider rollout. In this experiment, the potential value of economic evaluations (including some thoughts on real options) was contrasted with classical technical evaluations. The decision problem in this experiment was about the optimal timing of a circuit upgrade, a typical decision in the electricity distribution grid. To make such a decision risk based, some valuation of reliability is required. For the monetization of reliability reference was made to several studies, though these did not fully agree with each other. The value used in valuing reliability was that derived from quality regulation the Netherlands, a value of €0,25 per customer minute lost for a grid at Enexis quality.

The initial finding was that a large potential for postponing investments existed. Moving from static load limits to dynamic load limits (enhanced technical) would postpone the investment some 6 years, and inclusion of considerations with regard to risk (like load shedding or using the thermal equivalent value) could postpone the investment much further, about 20 years. However, as demonstrated in the epilogue to the experiment, the potential to make a different decision does not necessarily mean that there is much value to be gained. Postponing the investment would increase the energy losses, which would reduce the value increase of postponing the decision. Combining considerations on the energy losses with considerations on risk results in a much more nuanced outcome, as shown in the Figure 65.

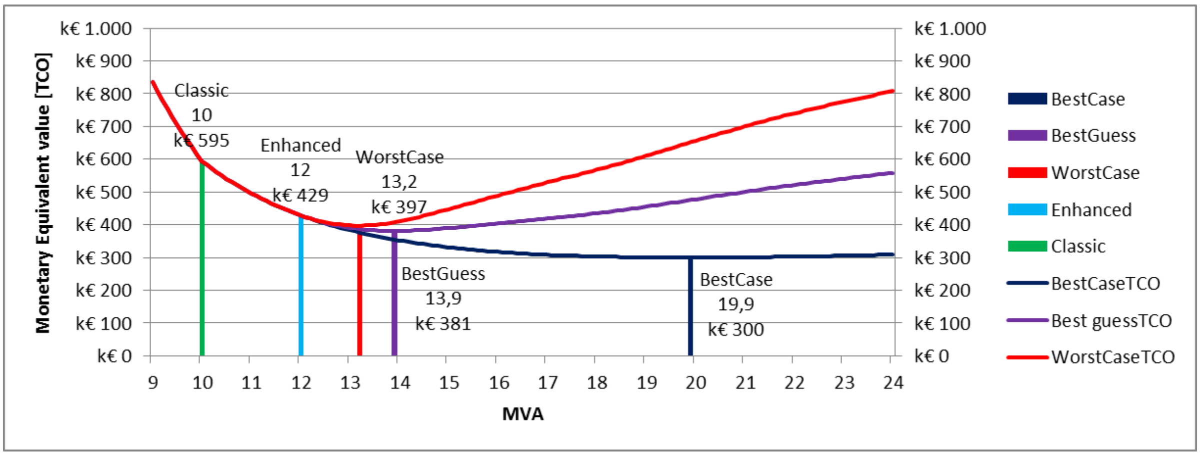


Figure 65: Comparing different outcomes, from classical (avoiding all risk), enhanced (risk avoidance given adequate response), worst case (optimal for worst assumptions on risk) to best case (optimal for best assumptions on risk, equivalent to not valuing risk at all). Figure based on graphs in chapter 5. The best guess is the average of best case and worst case. Optimization at a real interest rate of 5%.

Moving from the classical approach (avoiding the potential for overload compared to the nominal rating) to a risk based approach would bring a significant improvement, in the magnitude of 30%-50% of the original TCO, depending on the assumptions with regard to risk. Whereas the start of this range is in line with a postponement of 6 years at 5% interest rate ($1\text{-}0{,}95^{6}$ = 27%), the end is much less than a 20 year postponement of investment (equivalent with the best case scenario) would suggest ($1\text{-}0{,}95^{20}$ = 65%). Within the range of potential assumptions with regard to risk, it is possible to find a point that performs reasonably well under all assumptions, as it is not more than 11% worse than either of the extreme assumptions and less than 2% worse than the best guess optimum. This is less than many other uncertainties, like for example that in the opportunity cost of capital. As discussed in the epilogue to this experiment, a lower interest rate would decrease the importance of capital expenditure against energy losses, meaning the economic optimum would be reached at a lower value for peak power. That would bring the best case optimum closer to the worst case optimum, reducing the potential error of the enclosed optimum.

It also has to be stressed that the potential postponement of investment diminishes at higher growth rates of power demand. At a growth rate of 2% per year (as used in the analysis), a limit elevation of about 1MVA equals about 5 years, but at a rate of 7% per year this is reduced to slightly more than 1 year. Changing from the classic approach to a worst case approach then would delay the investment some 3-4 years. In different circumstances (like the 60s-80s), the classic approach thus may be regarded as a very robust strategy, though in the current circumstances with a low growth rate there is a significant value loss.

Based on these considerations two key findings can be formulated:

- In the current circumstances significant value can be found in using a risk based approach instead of classical approaches based on technical criteria.
- The optimal moment is not defined very sharply. Optimization of single assets therefore is more about indicating the right window of opportunity then about a precise point.

12.2 Portfolio optimization

The second experiment expanded the optimization of a single asset decision to the optimization of a set of decisions by means of a portfolio approach. In this approach, the best set of measures would have to be selected within constraints. The decision problem was the allocation of budget for the annual plan. The goal of this experiment was to gain understanding of the feasibility and acceptance of such an approach, the achievable consistency of decision making between teams and the potential value of such an approach.

The basic mechanism for a portfolio decision is ranking of projects according to their value for money, the yield. The optimal set of projects is then the set of projects that fits the constraints, in general limited to the available budget. To produce a ranking over all projects a consistent value system is a prerequisite, otherwise any comparison is impossible. However, as the experiment proved, that does not need to be absolute, in terms that it facilitates cost benefit analysis resulting in an accept or reject per individual project. Any system that produces an acceptable ranking is sufficient. In the experiment, a relative ranking (based on the weighted relative improvement per business value) was used. Furthermore, the portfolio decision does not require explicit knowledge on the risk position. That is only needed to establish whether the value system in combination with the budget results in an acceptable performance, but not for the selection of projects within constraints as such. The portfolio decision as it was performed in the experiment thus was a procedural optimization: making the best choice within constraints, not necessarily the best choice that could be made.

For reasons of practical feasibility, this experiment was conducted in several sub teams, for different parts of the asset base. Each team identified risks for which interventions would be developed. The achievable consistency of the value system was tested by integrating the results of these different teams into an overall ranking, mixing the priorities established per team. If teams accepted that some of their projects would be rejected because projects of other teams performed better, that would be a strong indication of the acceptance of a consistent value system. In the initial run of the experiment this acceptance was not achieved. With some fine-tuning of the value system an acceptable outcome for every team could be reached. However, key element in acceptance of the outcome was the option to bypass the ranking by manually selecting projects that had to be included in (or excluded from) the portfolio. In the end, only about 25% of the portfolio was ranked, the rest was manually selected (mostly obligatory projects like new connections). The outcome was (to a large extent) even accepted by the board of directors, despite the bypass options used by the teams.

This acceptance of the outcome was (with the benefit of hindsight) perhaps the biggest surprise. The best explanation lies in the spread of yield for the decisions in the portfolio. Given any budget limit, only a fraction of the portfolio is within a reasonable distance (say 20%) of the marginal yield at the budget limit. This is shown in Figure 66.

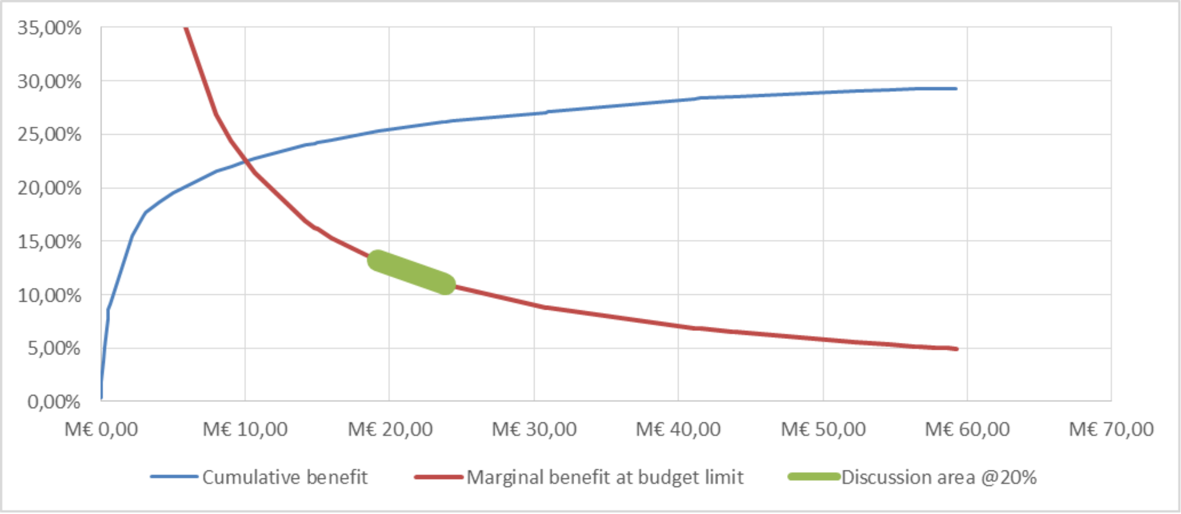


Figure 66: The development of value as a function of the budget of the portfolio decision, the marginal benefit of proposals and the discussion area of 20% around a budget limit (in this case of 20 M€).

Discussion typically focusses on this subset of the portfolio. Apparently it was accepted that the poor yielding projects never would make it to the portfolio, and nobody considered removing a high yielding project from the portfolio. This may be a result of peer pressure: trying to get all of the own poor yielding pet projects into the portfolio would not be accepted by the peers, as they would have a wish list themselves. A social rule derived from this experience could be that everyone would be allowed one pet project, under the provision that the total of pet projects would not exceed 5% of the budget. The logic behind this is that if someone is willing to stand up amongst peers, that person most likely has a point.

The portfolio decision also provided insight in the relation between whole system cost and performance by plotting the cumulative value of the ranked list of projects. Contrasting this with the cumulative value of an unranked list (for several hurdle rates) provided insight in the potential value of the portfolio decision, as discussed in the epilogue to the experiment. Estimates for the additional value that could be realized with a portfolio approach were in the 20% range. The marginal benefit at the budget limit in itself provided insight in the right hurdle rate for decision making. This allowed the development of an absolute value system instead of a relative one, facilitating aligned go/no-go decisions for individual projects without having to consider the whole portfolio.

This results in the following key findings on portfolio optimizations:

- A risk based optimization of the portfolio is feasible, in the sense that a value system can be formulated that allows comparison of many different types of projects for different subsets of the asset base.
- Portfolio decisions based on the ranking of projects can bring significant value if not all proposed projects can be executed.
- Optimization does not need to cover the whole portfolio, partial coverage due to resource constraints is acceptable. Over time the coverage can increase.
- Key element in accepting the outcome was the option to bypass the ranking. This removed the need for the value system to be 100% accurate and to cover all relevant aspects.

- Discussions were limited to a very small fraction of the portfolio, with some 95% being accepted on the basis of their benefit to cost ratio.
- However, portfolio decisions only give the best set based within constraints on the value system, they do not provide validation for the rightness of the value system or the constraints.

12.3 Structuring risk

Given the potential value of even “only” procedural portfolio decisions, expansion into a formal system optimization provided a tempting perspective. However, to execute a formal optimization a quantitatively adequate overview of the risk position is needed. At the time of the experiments no such adequate overview existed, nor did methods for providing that overview. The list of identified risks as used in the preparation of the portfolio decision only provided the set of most relevant risks. Many of the entries interacted at least for a part (for example by partially overlapping), and there was no method available to test whether all risks were included. The goal of this experiment was to find a method for providing an overview that could deal with the interactions between risks.

Several methods were reviewed to do so. This included an asset based bottom up approach (like FMEA) and an organizational structure based top down approach. Both are flawed, perhaps not fundamentally but at least pragmatically. Asset based bottom up strategies fail in recognizing all interaction between assets because of a combinatorial explosion (resulting in apparently emergent behavior), whereas the top down approach fails because of overlaps in responsibility for risk factors. Furthermore, both approaches use a very narrow definition of the concept of risk, whereas literature only seems to agree on the fact that there is no agreed definition of the concept of risk. Both methods thus may be fundamentally flawed as well.

A solution was found in the recognition of risk as a process[186] from cause to consequence, as inspired by Morgan et al. (2000). In that paper, the suggestion was made to test several categorization strategies in parallel (at least one per phase of the risk process) to find the one that best suits the decision problem at hand. This means that structuring risk is an optimization itself! In order to fit the decision problem in the energy distribution infrastructure, some modifications of the original chain were made. All phases of the chain of risk processes were used as categorization strategy, resulting in a list of risk factors per phase.

These categorizations per phase expanded the risk process into a risk graph. Each risk factor was a node, and the interrelations between the risks were the links. Each path from left to right through this graph then would be a risk. The interdependence between risks then would become apparent because of shared nodes. In this directed graph, additional interactions could be modeled by links going in the opposite direction. Failures of assets could be the cause of other asset to fail, and even interventions to limit the risk may cause other risks. This is demonstrated in the diagram below, combining several notions of the experiment.

---

[186] The original formulation was a chain of risk processes, which developed into a process of cause, asset, reaction and consequence in this research.

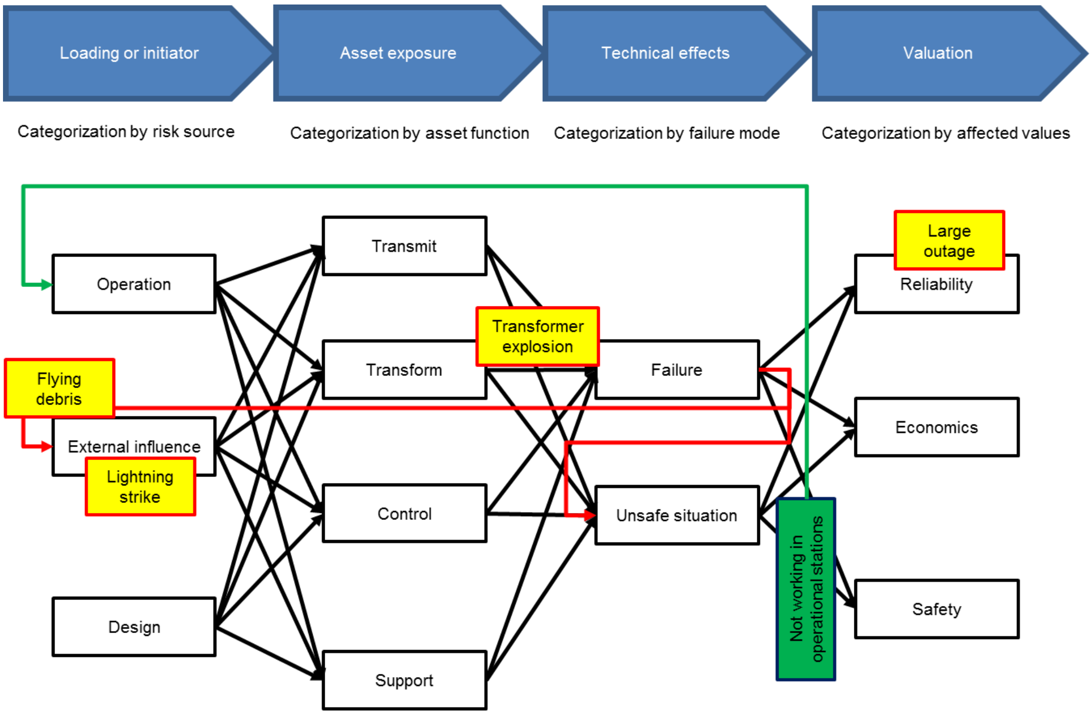


Figure 67: The notions on risk interdependency over the phases of the risk process combined. Overlap exist because of risks formulated in different phases of the process, like lightning strike, one of the causes of transformer explosions, which is one of the causes of large outages. The effects of the failure can be a cause themselves, i.e. sequential interdependency. Interventions to limit the risk (like not working in operational stations) may also be a cause for other risks, which in this case would be an increase of interruptions.

The key finding of this experiment was:

- Thinking of risk as a process helped in understanding the complexities in constructing a correct overview of the risks. It did not yet provide an immediate answer on what the best structure was for the risk overview, but at least it provided a structured framework to construct that.

12.4 Measuring risk

One of the most important values of the gas distribution grid is safety, as it forms the basis for many regulations. However, there was no good indicator to measure the actual safety performance of the distribution grid. In a review of all true accidents in the past decade, it became apparent that a direct measurement of safety by counting the (weighted) incidents that resulted in injuries or worse would result in a highly variant indicator. Only in the long run, such an indicator could give a representation with regard to the underlying quality of the gas grid, though it has to be realized that this quality may have changed in that same time. The goal of this experiment was to develop a performance indicator for the safety of the gas grid (hence the Safety Indicator), that would capture the underlying quality more timely.

This experiment provided a testing ground for the risk process, as that could help to understand the materialization of accidents from the failure modes, resulting from causes impacting on assets, and perhaps pinpoint the most important risk factors. By plotting the recorded incidents and accidents against the risk process, the riskiest paths through the risk graph could be identified, along with the early warnings in those paths. By counting the early warnings, a rising risk level could be detected before the number of real accidents increased. Theoretically[187], the earlier the warning in the process (i.e. failure modes or even causes), the higher the number of warnings and thus the better the prediction. Unfortunately this proved to be difficult to realize, as the quality of registration decreased by moving to the less severe occurrences. Unsafe situations often simply were not recognized or recorded. The measurement resolution thus proved to be an optimization itself. In the experiment, the optimal resolution proved to be the failure that required immediate repair because of the risk of the situation, as these work orders were already registered with reasonable (not perfect!) accuracy in the administration of the network operators. This concept of optimization of resolution is plotted alongside the incident triangle of the experiment in Figure 68.

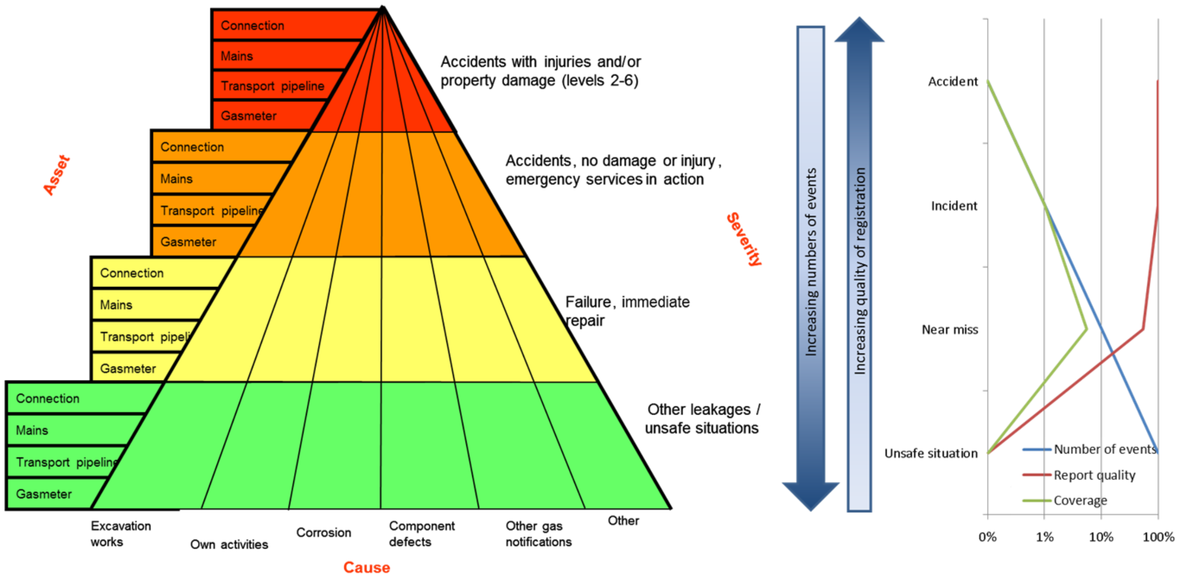


Figure 68: Incident triangle with the concept of optimization of the resolution. The numbers in the optimization graph are an illustration and not based on actual measurement.

A problem in measuring safety was that to come up with a single safety number, a way had to be found to add incidents of different severity levels. Furthermore, most safety incidents had a financial component in the form of property damage. The solution applied was alignment of severity levels in safety (from unsafe situations to fatalities) with a logarithmic classification of property damage. The rationale for this was that the propagation probability of incidents was also of logarithmic characteristic: an incident with a severity one level higher typically occurs ten times less. The resulting scheme (the start of the risk matrix!) allowed expressing every incident in an equivalent monetary value.

Below, the key findings are summarized:

- True accidents do not necessarily occur often enough to create a reliable measure of underlying risk. In order to predict risk more accurately, it is necessary to analyze risk at a more fundamental level.

---

[187] Under the assumption of limited propagation through the process: not every cause impacting on an asset results in a failure, not every failure results in an accident with an impact on values.

- The risk process is a useful tool for that analysis, as it helps in understanding how the accidents materialized, including an understanding of what the relevant risk paths, risk factors and precursors (combination of cause and asset) for accidents are.
- All accidents could be mapped to a relatively small number of precursors. The use of the risk process thus simplified the analysis significantly.
- Counting these precursors for accidents combined with an average impact per precursor provides a reasonable estimate for risk in the system. As the number of occurring precursors rises before it becomes apparent in the number of accidents, this is in a way a leading indicator.
- Based on the limited (both in quantity and quality) dataset, the prediction of risk only has a limited accuracy, but with an increasing dataset, this can be expected to improve over time (and thus is a robust method).
- The resolution of the Safety Indicator proved to be an optimization in itself, because of the diminishing quality of data at higher levels of resolution.

12.5 Whole system optimization

The fifth experiment was about formal optimization of the whole within constraints. This combined the results of the previous experiments: risk based optimization of individual assets from the first experiment, the prioritization of interventions to deal with the constraints, the risk process to structure the asset base into a limited number and the monetization factors from experiment 1 and 4. The research goals were about feasibility of a whole system optimization and acceptance of results. The experiment thus provided an opportunity to validate the value system by means of the resulting whole system performance.

The decision problem in this experiment was the optimization of the long term replacement strategy for the assets in the distribution infrastructure. In short, many assets were constructed in the booming economies of the 1960s and 1970s. With estimates on asset life ranging from 40 to 80 years, many assets could be expected to reach end of life within 10-20 years. In parallel, the infrastructure managers were confronted with organizational continuity issues. Many of the employees were contracted during the expansion problem, and a large part could be expected to retire on a short term, shortly before the expected rise in replacements because of end of life of the assets. Combined these two trend could result in a spiral of decline, where all preventive work would stop because all resources would be diverted towards corrective replacements. This created a sense of urgency towards developing an optimal long term replacement strategy.

Optimality was defined in this experiment as the lowest Total Cost of Ownership for the next 50 years, including the monetized performance on reliability and safety. By means of the risk process the asset base, consisting of thousands of different asset types, was grouped into a limited number (several dozen) of asset classes. Differentiating characteristics between these classes were the development of the failure rate and the failure consequences. Failure rates development mostly depends on the used technology, whereas consequences depended both on technology and position in the grid. Combined with the age profile per asset class and summed over all classes this resulted in a prognosis of the whole system performance. The risk profiles were calibrated against the actual whole system performance.

The optimal replacement moments for the asset cluster were determined using a marginal approach: suppose the asset would be replaced next year, would it be worthwhile to replace the asset right now. This allowed the optimal age to be calculated by means of a formula, allowing on the spot recalculations for other assumptions. The optimal replacement ages were implemented in a whole system simulation model to be confronted with the organizational constraints. If constraints were limiting the execution, resources would be allocated on the yield (risk reduction per unit of costs) of the replacements. We found that in the model the constraints in the form of available personnel effectively blocked the execution of preventive work. Just implementing the optimal replacement strategy would only have little impact on the total organizational value.

Several strategic alternatives were tested for relaxing the constraints, like advancing replacements, increasing recruitment capacity, adding flexible resources and increasing the efficiency for replacements. The value of those alternatives was evaluated in the same terms as the replacement of the assets, providing perfect consistency between asset decisions, portfolio decisions and strategy decisions. There was uncertainty in the assumptions with regard to the monetization factors, age profiles, failure consequences and end of life estimates. However, a sensitivity test proved that monetization factors and failure consequences did not matter much for optimization of individual assets. The optimum age would move a few years, but because of the shallowness of the optimum that would not impact TCO much. The optimal replacement age thus is more a window of opportunity than a precise moment. The estimated ageing rate however would impact asset optimality. However, for the strategy with regard to constraints to be wrong the estimated life spans would have to be at least 50% longer, a bet the organization was not willing to make.

We found that the whole system optimization had a very large significant value as shown in Figure 69. Continuing with a corrective replacement strategy (red line) would result in a very serious (almost an order of magnitude) deterioration of performance. Implementing a preventive strategy (purple) would limit this deterioration, but due to organizational constraints it would only become effective in the far future. Adjusting the constraints would result in a much better performance, as demonstrated by the green line, which demonstrates consistent ranking on yield. Unranked selection of interventions (the blue line), results in a significant deterioration of performance in periods of scarcity.

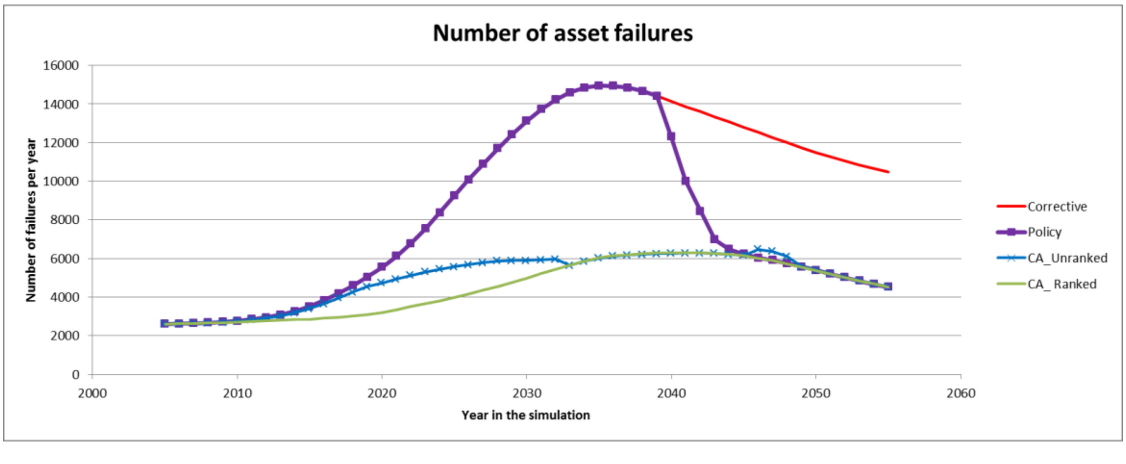


Figure 69: The effect of system optimization. The red line represents the number of failures at a corrective replacement policy, The boxed purple line a preventive strategy in current constraints which limit execution of the strategy. Adjusting these constraint (green line) results in a much better performance, though an unranked allocation of resources (crossed blue line) loses significant value.

The key findings of this experiment are:

- Using the risk process to structure the asset base allowed the risk position to be effectively modeled by a limited number of asset classes instead of thousands of asset types.
- In the whole system optimization, CBA can be used at all levels in perfect consistency: formal optimization of the asset replacement age, a procedural optimization of the replacement portfolio within the constraints and for formal optimization of the strategies for relaxing those constraints.
- Whole system optimization provides insight into the role of constraints that are neither visible from the individual asset nor from the portfolio.
- The outcomes are robust with regard to uncertainties in assumption about monetizing factors, failure consequences, age profile and development of the failure rate.
- The value of optimization is about 20-30% in total cost of ownership, both for assets, portfolios in time of scarcity as for the whole system.

12.6 Capturing the risk position

In both Measuring Risk as in Whole System Optimization, the risk process provided to be a valuable perspective for structuring risks so that the total risk in the system (i.e. risk position) could be established. However, the risks in these experiments were formulated in a single phase of the risk process per experiment, avoiding overlap between risks. What was not clear yet was if this would also work for a list of risks for which an overlap might exist, because the risks were framed in different phases of the risk process. This was addressed in the experiment of capturing the risk position. The research goal of this experiment was to see if the risk process would also be beneficial for structuring a potentially overlapping set of risks into an adequate representation of the risk position.

To test this idea, all identified risks so far (collected in several lists of a diverse format) were combined into a single structure risk register. To increase the richness of this collection, a process to collect new risks was implemented. For all risks, the potential impacts with their associated likelihoods were assessed. By running a Monte Carlo simulation for the whole register, a prediction of the whole system performance could be made, including the distribution of potential outcomes. This then could be compared with the actual performance over a number of years. This resulted in a serious overshoot in the estimated whole system performance. The observed distribution in the actual performance did not even overlap with the predicted distribution. Two potential explanations for this phenomenon were evaluated. Prime suspect was erroneous estimation of individual risks. However, given the large number of risks in the register it was somehow unlikely such a serious deviation could be the result of random estimation errors. Only a systematic bias could explain the deviation, but as most of the estimates were somehow based on reported incidents, this also was not very likely. This turned the attention towards potential overlap between risks in the register. In the risk analysis process, clear overlaps had been removed, like some risks being a specification of a more general formulation. The true cause of the overlap was much more hidden. It resulted from risks being framed in different phases of the risk process. A potential solution thus was framing all entries in the same phase, as was done for measuring risk and system optimization, and then to remove the overlaps. However, risks were framed in different phases because the perceived best solution would interrupt the risk process at that point. Framing all risks in the same phase thus would limit the potential scope of interventions, clearly not an optimal approach. Instead, the option was chosen to make the overlap transparent by linking the risks to the impacted assets if the asset was not yet part of the risk definition. By reviewing all risks to the asset, overlaps were more easily detectable because the list would be much shorter. Deactivating risks that were fully covered by others resulted in a reduction of the number of active risks from 443 to 284. The prediction of this trimmed set is much more in line with the actual observations. This is shown in Figure 70. It is important to recognize that even a limited trimming focused on the 10% most important (e.g. High and Medium) risks probably also would result in such an improvement. The orders of magnitude spreads in the expected value of risks means that this small subset most likely covers some 90% of the total risk position. For system level considerations this is good enough.

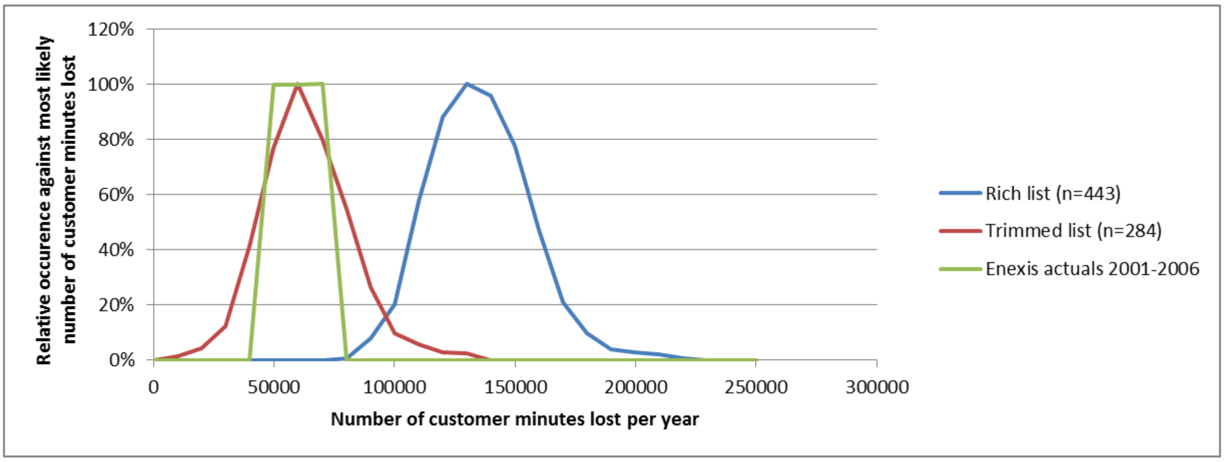


Figure 70: The effect of removing overlap by relating the risks of a rich list to the impacted asset and deactivating fully overlapping risks. The prediction of the rich list is shown in blue. The resulting trimmed list is not framed in a single phase, but exists of risks that are more or less independent. That this is good enough is demonstrated by the overlap between the prediction of the trimmed list (red line) and the actual performance of Enexis (the green line).

Below the key findings are summarized.

- Risk registers typically are (and should be) over complete, to support a wide range of possible interventions. Risks are formulated in the phase of the perceived best intervention.
- The resulting overlap can be made transparent (and managed if needed) by mapping risks to the impacted assets, as this limits the number of risks to be cross checked.
- A further decomposition of the asset base may result in a higher accuracy, however:
    - The actual risk position cannot be measured with full accuracy, and the prediction does not need to more accurate than this measurement.
    - The most important risks capture a large fraction of the risk position given the orders of magnitude spread in expected value per risk.
- Avoiding overlap by imposing a single definition of risk on the register would limit the set of potential interventions, clearly not the right way forward for optimization.

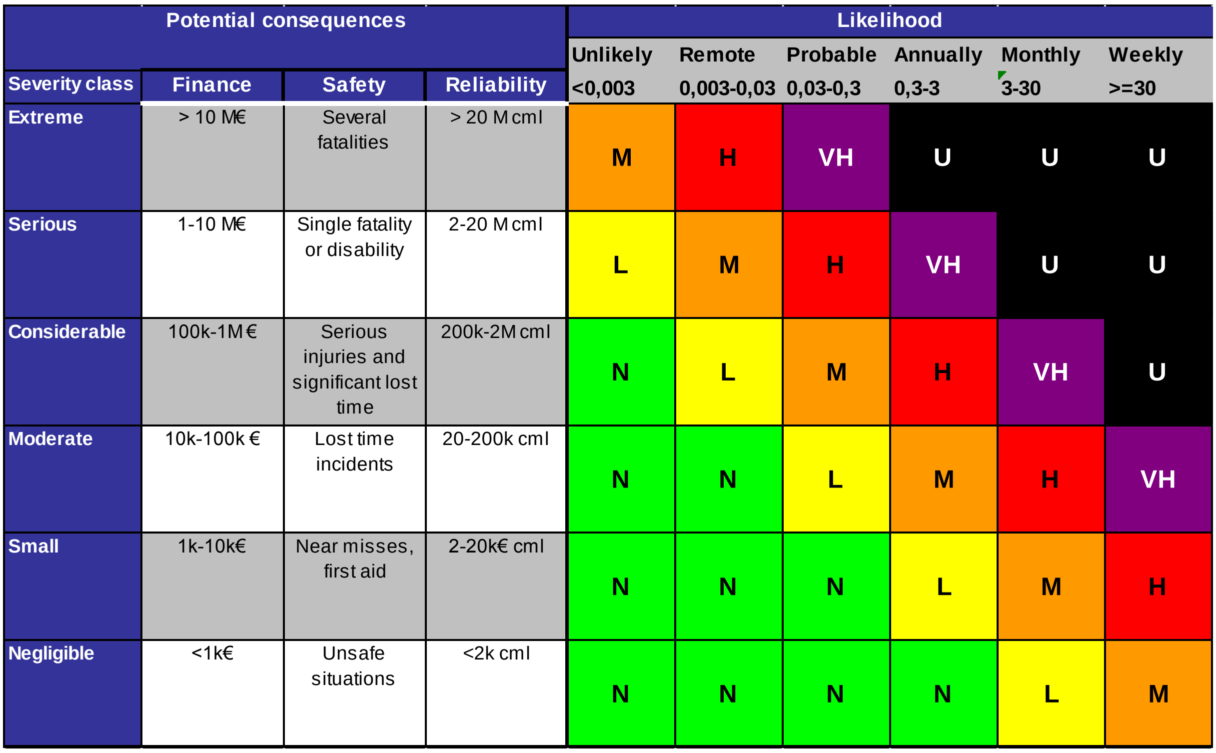
12.7 Capturing the value system

The final "experiment" conducted within this research was on the development of a robust method for capturing the value system. The key question in this experiment was how to provide a simple scheme that could be used both for prioritizing risk as for valuing interventions. A properly designed risk matrix proved to be capable to fulfill both purposes, at least, if used correctly. Properly designed means that the risk matrix follows the following 3 design rules:

1. Logarithmic scales to align with orders of magnitude that impacts and probabilities span. Additionally, these scales assure a constant relative uncertainty in estimation
2. Alignment of values based on the severity. If impacts score in the same category, decision makers should be indifferent. For impacts measured on a quantitative scale, the alignment should be constant in the per unit tradeoff: large impacts are more important than small impacts, but they do not have a higher value per unit[188]
3. Risk level based on expected values, also expressed on a logarithmic scale.

Additionally, it should be noted that the number of categories for severity and likelihood should (ideally) be equal to each other and to the number of risks levels. It does not make sense to assess probability on a 9-category scale if only three risk levels (Low, Medium, High) are used.

The matrix presented at the end of chapter 11 does not completely comply with these rules, as the number of likelihood categories is larger than the number of impact categories and the number of risk levels. Furthermore, given that expected values are typically expressed as the amount of misery per year, it makes sense to frame the likelihood categories in such a way that an annual event is in the middle of a category. In the matrix below, these small corrections have been applied. This matrix can most likely be used for the management of any set of normal risks in asset management.



	Potential consequences			Likelihood					
				Unlikely	Remote	Probable	Annually	Monthly	Weekly
Severity class	Finance	Safety	Reliability	<0,003	0,003-0,03	0,03-0,3	0,3-3	3-30	>=30
Extreme	> 10 M€	Several fatalities	> 20 M cml	M	H	VH	U	U	U
Serious	1-10 M€	Single fatality or disability	2-20 M cml	L	M	H	VH	U	U
Considerable	100k-1M €	Serious injuries and significant lost time	200k-2M cml	N	L	M	H	VH	U
Moderate	10k-100k €	Lost time incidents	20-200k cml	N	N	L	M	H	VH
Small	1k-10k€	Near misses, first aid	2-20k€ cml	N	N	N	L	M	H
Negligible	<1k€	Unsafe situations	<2k cml	N	N	N	N	L	M

Figure 71: A basic 6x6x6 matrix for the 3 values used in this thesis, after Enexis (2011).

---

[188] The typical example is an interruption of supply. The interruption of a whole city is more important than the interruption of a single street or even a single customer, but not from the perspective of the customer. The lost value for the customer does not depend on how many other customers are interrupted at the same time.

As finance typically is one of the business values, this provides a clue with regard to the right monetization factors. From the matrix above, the following monetization factors can be derived

1. Safety: 1-10M € per human life
2. Reliability: 0,50€ per customer minute lost

These values can be used in cost benefit analysis of interventions. The common practice of intervention decisions based on risk level alone cannot be advised for more than the unacceptable risks. First of all, that does not recognize the cost of interventions, and secondly, an order of magnitude accuracy typically is not good enough. Better instruments that make use of the expected value, like for example cost benefit analysis, are necessary.

The key findings of this experiment are summarized below

- A properly designed risk matrix is a powerful tool for prioritizing risk in the selection of risks to be analyzed even if there is order of magnitude uncertainty regarding impact and likelihood.
- The core elements of the value system for asset management are finance, safety and reliability, at least with regard to normal risks. In specific cases aspects like reputation, compliance and sustainability can be of importance.
- The risk matrix is an order of magnitude instrument. Decision making typically requires more accurate tools.

13 Discussion per knowledge gap

In this chapter, the findings will be discussed in the context of this thesis' theoretical framework from chapter 3 and the knowledge gaps as presented in chapter 4:

1. Risk management value system
2. Risk position
3. Acceptance
4. Value of risk based optimization

13.1 Risk management value system

In chapter 11 on **capturing the value system** we found that a well-designed risk matrix is an adequate representation of the risk management value system. It provides order of magnitude risk evaluation criteria for selecting the risk to analyze and mitigate. The alignment of values in the risk matrix provides the monetization factors for the non-financial values. The risk matrix does not need to be complicated. We found a three value six level matrix with expected value based risk evaluation criteria to be sufficient. This was demonstrated for **single asset optimization** (chapter 5), **portfolio optimization** (chapter 6) and **whole system optimization** (chapter 9).

The adequacy of such a relatively simple tool is in line with part of the literature and contradicts other parts. Alignment with literature is reached with regard to the monetization factors. For normal risks, Cost Benefit Analysis which uses monetization factors is the right approach (Klinke and Renn, 2002). The monetization factors embedded in the resulting risk matrix also align with literature. The value of reliability was initially derived from literature, but adjusted (doubled!) in its application. We regard this still as an alignment, because optimization indicates a window of opportunity which is relative insensitive to variation. This is demonstrated by Figure 4 (theoretical framework) and Figure 48 (whole system optimization). Even halving or doubling a monetization factor would only result in a small shift of the optimal moment, still within the original window of opportunity. With regard to the value of a human life, the value of € 1-10M is aligned with (de facto) industry standards. Assuming on average 40 years remaining life and the geometric average value of the matrix (€3M) results in some € 80k per life year. This aligns with the most common value in Figure 9 by Tengs et al. (1995), i.e. a value of $ 10-100k per life year.

However, with regard to the number of values to include alignment with literature is much less. If statements are made (ISO, 2009a, COSO, 2004b, Klinke and Renn, 2002) the values should address anything any stakeholder values. This is in sharp contrast with the three values that were found to be adequate in this thesis. There is also no alignment on the risk matrix being an adequate tool in the first place (Cox, 2008), though the common use is recognized. The third issue is the use of expected values for judging risk levels. This aligns with the principles of CBA, but misaligns with any non-objective notion on risk (Slovic, 1987, Aven, 2009, Hanson, 2004).

The explanation for this mismatch between theory and practice can be found in the pragmatic application of the theories. First of all, the risk matrix is not the only instrument in use for decision making. If it results in a wrong answer, this can be corrected somewhere else in the decision making process. This removes the need for the matrix to be 100% correct, which (in line with literature) indeed is (most likely) impossible. The second argument is that actual optimizations are not that sensitive to the exact valuation of risk, because of their window-of-opportunity like outcome. Even being 50% off does not necessarily result in bad decisions. The ranking of interventions in portfolio decisions is even less sensitive. This means that "forgetting" a value in the value system is not necessarily a fatal flaw, especially not if an umbrella value like reputation is used. However, if that forgotten value would be very differentiating it could develop into an issue, as happened in the **portfolio optimization** of chapter 6. Yet, because of the embedding in a decision process it was possible to resolve the issue on the spot. It has to be recognized that risk assessments also can be used in isolation, which is a topic of much literature. The value system then is used to give an absolute judgment on the tolerability of risk (especially to those not benefitting from the risk) and a much higher accuracy is needed. On the other hand, given the unavoidably subjective nature of risk valuation, it may be reasonable to question the use of value systems for absolute judgements. In **capturing the value system**, chapter 11, many of the suspected flaws of the risk matrix could be attributed to errors in design and use. Perhaps the use of risk assessments for absolute judgments is a similar incorrect use.

13.2 The risk position

Literature was found to be scarce with regard to capturing the risk position. From many sources it was stressed that all risks should be captured, as risks that are not identified cannot be managed. Yet, no robust method for identifying the relevant risks was presented, only a suggestion for methods to apply. What comes closest to an instruction for building a precisely complete risk register that captures the risk position was the set of attributes for an ideal categorizing system by Morgan et al. (2000) as discussed in chapter 11. The basic idea was that large volumes of identified risks could be reduced to a manageable number by providing a structure afterwards. The risk process (from cause to consequence) was presented as a means to identify categorization strategies.

In the experiments several approaches were tested for capturing the risk position by means of such risk process based categorization. In the experiment on **measuring risk** (chapter 8), the categorization of risk was fully based on the incidents that occurred in the past years. The reasoning behind this was that anything that did not happen in the past 10 years was not important. In the experiment on **whole system optimization** (chapter 9), categorization was much more top down. Given the focus on asset replacement, the core structure was the asset hierarchy. Finally, in the experiment on **capturing the risk position** (chapter 10) structure was provided ex-post to remove the significant overlap that resulted from a continuous broad identification of risk throughout the whole organization. Common ground for all three approaches was the asset phase of the risk process as the basic structure. All three approaches resulted in adequate representations of the risk position[189]. Interestingly, both in **measuring risk** and in **whole system optimization** the risk position could be represented with a limited number of entries, between 10 and 100. This complies with both the potentially conflicting desirable attributes (large enough and few enough), contradicting the notion expressed by Morgan et al. (2000) that in reality this often is not possible. In **capturing the risk position**, the remaining number of risks was much higher, though in the used approach removal of overlap between risks took place within smaller subsets of the risk register, organized per asset.

---

[189] Though it may be debated whether measuring risk is capturing risk, given that the categorization used is a sub division of the risk position and not an independent approach

This leads to the paradoxical result that whereas in theory there should be no such thing as a precisely complete risk register, in practice the risk position of the distribution infrastructure can be adequately represented with a limited list of risks. There is even more than one approach available to achieve this result.

The most important explanation for this result can be found in the risk level of the risks that populate any risks register. As was shown in Figure 60 of chapter 11 on **capturing the value system**, risk registers typically contain a small number of important risks with a large expected value and many more risks with a smaller expected value, orders of magnitude below the most important risks. The cumulative value at risk found in a ranked list (either by risk level or expected value) quickly approaches the total value at risk in the register, often with only the most important levels (medium and above) included. As the expected value of a medium risk typically is ten times smaller than that of a high risk, the number of medium risks would need to be ten times larger to capture a similar value at risk.

The second explanation is that the total risk position (in the sense of the total expected amount of misery) is not known accurately either. Both for the safety of the gas grid as for the reliability of the electricity grid the actual performance showed significant variation on a year to year basis. Validation against such an uncertain number is comparing the distribution, as demonstrated in chapter 10. The standard deviation of performance is at least in the 10-20% range. Combined with the notion of the cumulative value at risk, a small set easily captures 90% of the value at risk.

The third possible explanation is the normality of risk in the distribution infrastructure, as opposed with non-normal risks which are often the topic of the theoretical considerations. Normal risks can be judged by CBA and consequently can be ranked according to their expected value. Furthermore, in a mature and widespread asset base like the distribution infrastructure for electricity and gas the vast majority of risk [190] has happened somewhere in the world in the past, reducing the concept of probability to a relative frequency. Under these conditions, even the distinction between risk as a concept (the event potentially resulting in misery) and the way it is measured (exposure/expected value or risk level) may lose practical meaning, resulting in the simplification risk=effect*probability which is so widely used in the energy distribution sector. However, in risks where probabilities and impacts are difficult or impossible to assess (e.g. probability in terrorism threats, consequences in climate change), this does clearly not hold, as argued by Klinke and Renn (2002).

The notion of chapter 10 of allowing multiple definitions of the concept of risk so that no valuable interventions are excluded does not align with literature on precisely defining risk. It also does not align with the idea of using a single phase of the risk process as categorization structure. Interestingly, many of the definitions we found in literature can be plotted on the risk process if multiple levels of aggregation are considered. This is shown in Figure 72. Risks framed as a threat is placed in the cause phase of the risk process. Event like risks are positioned at the level of the

---

[190] What is meant here is the materialization of failure modes and their causes. Specific non-normal risks like climate change may alter the probability of typical failure modes, but it is unlikely that they will result in new failure modes.

individual asset. The idea of risk = probability*effect has meaning on an aggregated level of assets. The variation of performance (Brown and Humphrey, 2005) or deviation from objectives (ISO, 2009a) regards the whole system. In a way this is not a separate risk, it is just the potential distribution of outcomes based on entries in the risk register as demonstrated in capturing risk.

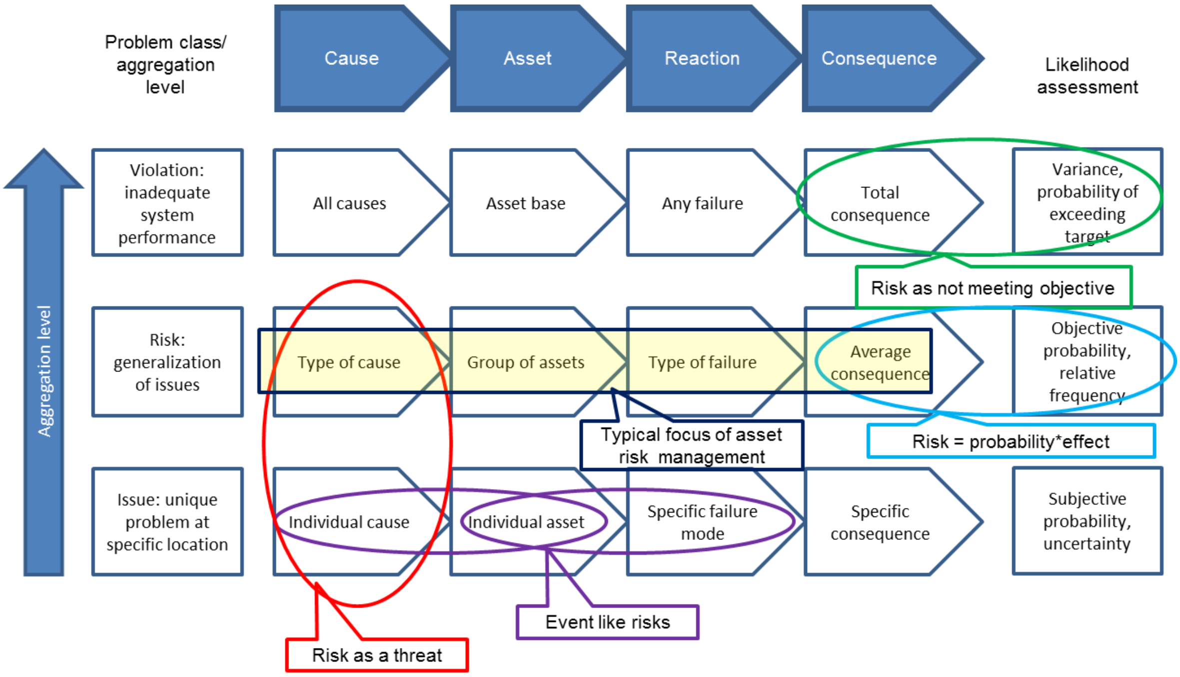


Figure 72: Several definitions of risk plotted in the risk process at different aggregation levels.

The likelihood assessment changes with the aggregation level. For unique issues, objective probability has little meaning and subjective probability or uncertainty is more adequate. For groups objective probability often is expressed as relative frequency. Risks in asset management are typically formulated on the group level. For system level problems (violations), likelihood assessments tends to be expressed in variance, used for calculating the probability of not realizing an objective.

As demonstrated, the use of multiple concepts of risk does not form a barrier for understanding the total risk. Perhaps it is even the other way around, that for understanding the total risk in a system it is vital to use multiple concepts of risk. This in essence was already advocated by the Stern et al. (1996), in their opinion that relevance to a decision cannot be determined a priori by a formal definition of risk. It also aligns with the constructive school: risk is an imposed structure upon reality to help in decision making about potential problems.

13.3 Acceptance

We found in the experiments on **portfolio optimization** (chapter 6) and **whole system optimization** (chapter 9) that acceptance of Risk Based Optimization was not a difficult issue. In the portfolio decision, some (initial) hesitation was encountered, but acceptance could be created with small modifications, like adjustments to the value system or bypassing the process for some interventions. In **whole system optimization** doubts were issued on the adequacy of the assumptions, but a (on-the-spot) change of those assumptions demonstrated the robustness of results and thus their acceptance. This acceptance of RBO is in line with what should be expected for normal risks. The presumption of risks being normal therefore seems to be confirmed.

However, this is not necessarily the case. As mentioned, the results of CBA were accepted after a discussion on their validity in the light of uncertainty, followed by modifications to the method of RBO or even a negotiated bypass of RBO. This is much more a discourse based strategy, several steps up the risk escalator. This suggests that either the risks were not entirely normal, or that CBA is not entirely the right approach for normal risks. A resolution may be found in the notion Klinke and Renn (2002) make that discourse is important in any decision approach. For the portfolio decision, the discourse was internal, as prescribed by the risk escalator for normal risks. In the whole system optimization, the discourse was more participatory, allowing participants to test their own assumptions on the problem and its resolution. This aligns with the notions of discourse for a Cassandra class risk, which asset ageing (being a problem in the far future) in a way is. Despite the small basis for reflection, the conclusion therefore seems justified that the experiments confirm the theory of the risk escalator.

However, the inclusion of deliberation into Risk Based Optimization sacrifices consistency of decision making. In literature, the relevance of consistency has been advocated by Morgan (1993) and Merkhofer (1987) for decision making in general and by Woodhouse (2014) specifically for asset management. The response to this concern is pragmatic. The corrections resulting from deliberation had only a small impact on the end result, especially given the uncertainties that inevitably surround the decisions. From a cost benefit perspective, it is therefore much better (cheaper!) to accept these small corrections than to debate them extensively or to refine the value system so that the RBO would exactly match the choice that was made. The allowance of inconsistency in decision making then is consistent with the whole idea of optimization, but from a slightly different perspective.

Building on the hypothesis that organized inconsistency is vital to acceptable decision making, a follow-on question would be on how to optimally organize those inconsistencies. Should it be by integrating approaches (as for example Merkhofer mentioned), should it be in parallel, or perhaps sequential? And how does this change in different settings?

In the experiments, the rational core of RBO was embedded in a social process. Vital seemed to be that deliberation did only interfere at the start of RBO (by value system and framing of risks) and in accepting the end result, and not at every stage in between. Furthermore, the setting effectively countered peer pressure and groupthink (Janis, 1982) . These are characteristics of any well-designed process, though not yet included in the asset management domain.

13.4 Value of risk based optimization

Three experiments in this thesis explicitly addressed the added value of risk based optimization. In the experiment on **single asset optimization** (chapter 5), we found a 20-30% improvement compared to (then) current practices. A similar result was found for the **optimization of the portfolio** (chapter 6). A ranked allocation of budget would improve the performance by some 20% compared by unranked serial decision making. In the experiment on **whole system optimization** (chapter 9), an improvement was found of about 20% on the system level, about 20% for optimization of the portfolio in times of scarce resources and about 20% in the optimization of individual assets. This is a very consistent finding with regard to the value of risk based optimization.

The results from the experiments are not entirely in line with the results we found in literature. In the introduction to this thesis, reference was made to the UK utility industry, where asset management was presented as the solution for dealing with income reductions in the range of 10-

50%. A similar number was encountered in the discussion of the cost developments in North sea Oil, where cost reductions of some 70% per unit were reported (Woodhouse, 2014). Unfortunately, the reported numbers in literature were about the value of implementing asset management, and not just on the value of Risk Based Optimization.

An explanation to this partial alignment was encountered in the experiment on **whole system optimization**. One of the strategies tested was a general improvement of efficiency in executing the interventions. This roughly doubled the value of the strategies to which it was applied. Doubling the improvement of just RBO as found by this thesis results in an overall improvement in the 50% region, very much in line with the reported improvements of implementing asset management. This suggests that applying RBO accounts for roughly half the value of asset management. This aligns with the importance of balancing costs, performance and risk as is widely advocated in asset management literature.

14 Conclusion

Following the liberalization of the Dutch energy markets, many Distribution Network Operators implemented asset management to deal with the foreseen income reductions. A central element in asset management is managing the risks associated with the asset base, often expressed as managing the balance of costs, performance, risk and opportunities.

When the concept of asset management is applied on the infrastructure for energy distribution, most of the management attention is on controlling risk. Most risks are regarded as normal risks that can be objectified. This raised the question whether the management of infrastructures for energy distribution could be regarded as a whole system cost benefit consideration with regard to risk (Risk Based Optimization, RBO). Managing risk, however, is a problematic concept. We found that, in literature there was no precise agreement on how to approach this. Behind some superficial differences a more fundamental conflict was hidden, on what a risk precisely is, and how a good decision should be made about risk. Many different definitions can be used for risk, with most of these recognizing the importance of uncertainty. The most important difference between the definitions was found to be that of risk as a concept and the way this concept is measured.

We have selected Cost Benefit Analysis (CBA) as the appropriate approach for decision making in asset management, after a thorough literature study. However, CBA is not sufficient to achieve a full risk based optimization of the asset base: theoretically it is possible to account for non-financial effects in CBA, but there is no generally accepted scheme for incorporating them into such analysis. Besides, CBA does not provide a means for evaluating risks on their importance as such. If risk evaluation criteria were available, the most important problems could be selected from an overview of problems. However there was no agreed method for generating such an overview. It was also not clear at the start of the research whether using a CBA for every decision would be accepted. And finally, it was not clear what the added value of Risk Based Optimization would be in an energy distribution infrastructure. A pivotal point in understanding these knowledge gaps was that the gaps are not independent in practice. For example, a very sophisticated value system may be generally accepted but may also be very difficult to apply. On the other hand, a very simple value system (e.g. only financial) may be easy to apply, but results may not be accepted.

Based on these knowledge gaps the central research question was:

To what extent is formal risk based optimization of the whole system feasible in managing assets of the infrastructure for energy distribution?

This central question was divided into 4 sub questions:

1. What is an adequate representation of the value system that facilitates both CBA as the selection of most important risks?
2. What is an adequate representation of the risk position?
3. What is the effectiveness of applying Risk Based Optimization by means of these adequate representations?
4. How robust is this effectiveness of Risk Based Optimization?

Several experiments were conducted to answer the questions. The findings on these experiments were presented and discussed in chapters 12 and 13. Here we will present the main conclusions for each sub question.

1.What is an adequate representation of the value system that facilitates both CBA as the selection of most important risks?

The experiments revealed that a fairly simple method to represent the value system was found to be the risk matrix, provided that some basic rules are followed in the design of the matrix:

1. Logarithmic scales to align with the orders of magnitude that impacts and probabilities span. Additionally, these scales assure a constant relative uncertainty in estimation.

2. Alignment of values based on the severity of impacts. If impacts score in the same severity category, decision makers should be indifferent. For impacts measured on a quantitative scale, the alignment should be constant in the per unit tradeoff: more impact at the same time does not increase the value per unit.

3. Risk level based on expected values, also expressed on a logarithmic scale.

The matrix of Figure 71 holds such a robust representation. It contains three values, 6 severity classes, 6 likelihood categories and 6 risk levels. This is an adequate representation of the value system because:

- The values indicate what the decision maker cares about.
- The severity classes indicate what the decision maker considers equally bad. The risk levels indicate what problems should be addressed first.
- The per unit alignment allows pricing of risk, to be used in cost benefit analysis of mitigations.

The values finance, safety and reliability were found to be most relevant for the distribution infrastructure for electricity and gas. In some cases and for some risks, other values were found to be of importance, like compliance (e.g. when regulations change), sustainability and reputation. It was found that ’reputation’ could be used as an umbrella value for everything not captured by finance, safety and reliability.

One of the key conclusions of this research is that the value system does not need to be complicated. Since optimizations are by nature balance points, our research consistently showed that being slightly off does not impact the total value substantially. We proved that as long as the risk is captured, it is not very relevant that all of the impact is captured. Even at some low values such as 50%-75% of the value at risk being captured by the value system, the system optimization resulted in more than 90% accurate outcomes.

2. What is an adequate representation of the risk position?

Representing risk as a process of cause, asset, reaction and consequence, was shown to structure the large number of risks into a limited set while still providing an adequate representation of the risk position. We showed that the risks in this limited set do not need to be framed all in the same phase of the risk process. The only requirement was found to be that the risks are more or less independent so that summing their expected values does not result in double counting.

Our conclusion challenges literature in the sense that it generally is assumed that risks registers need to contain large volumes of risks to appropriately capture the risk position. We found however, that at any level of decision making a limited set will suffice, but that the aggregation level of the risks will change: in modeling the whole system performance, a small number of highly aggregated risks form an adequate representation. In modeling an individual asset (e.g. for designing the maintenance concept) the risks typically are the failure modes of the constituting parts. However, in understanding the system performance, we showed convincingly that such level of detail is not necessary.

3. What is the effectiveness of applying Risk Based Optimization by means of these adequate representations?

The effectiveness of applying RBO, measured in the cost improvements gained and the acceptance of its results, was found to be high. RBO reduced the Total Cost of Ownership of the assets by some 20-30% with wide-spread acceptance of results. This value was found for optimizations on individual assets, for optimization of the portfolio of interventions and for the optimizations of system level strategies. In this research, RBO was used as a starting point or input for the actual decision making. The actual decision was made in a social decision making process in which the RBO results were widely accepted with minor exceptions. This results in the somewhat paradoxical conclusion that RBO was found to be very effective as long as it was not presented as providing the right answer from the outset without human consideration and discussion.

4. How robust is this effectiveness of Risk Based Optimization?

The effectiveness of risk based optimization was found to be very robust. We measured robustness by several dimensions. First, optimization was found to be able to handle numerical errors quite well. If the value system was not entirely right (only capturing 50%-75% of the value at risk), or in cases where the failure mechanism were not understood fully, optimization could still produce very (>90%) accurate answers. This robustness seems to be a fundamental property of our RBO. Many of the optimums were found to be quite flat (shallow). The outcome of optimization was therefore more like a window of opportunity than an exact point in time or an exact portfolio content.

The second dimension of robustness was found at higher levels of aggregation. Even when the random error in the assumptions for individual interventions was significant (say 20%), this almost disappeared when considering a set of interventions, simply because the errors cancelled each other out.

The third dimension of robustness we considered was about tolerance for errors in the decision approach. In our experiments on the portfolio decisions, only very small parts of the RBO outcomes were rejected by decision makers, typically less than 5% of the budget. Even if these personal interventions would not have any value at all, we found that the (potentially high) cost of fine-tuning the RBO process would not be justified.

The fourth dimension of robustness we considered is the ability to adapt the methods to achieve an even better performance. The tools developed in this research are flexible and scalable by nature. The risk matrix can be adjusted if the decision requires, either by realigning values, changing the resolution and scope, or changing the values of interest. In constructing the risk position, risks can be framed in a different phase of the risk process, allowing other interventions to become valuable. Additionally, it is possible to move towards a more detailed representation. The method of RBO thus is both flexible and scalable. This is further supported by the fact that the same basic concept was used at all discussed levels of aggregation.

To what extent is formal risk based optimization of the whole system feasible in managing assets of the infrastructure for energy distribution?

This research has shown that the answer to this question is that there are no fundamental barriers to risk based optimization of the whole system. This means that our conclusion is that optimization of the whole system is feasible to a very large extent (if not fully) for asset bases with predominantly normal risks. We argued that this conclusion holds because the rational core of Risk Based Optimization was connected in a flexible way to the social decision making process, to address ever present subjectivity and uncertainty. We operationalized this flexibility by allowing (re)structuring of the risk position, adjusting the representation of the value system and even bypassing RBO for certain decisions. Recognizing and dealing with subjectivity within the asset management process increased acceptance.

Contrary to common belief that uncertainty in decision making needs to be battled with more detailed and more complicated models, we found that explicit inclusion of uncertainty allows for relatively simple tools. The outcomes of optimization with these simple tools were not perfect and indisputable, but accurate enough given the uncertainties that inevitably surround real life decisions. And they were accurate enough to form a strong basis of asset management decision making.

After all, risk management is not about being right all the time, but allowing things to go wrong where they do not matter much.

15 Recommendations

Based upon this research many recommendations can be made for both practice and science. In this chapter, the most important ones will be listed and briefly discussed. In chapter 16 we will provide a more detailed reflection on how these recommendations were identified. First the most relevant recommendations for further research will be listed, followed by the most relevant recommendations for practice.

15.1 Recommendations for further research

1. Repetition of this research in other asset bases, at least in other infrastructures. The research could focus on whether the presented value system can be adapted to fit the decisions, whether the risk position of that asset base can be captured using the risk process, and whether optimization (with organized inconsistencies!) would deliver similar value both for assets, portfolios and systems. If the results can be somehow replicated, it is a strong indication of general applicability of the findings, whereas different results could help understand the critical conditions for optimization to work.

2. A more fundamental research regarding a quantitative value system used in general decision making. Even though the three-value system worked in this research, its application in other industries is impossible without at least adaptation of the reliability metric. But what should that be, and how should it be priced? Perhaps the research Tengs et al. (1995) did can be repeated for other values like sustainability, compliance, reputation and so on. Or research for other infrastructures could be done, e.g. what is the value for users of being out of telecom, roads and so on. These then all could be combined into a generally accepted risk matrix covering and aligning many values. Asset managers then could use this as a reference, selecting the relevant values and scaling them to fit their organization. If a simple system as developed in this thesis works this well for a large infrastructure operated in the public domain, perhaps a system to capture all relevant values in the world is not that much more complicated. Just being able to understand why people care can help in finding common ground. It must be emphasized that this is not about optimizing the world, but about creating a common language why we care about certain aspects of the world.

3. Research on the automated structuring of a rich list of risks, preferably with some quality measures of the achieved level of independence and possibilities to compare different numbers of risk to construct the risk position. An idea that somehow did not surface until late in the research was that the risk position can be represented with a limited number of risks. Both in measuring risk and in replacement optimization it was limited to a few dozen risks, but in capturing the risk positon several hundreds were reviewed. Perhaps the right approach would have been to plot all identified risks in a (Bayesian) relational diagram and to identify more or less independent clusters. It would require a thorough review of potentially supporting tools.

4. Research on the integrating of fundamentally conflicting decision approaches in organizations of different size and complexity. A very intriguing area identified by this thesis is the acceptance of inconsistency in decision making. The process to do so may be radically different for small organizations than for large ones. But the whole idea of inconsistency itself may be very difficult to research, as the organizations may reject the whole notion of inconsistency. Many advocate the

importance of consistent decision making (one version of the truth!) and letting go of that concept may be very challenging.

5. Research about the (non)normality of risks and the way to express that. Research into non-normality becomes important with the advent of smart grids. Such smart grids will coordinate the behaviours of the asset base after which it will no longer behave randomly, and will stop to exhibit normally distributed failures. As discussed, it seems that very many non-normal risks still have a normal core, making non-normality an attribute of risk. But that indicates that the amount of non-normality is measurable and that a more precise match of decision approaches can be made than just with thresholds. It would also allow assessment of non-normality before decision making, and thus decrease the potential errors.

6. Research on the definition of a theoretically absolute optimality. If such a measure could be created, the quality of asset management organizations could be expressed much more direct than by means of maturity or compliance with standards. In this thesis it was proven that even under uncertainty optimization could reach very good results, especially in terms of the achievable optimum given the optimization variable (the relative optimality). However, in that optimization still some aspects would be fixed manually. And even if changing these would not significantly alter the decision, It would not mean that the total cost of ownership could not be lowered. That could mean that it may be possible to determine a theoretically absolute optimality, at least for individual assets, but perhaps even for whole systems. This is illustrated in Figure 73.

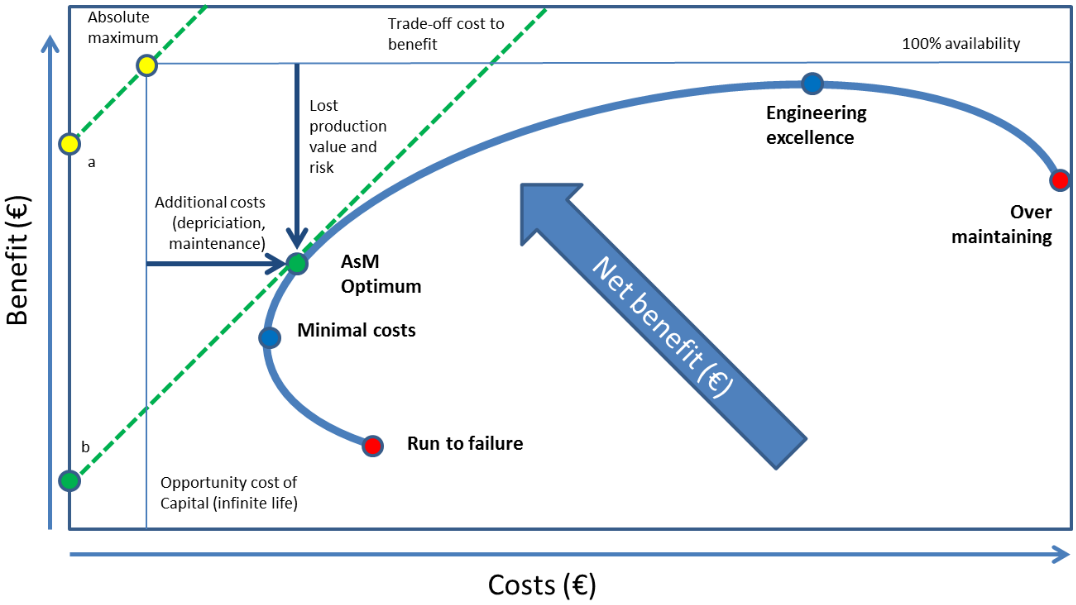


Figure 73: The idea of absolute optimality visualized, after Wijnia (2015).

The blue line represents the optimization trajectory, with run to failure (i.e. infinite planned interval) at one end and over maintaining (i.e. very short intervals) at the other extreme. Starting at run to failure, adding a little bit of maintenance will prevent some faults. Assuming repairing faults is more costly (in real cash) than preventive actions this will reduce the costs. At a certain moment the rise of planned costs will equal the reduction of unplanned costs. This is the cost minimum. This does not need to coincide with the maximum availability (and thus production time and gross benefit). That will be reached when the increase of planned outage time equals the reduction of unplanned outages. This is a second optimum, generally regarded as engineering excellence. The asset management optimum is somewhere in between, where the increase of the equivalent value of planned cost and downtime equals the reduction of the unplanned counterparts. At this point, the asset produces the highest net benefit, point b on the benefit axis. The highest net benefit the asset could ever deliver is when there is no additional cost outside the investment and the availability is 100%, represented by point a. The theoretically absolute optimality of the asset then can be represented by the ratio between achieved net benefit and the theoretically possible net benefit.

15.2 Recommendations for practice

1. Optimization should be part of the definition of asset management. "Coordinated activities to realize value from assets", as it is currently defined in ISO 55000, is not precise enough to specify asset management. Any asset that is not abandoned would qualify as being managed according to this definition. This clearly is not in line with the ideas as ventilated in the past 40 years or so of asset management science. A simple revision would be to include optimal in the current definition: Asset management is the coordinated activity to realize optimal value from assets.

2. A practical recommendation to asset managers is to focus more on the certainty of the decision being right than on the accuracy of numbers used in the decision. Uncertainties in the assumptions surrounding the decision may cause significant changes in the predicted costs and benefits, but if all alternatives are sensitive in the same way this uncertainty does not influence the optimal decision. Asset management is about avoiding the paradox of choice: trying to rationalize the decision when in reality it does not really matter which option is chosen.

3. Given the relative simplicity of optimization, it may be very useful for practice to develop a public set of reference models for value systems, risk registers, asset optimization models, portfolio decision tools including social choice, and whole system optimization concepts. Many of these exist in the commercial domain, but are difficult to challenge because of proprietary aspects and thus may be fundamentally wrong.

4. For the regulation of energy distribution companies in the Netherlands, there is no argument why there could not be a value system prescribed by the regulator. Even the central definition of the risks that span the risk position should be possible. For quality regulation then it could be enough that the asset managers prove that they optimize the risks according to the value system.

16 Reflection

16.1 On the approach

The approach followed in this thesis was basically a reversal of the risk management process. The argument used in chapter 4 was that such a straightforward approach would not reach a natural end point. Only by understanding the decision to be made, it could be determined what level of understanding of the risk position was needed and what kind of a value system would facilitate optimization. Given the results we achieved with this research, the reversed approach was effective. However, it does not immediately confirm that it was a necessity, though support can be found in this research. Most relevant in that respect is the finding that a relatively simple representation of the value system that captures only 50-75% of the value at risk still produces near optimal decisions. This finding can only be reached by starting from the decision. A similar argumentation can be found in representing the risk position. Only by starting from the total value at risk it can be derived that a couple of dozen of risks are enough for an adequate representation. These are strong arguments for a more exploratory reversed approach.

Contradicting this argumentation is that these conclusions could have been postulated at the start of this research on the basis of practical feasibility as hypotheses: it is possible to represent the value system with a “limited resolution” risk matrix, and it is possible to represent the risk position with a limited set of risks. These hypotheses could have been tested in the straightforward approach, removing the need to start with the decision. Given the state of knowledge at the start of this research project these would have been bold statements, but they were not entirely out of reach. Yet, whether that would have resulted in other experiments is unlikely. The state of knowledge did not allow for omitting an exploratory phase.

What seems fair to state though is that for similar research projects in other (similar) asset bases there is no need for a reversed approach. Starting with a relatively simple risk matrix and identification of the top 20 risks followed by risk based decision making on the interventions allows for a quick validation of the conclusions of this thesis. For asset bases with much less normal risks, in which the conclusions cannot be assumed to be right, the reversed approach probably still would be the better option.

That the reversed approach was effective does not mean it could not have been conducted more efficiently. The key insight with regard to robustness was touched upon in each and every decision made. However, this did not immediately result in the conclusion that this was a fundamental property of optimization, and not just a fortunate coincidence for the experiments conducted for this research. A more fundamental sensitivity analysis on all relevant aspects probably would have revealed that much earlier.

16.2 On inconsistency

One of the difficulties in this research was that practice requires flexibility with regard to the view on reality. Stakeholders and their perceived risks simply cannot be excluded on the basis of a deviation from the agreed upon view, because then they will be forced to find another way to influence decision making (hence the umbrella value reputation). Pragmatic solutions can be found, as was demonstrated by the flexible connection of RBO to decision making. However in research the inherent inconsistencies from applying two conflicting theories at the same time in an entangled hierarchy is much more troublesome. Perhaps the methodologies to study complex systems could have helped, in which the need for more than one perspective is recognized (Nikolic, 2009).

16.3 On generalization

This research was conducted for the asset base of Enexis, and strictly speaking our conclusions have no further reach than this asset base. However, as the asset base of Enexis is not uniform but a collection of many constituting smaller asset bases, it may be assumed that the conclusions hold for any collection of small asset bases in the same sector, that is, for the entire energy distribution sector in the Netherlands.

Arguments can be found for further generalization. Limitation of the value system to finance, safety and reliability may very well apply to any passive asset as the asset will only be “noticed” if it is not functioning properly. Normality of risks may be a characteristic of any matured asset base because of the familiarity with the risks in such an asset base. The robustness of RBO seems to be a fundamental property of an optimum. Mixing decision approaches is relevant in any decision situation with a variety of views on reality, i.e. many different stakeholders. The conditions under which the conclusions hold therefore may apply to a wide range of asset bases, at least to infrastructures in the public domain.

However, this only holds for the asset bases in their matured state. If the asset base is confronted with new forms of use or new technologies, the normality of risk may no longer hold, and thus the conclusions of this thesis. An example may be the introduction of smart technologies (Verzijlbergh, 2013, Veldman, 2013) which could result in completely different interconnected failure modes. Managing such an interconnected asset base may require a completely different (and more complicated) set of tools.

16.4 On normal risk

In this research project, the explicit presumption was made (based on a literature review) that risks in the energy distribution infrastructure are normal and CBA is applicable. However, the conclusion that because CBA was accepted the risks must have been normal, does not hold. Perhaps the right question is if this presumption was needed at all. Suppose that the corrective actions of applying CBA would have been larger than the value of optimization, would that have been a reason to reject CBA for non-normal risks? That would only be the case if the non-normality of risks could have been assessed with certainty beforehand. If not, then apparently all encountered risks have to be addressed with the same approach. The question then should be more on what the best approach is for a combination of normal and non-normal risks. This may very well depend on the fraction of non-normal risks. However, given the costs of other decision methods it also may be that even for portfolios only consisting of non-normal risks the best approach still would be to start with CBA.

16.5 On continual improvement versus optimization- a silent paradigm change?

In hindsight, a recurring motive in this research was the competition between the viewpoint of continual improvement and that of optimization. Given the extent to which risk based optimization is possible in asset management at any level of decision making, it is remarkable that it receives so little attention in literature. It is as if the whole concept of optimization has been replaced with the concept of continual improvement, which is at the heart of the management system standards. However, whereas optimization has a best answer (at least in the form of an window of opportunity that is good enough), continual improvement lacks such an end point.

This change of paradigm may be an indication that in practice continual improvement is more productive. Yet, in scientific terms such a silent change of paradigm is very concerning. If positions change, it should be because of fierce and loud debates, accompanied by very conclusive experiments. If such an experiment is not possible, the debate should at least focus on understanding the conditions under which the viewpoints are more valuable. It may e.g. just be that optimization holds in stable environments, like the management of infrastructures, and that continual improvement is much better in volatile environments.

In a way, this thesis demonstrated that optimization and continual improvement go very well together. The research itself, being experimental and exploratory, is an example of continual improvement. Each experiment resulted in a better understanding, resulting in a design for a new experiment. Many of the outcomes of the experiments were also examples of continual improvement, or at least, the results were acceptable because they lent themselves for future improvement. Single asset decisions could be improved with a better understanding of the value system, the portfolio decision could be improved by modeling the risk for more programs, the safety indicator could be improved by adding more incident data. Yet, continual improvement is also a bit cautious or even fainthearted. Any coordinated improvement is continual improvement, but in practice the rate of progress would seem to be a vital factor. Furthermore, the viewpoint of continual improvement can be very wrong if the assumption is made that past improvements can be repeated in future. The example given in the introduction was benchmark only regulation resulting in income cuts for the network operators. If the profitability was increased by mortgaging the future, the resulting income cut in the next period would induce the company to do so even more, resulting in a potential spiral of decline. Other examples are assets that meet some natural constraint, like passenger airplanes flying slightly slower than the speed of sound or power plants operating at a maximal efficiency dictated by the Carnot cycle. Though it has to be recognized that in some parts of technology continual improvement seems limitless, as demonstrated by Moore’s Law for computer capacity.

The viewpoint of optimization is better equipped for dealing with those natural constraints. In a power plant, a good measure for the quality of the plant would be the achieved thermal efficiency versus the theoretically possible Carnot efficiency given the used temperatures. But the absolute efficiency is a valuable indicator as well. Yet, optimization as a viewpoint also has its drawbacks. A significant one may be organizational. If an asset is operated at its optimum, there is little incentive for trying to improve the way of operating the asset. Optimization thus may generate very lazy asset managers. In a way, the existing technical standards could be regarded as optimal intervention in a different era, which simply never were adjusted to changing opinions in the stakeholders because they were considered to be optimal. A more theoretical drawback is that the optimum is time and context dependent and relies as much on subjective value judgments as is does on “objective” assessments. The optimum in a multi value system does not exist as fact of the world, so how could optimization ever result in an undisputed answer? Furthermore, technological developments or scientific breakthroughs may push the boundary of what is achievable beyond the wildest fantasies existing today. Again, what does optimality mean if the reference point is not fixed in any way?

Which one of these viewpoints is better, more true or more valuable cannot be answered by this thesis. Continual improvement was a necessity for getting at the results achieved. Yet, the assumption of the existence of an optimum was much more useful in designing the experiments, as it opened up a whole range of questions about what the optimum was, how it best could be approached and how certain it was that the resulting answer actually was the best that was possible. Given that both viewpoints provided a valuable contribution to this thesis, there may not even be an ultimate answer.

However, for a healthy debate on the pros and cons of optimization and continual improvement, much more attention should be given to optimization. Several major areas for further development of optimization in the science of asset management are out there to be researched by the community.

LITERATURE AND PUBLICATIONS

17 Literature

Abraha, H. & Liyanage, J. 2015. Review of Theories and Accident Causation Models: Understanding of Human-Context Dyad Toward the Use in Modern Complex Systems. *In:* Lee, W. B., Choi, B., Ma, L. & Mathew, J. (eds.) *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012).* Springer International Publishing.

ACM. 2014a. *Bijlagen x-factormodel regionaal netbeheer elektriciteit 2011-2013:* [Online]. Available: https://www.acm.nl/nl/publicaties/publicatie/13305/Bijlagen-x-factormodel-regionaal-netbeheer-elektriciteit-2011-2013/ [Accessed January 12 2015.

ACM 2014b. Netcode Elektriciteit *In:* Acm (ed.).

Ajah, A. N. 2009. *On the Conceptual Design of Large Scale Process & energy Infrastructure Systems: Integrating Flexibility, Reliability, Availability, Maintainability and Economics (FRAME) performance metrics.* Ph.D., Delft University of Technology.

ALM-43-7494-C Reliability Centered Maintenance.

Amadi-Echendu, J. E., Willett, R., Brown, K., Hope, T., Lee, J., Mathew, J., Vyas, N. & Yang, B.-S. 2010. What Is Engineering Asset Management? *In:* Amadi-Echendu, J. E., Brown, K., Willett, R. & Mathew, J. (eds.) *Definitions, Concepts and Scope of Engineering Asset Management.* Springer London.

Australian Energy Regulator 2014. State of the energy market 2014.

Aven, T. 2009. Misconceptions of Risk. *Misconceptions of Risk.* John Wiley & Sons, Ltd.

Aven, T. 2010. On how to define, understand and describe risk. *Reliability Engineering & System Safety,* 95**,** 623-631.

Aven, T. 2011. On the new ISO guide on risk management terminology. *Reliability Engineering & System Safety,* 96**,** 719-726.

Aven, T. 2012. The risk concept—historical and recent development trends. *Reliability Engineering & System Safety,* 99**,** 33-44.

Aven, T. & Renn, O. 2009. On risk defined as an event where the outcome is uncertain. *Journal of Risk Research,* 12**,** 1-11.

Aven, T., Renn, O. & Rosa, E. A. 2011. On the ontological status of the concept of risk. *Safety Science,* 49**,** 1074-1079.

Baarsma, B. E., Berkhout, P. H. G. & Hop, J. P. 2004. Op prijs gesteld, maar ook op kwaliteit: de prijs van stroomonderbrekingen- op zoek naar phi,. Amsterdam: SEO.

Bakir, V. 2005. Greenpeace v. Shell: Media exploitation and the Social Amplification of Risk Framework (SARF). *Journal of Risk Research,* 8**,** 679-691.

Beer, T. & Ziolkowski, F. 1995. Environmental risk assessment: an Australian perspective. Barton: Supervising scientist.

Berndsen, d. R. C. D. M., van der Rijken, d. m. T. & Wiersema, S. C. M. 2012. Tariefregulering in retrospectief: Inventariserend en structurerend feitenonderzoek. Berenschot.

Brealey, R. A. & Myers, S. C. 2000. *Principles of corporate Finance,* New York, McGraw-Hill.

Brown, R. E. & Humphrey, B. G. 2005. Asset management for transmission and distribution. *IEEE Power and Energy Magazine,* 3**,** 39-45.

BSI 1964. British Standard BS 3811: 1964 - Glossary of Maintenance Management Terms in Terotechnology London.

BSI 2004a. PAS55-1 Asset Management. *Part 1: Specification of the optimal management of physical infrastructure assets.* London.

BSI 2004b. PAS 55-2 Asset management *Part 2: Guidelines for the application of PAS 55-1.*

BSI 2008a. PAS 55-1:2008 Asset Management. *Part 1: Specification for the optimised management of physical assets.*

BSI 2008b. PAS 55-2:2008 Asset Management. *Part 2: Guidelines for the application of PAS 55-1.*

Burns, P. 2010. *The History Project* [Online]. Available: https://www.amqi.com/index.php/historyproject [Accessed june 24 2011].

Carr, M. J., Konda, S. L., Monarch, I., Ulrich, F. C. & Walker, C. F. 1993. Taxonomy-based risk identification. DTIC Document.

Carreras, B. A., Lynch, V. E., Dobson, I. & Newman, D. E. 2003. Blackout Mitigation assessment in power transition systems. *Thirty-sixth Hawaii International Conference on System Sciences.* Hawaii.

Claes, P. F. 2001. *Risico Management,* Leiden, Stenfert Kroese.

Clark, P. S. 2004. Inquiry into national workers' compensation and OH&S frameworks. *Workplace Safety and Health Management*.

Cohen, K. J. & Hammer, F. S. 1967. APPLICATIONS OF FINANCIAL THEORY - LINEAR PROGRAMMING AND OPTIMAL BANK ASSET MANAGEMENT DECISIONS. *Journal of Finance,* 22**,** 147-165.

COSO 2004a. *Enterprise risk management- Integrated Approach: Executive summary*.

COSO 2004b. Enterprise Risk Management-Integrated Framework Executive summary. Committee of Sponsoring Organizations of the Treadway Commission.

Council of European Energy Regulators 2015. CEER Benchmarking Report 5.2 on the Continuity of Electricity Supply Data update. CEER.

Cox, L. A., Jr. 2008. What's Wrong with Risk Matrices? *Risk analysis,* 28**,** 16.

Cox, L. A., Jr. 2012. Why Frequency is Not Well Defined for Engineering Systems with Nonexponential Failure Times. *Risk Analysis,* 32**,** 368-372.

Crespo Marquez, A. & Gupta, J. N. D. 2006. Contemporary maintenance management: process, framework and supporting pillars. *Omega,* 34**,** 313-326.

Cullen, L. 1990. Public inquiry into the Piper Alpha disaster, Vols. 1 and 2 HMSO.

de Bruijne, M. 2004. Reliability against the odds: California's electricity crises. *1st Annual CZAEE International Conference "Critical Infrastructure in the energy sector: Vulnerabilities and protection".* Prague.

de Vries, L. J. 2004. *Securing the public interest in electricity generation markets: the myths of the invisible hand and the copper plate.* Delft University of Technology.

Deloitte. 2010. *Cloud Computing Risk Intelligence Map* [Online]. Available: http://www.isaca.org/Groups/Professional-English/cloud-computing/GroupDocuments/Deloitte%20Risk%20Map%20for%20Cloud%20Computing.pdf [Accessed December 09 2014].

EGIG 2008. 7th EGIG report.

Electricity Association 2000. *Electricity industry review 4*, Electricity Association, Business Information Centre.

Emmanouilidis, C., Liyanage, J. P. & Jantunen, E. 2009. Mobile solutions for engineering asset and maintenance management. *Journal of Quality in Maintenance Engineering,* 15**,** 92-105.

Energiegids, D. 2013. *Netbeheerders* [Online]. Available: http://www.deenergiegids.nl/marktpartijen/Netbeheerders.aspx [Accessed 11 january 2013].

Enexis 2011. Kwaliteits- en Capaciteitsdocument Gas 2012-2021 deel 1.

Enexis 2015. Kwaliteits- en Capaciteitsdocument Elektriciteit 2016-2025.

Essent Netwerk 2007. Kwaliteits- en Capaciteitsdocument Elektriciteit 2008-2014.

Essent Netwerk Noord 2002a. gebeurtenissen bij file 110.

Essent Netwerk Noord 2002b. Validatie risico's 20021021 v3 met uitwerking resultaat.

European Committee for Standardization 2001. EN 13306:2001 Maintenance Terminology. Brussels.

Feng, D., Gan, D., Zhong, J. & Ni, Y. 2007. Supplier asset allocation in a pool-based electricity market. *IEEE Transactions on Power Systems,* 22**,** 1129-1138.

Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., Combs, B. & 1987. How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences,* 9**,** 127-152.

Frey, D., Herder, P., Wijnia, Y., Subrahmanian, E., Katsikopoulos, K. & Clausing, D. 2009. The Pugh Controlled Convergence method: model-based evaluation and implications for design theory. *Research in Engineering Design,* 20**,** 41-58.

Frey, D., Herder, P., Wijnia, Y., Subrahmanian, E., Katsikopoulos, K., Neufville, R., Oye, K. & Clausing, D. 2010. Research in engineering design: the role of mathematical theory and empirical evidence. *Research in Engineering Design,* 21**,** 145-151.

Frontinus, S. J. 97 AD. De Aquaeductu Urbis Romae (On the water management of the city of Rome). Translated by R. H. Rodgers. University of Vermont. 2003.

Garrick, B. J. 2012. Response. *Risk Analysis,* 32**,** 373-373.

Goel, H. 2004. *Integrating Reliability, Availability and Maintainability in Conceptual Process Design.* Ph.D., Delft University of Technology.

Grey, W. & Shi, D. 2001. Value Chain Risk Management. IBM T.J. Watson Centre.

Gulski, E., Smit, J. J. & Wester, F. J. 2005. PD knowledge rules for insulation condition assessment of distribution power cables. *IEEE Transactions on Dielectrics and Electrical Insulation,* 12**,** 223-239.

Haimes, Y. Y., Kaplan, S. & Lambert, J. H. 2002. Risk Filtering, Ranking, and Management Framework Using Hierarchical Holographic Modeling. *Risk Analysis,* 22**,** 383-397.

Hallikas, J., Karvonen, I., Pulkkinen, U., Virolainen, V.-M. & Tuominen, M. 2004. Risk management processes in supplier networks. *International Journal of Production Economics,* 90**,** 47-58.

Hanson, S. O. 2004. Fallacies of risk. *Journal of Risk Research,* 7**,** 353-360.

Hazelrigg, G. 2010. The Pugh controlled convergence method: model-based evaluation and implications for design theory. *Research in Engineering Design,* 21**,** 143-144.

Heimann, C. F. L. 1997. *Acceptable risks: Politics, Policy and risky technology,* Ann Arbor, The University of Michigan Press.

Heinrich, H. W. 1931. *Industrial accident prevention: a scientific approach* New York, McGraw-Hill, .

Henrard, L. 2004. Risk managament of a financial conglomerate. *Fortis Enterprise Risk management Symposium.* Chicago.

Herder, P. M. & Wijnia, Y. 2012. A Systems View on Infrastructure Asset Management. *In:* Van Der Lei, T., Herder, P. & Wijnia, Y. (eds.) *Asset Management.* Springer Netherlands.

Hermkens, R. J. M. 2005. Veiligheidsindicator Gasdistributienetwerken - Implementatie van de veiligheidsindicator. Apeldoorn: Kiwa-Gastec Technology.

Hermkens, R. J. M. & Pulles, C. J. A. 2007. Statuut veiligheidsindicator V01. Apeldoorn: KIWA

Hesselmans, A. N. 1995. *'De Ware Ingenieur': Clarence Feldmann, Delfts Hoogleraar en grondlegger van de provinciale elektriciteitsvoorziening,* Utrecht, Stichting Histosearch.

Hodkiewicz, M. R. 2015. The Development of ISO 55000 Series Standards. *In:* Tse, P. W., Mathew, J., Wong, K., Lam, R. & Ko, C. N. (eds.) *Engineering Asset Management - Systems, Professional Practices and Certification.* Springer International Publishing.

IAM 2002. *International Infrastructure Managament Manual*, UK Institute of Asset Management.

IAM 2006. Guidelines for Patrons & Terms of Reference for ‘The Patrons of the IAM. Institute of Asset Management.

IAM 2013. IAM Strategy 2014-19. Institute of Asset Management,.

IAM 2014. Asset Management - an anatomy Version 2. The institute of asset management.

IEC 1985a. IEC 812: Analysis techniques for system reliability- procedure for failure mode and effect analysis.

IEC 1985b. IEC 60853: Calculation of the cyclic and emergency current rating of cables. *Part 1: Cyclic rating factors for cables up to and including 18/30 (36) kV.*

IEC 2002. IEC 60287: Electric cables- calculation of the current rating. *Part 1-3: Current rating equations(100% load factor) and calculation of losses- current sharing between parallel single-core cables and calculation of circulating current losses.*

Institute of Asset Management 2008. Asset Management Part 1: Specification for the optimized management of physical assets. *In:* Bsi (ed.) *PAS 55-1.*

Institute of Municipal Engineering Australia 1994. National asset management manual. South Melbourne: The Institute, .

ISO 2002. ISO/IEC Guide 73: 2002.

ISO 2009a. ISO 31000: Risk Management- Principles and guidelines.

ISO 2009b. ISO/IEC 31010: Risk management: risk assessment techniques (final draft).

ISO 2014a. ISO 55000 *Asset Management-Overview, principles and terminology.* Geneva.

ISO 2014b. ISO 55001 *Asset Management-Management systems-requirements.* Geneva.

ISO 2014c. ISO 55002. *Asset Management-Management systems-Guidelines for the application of ISO 55001.* Geneva.

Janis, I. L. 1982. *Groupthink, Psychological Studies of Policy Decisions and Fiascoes,* Boston, Houghton Mifflin Company.

Johnson, W. G. 1973. The Management Oversight & Risk Tree - MORT SAN 821-2. U.S. Atomic Energy Commission, Division of Operational Safety.

Johnson, W. G. 1975. MORT: The Management Oversight and Risk Tree. *Journal of Safety Research*.

Johnson, W. G. 1980. *MORT safety assurance systems,* New York, M. Dekker.

Jonkman, S. N., van Gelder, P. H. A. J. M. & Vrijling, J. K. 2003. An overview of quantitative risk measures for loss of life and economic damage. *Journal of Hazardous Materials,* 99**,** 1-30.

Jonsson, P. 2000. Towards an holistic understanding of disruptions in Operations Management. *Journal of Operations Management,* 18**,** 701-718.

Kahneman, D. & Tversky, A. 1979. Prospect theory, an analysis of decision under risk. *Econometrica,* 47**,** 263-289.

KEMA 2004. Wensstromen: Gewenste kwaliteit- De waardering van de kwaliteit van levering van elektrische energie door aangeslotenen. Arnhem: KEMA.

Klinke, A. & Renn, O. 2002. A new approach to risk evaluation and management: Risk Based, precaution based and discourse based strategies. *Risk analysis,* 22**,** 1071-1094.

Knox, N. W. & Eicher, R. W. 1992. Mort User's Manual: For use with the Management Oversight and Risk Tree analytical logic diagram.

Komonen, K., Koppinen, T. & Kortelainen, H. 2008. A Strategic Asset Management Model: Determination of Corporate Strategy for Physical Assets. *Euromaintenance 2008.* Brussels, Belgium.

Komonen, K., Kortelainen, H. & Räikkönen, M. 2012. Corporate Asset Management for Industrial Companies: An Integrated Business-Driven Approach. *In:* Van Der Lei, T., Herder, P. & Wijnia, Y. (eds.) *Asset Management.* Springer Netherlands.

Komonen, K., Räikkönen, M., Kunttu, S., Heikkilä, A. & Ahonen, T. 2010. Investments, capacity and maintenance: ways to safely increase capital turnover. *Euromaintenance 2010.* Verona, Italy

Korn, M. S. & Veldman, E. 2008. The benefits of continuous risk management. *International conference on infrastructure systems: Building networks for a brighter future.* Rotterdam.

Körvers, P. M. W. 2004. *Accident Precursors.* Ph.D., Technische Universiteit Eindhoven

Krol, J. 2000. *Power to the people.* M.Sc. Master, Rijksuniversiteit Groningen.

Le Du, M., Rassineux, B. & Cochet, P. 2002. The French power network facing th 1999 storms. *Power Systems and Communications Infrastructures for the future.* Beijng.

Lee, J. 2007. *Off Goes the Power Current Started by Thomas Edison* [Online]. Cityroom New York Times. Available: http://cityroom.blogs.nytimes.com/2007/11/14/off-goes-the-power-current-started-by-thomas-edison/ [Accessed 2010 04 29.

Leitch, M. 2010. ISO 31000:2009—The New International Standard on Risk Management. *Risk Analysis,* 30**,** 887-892.

Lerner, E. J. 2003. What's wrong with the electricity grid. *Industrial Physicist.*

Löfstedt, R. E. & Renn, O. 1997. The Brent Spar Controversy: An Example of Risk Communication Gone Wrong. *Risk Analysis,* 17**,** 131-136.

Lynn, M. 2002. BPs Approach to Operations Excellence. *Performance Excellence in the Chemical Iindustry, the 23rd annual European AIChE Colloqium.* The Hague.

Marx, I. 2010. Combining the best of both worlds. *IEEE Industry Applications Magazine,* 16**,** 30-34.

Masera, M., Wijnia, Y. C., de Vries, L. J., Kuenzi, C., Sajeva, M. & Weijnen, M. P. C. 2006. Governing Risks in the European Critical Electricity Infrastrcuture. *In:* Gheorghe, A. V., Masera, M., De Vries, L. J. & Weijnen, M. P. C. (eds.) *Critical Infrastructures at Risk, Securing the European Electric Power System.* Dordrecht: Springer.

McGill, W. L., Ayyub, B. M. & Kaminskiy, M. 2007. Risk Analysis for Critical Asset Protection. *Risk Analysis,* 27**,** 1265-1281.

McNichol, T. 2006. *AC/DC: The Savage Tale of the First Standards War*, Wiley.

Meadows, D. H., Meadows, D. L., Randers, J. & Behrens, W. W. I. 1972. *The Limits to growth: a global challenge*, Club of Rome.

Merkhofer, M. W. 1987. *Decision sciences and Social Risk Management: A comparative approach of cost-benefit analysis, decision analysis and other formal decision-aiding approaches,* Dordrecht, D. Reidel Publishing company.

Ministry of commerce of New Zealand 1998. Inquiry into the Auckland Power Supply Failure.

Morgan, M. G. 1993. Risk analysis and management. *Scientific American,* 269**,** 32-35, 38.

Morgan, M. G., Florig, H. K., Dekay, M. L. & Fischbeck, P. 2000. Categorizing risk for risk ranking. *Risk analysis,* 20**,** 49-58.

Morgan, M. G. & Henrion, M. 1992. *Uncertainty- A guide todealing with uncertainty in Quantitative Risk and Policy Analysis,* Cambridge (UK), Cambridge University Press.

National Research Council 1995. *Technical Bases for Yucca Mountain Standards,* Washington, DC, The National Academies Press.

Nemhauser, G. L. & Wolsey, L. A. 1988. *Integer and combinatorial optimization*, Wiley-Interscience.

NEN 2009. NTA 8120:2009 Assetmanagement - Eisen aan een veiligheids-, kwaliteits- en capaciteitsmanagementsysteem voor het elektriciteits- en gasnetbeheer.

Netbeheer Nederland 2014. Energietrends 2014. Netbeheer Nederland.

New South Wales, Public Works Department. Policy Division, Premier's Department, Public Works Department, Treasury & Capital Works Committee 1993. Total asset management manual : capital works investment. Sydney, Australia: Policy Division, Public Works NSW

Newman, M. E. J. 2003. The structure and functionof complex networks. *SLAM Review,* 45**,** 167-256.

Nichols, D. L. 1973. Mishap Analysis: An Improved Approach to Aircraft Accident Prevention. *Air University Review,* July – August.

Nikolic, I. 2009. *Co-evolutionary method for modelling large scale socio-technical systems evolution.* Doctor, TU Delft, Delft University of Technology.

NMA 2007. 102610_Methodebesluit_voor_de_regionale_netbeheerders_elektriciteit_vierde_regulerings priode. *In:* Nma (ed.).

Noordwijk Risk Initiative Foundation 2009. MORT User's Manual for use with the MORT analytical logic diagram,. Delft: NRI.

Office of Management and Budget 2001. Spec 24 Ranking regulatory investments in public health. *Treasury and General Government Appropriation Act.*

Overbeeke, P. v. 2001. *Kachels, geisers en fornuizen.* PhD, Technische Universiteit Eindhoven.

Overijssel, H. C. 2013. 0735 Gasfabriek en Waterleiding- en Bosbedrijf van Zwolle, 1.1.1.8. Opheffing gasbedrijf. laatste wijziging 28-02-2013 ed.

Perrow, C. 1984. *Normal accidents: Living with high risk technologies,* Princeton, Princeton University Press.

Perrow, C. 1999. *Normal Accidents: Living with high-risk technologies,* New Jersey, Princeton University Press : [distributor] John Wiley and Sons Ltd.

Pulles, C. J. A. & van Eekelen, R. N. 2009. Bevindingen kwaliteitsterm Gas: Onderzoek naar de invulling van de kwaliteitsterm voor Gas. Apeldoorn: KIWA.

Reich, Y. 2010. My method is better! *Research in Engineering Design,* 21**,** 137-142.

Rijksoverheid 2005. Regeling inzake tariefstructuren en voorwaarden elektriciteit_tcm7-15329. *In:* Zaken, M. V. E. (ed.) *WJZ 5001015.* Staatscourant 13 januari 2005, nr. 9 / pag. 11: Staatscourant.

Rijksoverheid 2008. Wet informatie-uitwisseling ondergrondse netten.

Rijksoverheid 2013. Regeling van de Minister van Economische Zaken van 8 februari 2013, nr. WJZ/ 12357329, tot wijziging van enkele regelingen in verband met uitvoering van het marktmodel. February 11 2013 ed.

RIVM 2003. Coping rationlly with risks. National Institute for Public Health and the Environment.

Saaty, T. L. & Vargas, L. G. 2001. *Models, methods, concepts & applications of the analytic hierarchy process / by Thomas L. Saaty and Luis G. Vargas,* Boston ; London, Kluwer Academic Publishers.

Saldaña, M. A. M., Herrero, S. G., Campo, M. A. M. d. & Ritzel, D. O. 2003. Assessing Definitions and Concepts Within the Safety Profession. *The international Electronic Journal of Health Education,* 6:1-9.

Shell Global Solutions 2002. RRM Risk and reliability Management: an Overview.

Short, T. 2002. Reliability indices. *T&D World Expo 2002.* Indianapolis.

Singh, B., Jukes, P., Poblete, B. & Wittkower, B. 2010. 20 Years on lessons learned from Piper Alpha. The evolution of concurrent and inherently safe design. *Journal of Loss Prevention in the Process Industries,* 23**,** 936-953.

Slovic, P. 1987. Perception of risk. *Science,* 236**,** 280-285.

Slovic, P. & Weber, E. U. 2002. Perceptions of Risk posed by extreme events. *Risk management Strategies in an uncertain world.* New York.

Stacey, R. D. 1990. *Dynamic Strategic Management for the 1990s - balancing opportunism and business planning,,* Londen, Kogan Page Ltd.

Stern, P. C., Fineberg, H. V. & Committee on Risk Characterization - National Research Council 1996. *Understanding Risk: Informing Decisions in a Democratic Society,* Washington D.C., The National Academies Press.

Sugden, R. & Williams, A. 1978. *The principles of practical Cost Benefit Analysis,* Oxford, Oxford University Press.

Tengs, T. O., Adams, M. E., Pliskin, J. S., Safran, D. G., Siegel, J. E., Weinstein, M. C. & Graham, J. D. 1995. Five-Hundred Life-Saving Interventions and Their Cost-Effectiveness. *Risk Analysis,* 15**,** 369-390.

Thackara, A. D. 1975. Terotechnology - What it is all about. *Chart Mech Eng,* 22**,** 88-90.

The Institute of Risk Management (IRM), The Association of Insurance and Risk Managers (AIRMIC) & The National Forum for Risk Management in the Public Sector (ALARM) 2002. A Risk Management Standard.

Tombs, S. & Whyte, D. 1998. Capital Fights Back: risk, regulation and profit in the UK offshore oil industry. *Studies in political economy,* 57**,** 73-101.

Tversky, A. & Kahneman, D. 1981. the framing of decisions and the psychology of choice. *Science,* 211**,** 453-458.

Tweede Kamer der Staten Generaal 2007. Aanhangsel van de Handelingen 623. Den Haag.

U.S. Energy Information Administration 2013. Electric power sales, revenue, and energy efficiency Form EIA-861 detailed data files Release Date for 2013: February 19, 2015 ed.

United Nations 1992. RIO DECLARATION ON ENVIRONMENT AND DEVELOPMENT. United Nations.

van Akkeren, R. & Wijnia, Y. C. 2005. AsM-PR-veiligheidsindicator Gas (AN)-c-I.6.doc (dutch). Essent Netwerk B.V.

van den Noort, J. 1993. De verrassend veelzijdige geschiedenis van overheidsbedrijven. *Historische Bedrijfsarchieven Openbare Nuts- en Communicatiebedrijven. Een geschiedenis en bronnenoverzicht.* Amsterdam.

Van der Lei, T., Wijnia, Y. C. & Herder, P. M. 2010. TOWARDS AN INTEGRAL ASSET MANAGEMENT FRAMEWORK USING ASSET CHARACTERISTICS, ASSET ENVIRONMENT AND LIFECYCLE PHASES AS LEADING PRINCIPLES. *the 5th World conference on engineering asset management* Brisbane, Australia: Springer.

Veldman, E. 2013. *Power Play.* PhD, Eindhoven University of Technology.

Verzijlbergh, R. 2013. *The Power Of Electric Vehicles.* PhD, Delft University of Technology.

Vlek, C. 2013. How Solid Is the Dutch (and the British) National Risk Assessment? Overview and Decision-Theoretic Evaluation. *Risk Analysis,* 33**,** 948-971.

Waeyenbergh, G. & Pintelon, L. 2002. A framework for maintenance concept development. *International Journal of Production Economics,* 77**,** 299-313.

Wang, J. X. & Roush, M. L. 2000. *What every engineer should know about risk engineering and management,* New York, Dekker.

White, E. N. 1975. Terotechnology (Physical asset management). *Mining Technology,* 57**,** 5.

Whiting, J. F. 2001. Risk Tolerability Framework - Developing and implementing a practical workable framework for your workplace;. *APOSHO.* Taipei, Taiwan.

Wijnia, Y. 2012. Asset Risk Management: Issues in the Design and Use of the Risk Matrix. *In:* Mathew, J., Ma, L., Tan, A., Weijnen, M. & Lee, J. (eds.) *Engineering Asset Management and Infrastructure Sustainability.* Springer London.

Wijnia, Y. & de Croon, J. 2015. The asset management process reference model for infrastructures *In:* Amadi-Echend, J. E. & Mathew, J. (eds.) *9th World conference on engineering asset management.* Pretoria, South africa: Springer.

Wijnia, Y., de Croon, J. & Liyanage, J. P. 2014a. Application of a Unified Reference Model Across Asset Types: Comparative Cases. *In:* Lee, J., Ni, J., Sarangapani, J. & Mathew, J. (eds.) *Engineering Asset Management 2011.* Springer London.

Wijnia, Y., de Croon, J. & Liyanage, J. P. 2014b. Towards an Asset Management Reference Model: Basis for a Unified Approach. *In:* Lee, J., Ni, J., Sarangapani, J. & Mathew, J. (eds.) *Engineering Asset Management 2011.* Springer London.

Wijnia, Y. & Korn, M. 2007. Het risicoproces en het risicoregister inclusief demonstratie. Arnhem, Themadag Asset management 1 2007.

Wijnia, Y. C. 2004. The Challenges of a risk based asset management organisation. *Strategic asset management***,** 832-837.

Wijnia, Y. C. 2005. Risky business. *Assets.*

Wijnia, Y. C. 2007. The Human factor in Asset Management Process development. *2nd World Conference on Engineering Asset Management (WCEAM 2007).* Harrogate, United Kingdom.

Wijnia, Y. C. 2009. Asset Management for Infrastructures in Fast Developing Countries. *Infrastructure systems 2009: Developing 21st century Infrastructure Networks.* Chennai.

Wijnia, Y. C. 2010. Historie Risicomatrix. Rosmalen: Enexis.

Wijnia, Y. C. 2013. Dealing with uncertainty in the asset replacement decision. *In:* Mathew, J. & Tse, P. (eds.) *8th World conference on engineering asset management.* Hong Kong: Springer.

Wijnia, Y. C. 2014. Tactical Risk Management *CESUN, 4th International Engineering Systems Symposium.* Hoboken, New Jersey.

Wijnia, Y. C. 2015 Towards quantification of optimality in asset management (to be publsihed). *WCEAM2015.* Tampere, Finland.

Wijnia, Y. C. & Herder, P. M. 2004. Modeling Interdependencies in electricity infrastructure risk. *1st Annual CZAEE International Conference "Critical Infrastructure in the energy sector: Vulnerabilities and protection".* Prague.

Wijnia, Y. C. & Herder, P. M. 2005. Options for real options: Dealing with uncertainty in investment decisions for electricity networks. *International conference on Systems, Man and Cybernetics.* Hawaii.

Wijnia, Y. C. & Herder, P. M. 2009. The State of Asset Management in the Netherlands. *World Conference on Engineering Asset Management.* Athens: Springer.

Wijnia, Y. C. & Herder, P. M. 2013. Flipping a coin in infrastructure decisions. *In:* Brown, K. (ed.) *International Conference on Strategic Infrastructure Asset Management for Deltas 2013.* Sydney, Australia.

Wijnia, Y. C. & Hermkens, R. J. M. 2006. Measuring safety in gas distribution networks. *Proceedings of the 1st World Congress on Engineering Asset Management, WCEAM 2006.*

Wijnia, Y. C., Hermkens, R. J. M. & Flonk, J. 2010. The Safety Indicator: Measuring Safety in Gas Distribution Networks. *In:* Amadi-Echendu, J. E., Brown, K., Willett, R. & Mathew, J. (eds.) *Definitions, Concepts and Scope of Engineering Asset Management.* Springer London.

Wijnia, Y. C. & Huisma, J. 2007. Changing behavior for effective Asset Management. *2nd World Conference on Engineering Asset Management (WCEAM 2007).* Harrogate, United Kingdom.

Wijnia, Y. C., Korn, M. S., de Jager, S. Y. & Herder, P. M. 2006. Long Term optimization of asset replacement in energy infrastructures. *2006 IEEE Conference on Systems, Man, and Cybernetics.* Taipei, Taiwan.

Wijnia, Y. C. & Nikolic, I. 2007. Assessing Business Continuity Risk in IT. *2007 IEEE Conference on Systems, Man, and Cybernetics.* Montreal.

Wijnia, Y. C. & Peters, J. C. F. M. 2008. Integrating sustainability into risk based asset management. *International conference on infrastructure systems: Building networks for a brighter future.* Rotterdam.

Wijnia, Y. C. & Warners, J. P. 2006. Prioritizing investment. The value of portfolio decisions in electricity infrastructure management. *29th IAEE Annual International Energy Conference 2006: 'Securing Energy in Insecure Times'.* Potsdam.

Woodhouse, J. 2014. Asset Management is Growing up. Tutorial at the 9th WCEAM, Pretoria.

World Bank 2015. World development indicators. Washington, DC: World Bank.

Yellman, T. 2012. On Frequency. *Risk Analysis,* 32**,** 363-367.

18 Publications

Frey, D., Herder, P., Wijnia, Y., Subrahmanian, E., Katsikopoulos, K. & Clausing, D. 2009. The Pugh Controlled Convergence method: model-based evaluation and implications for design theory. *Research in Engineering Design,* 20**,** 41-58.

Frey, D., Herder, P., Wijnia, Y., Subrahmanian, E., Katsikopoulos, K., Neufville, R., Oye, K. & Clausing, D. 2010. Research in engineering design: the role of mathematical theory and empirical evidence. *Research in Engineering Design,* 21**,** 145-151.

Frey, D. D., Herder, P. M., Wijnia, Y. C., Subramanian, E., Katsipopoulos, K. & Clausing, D. P. 2007. An evaluation of the Pugh Controlled Convergence Method. *ASME Design Engineering Technical Conference.* Las Vegas.

Herder, P. M. & Wijnia, Y. 2011. A systems view on Infrastructure Asset Management. *In:* Van Der Lei, T., Herder, P. & Wijnia, Y. (eds.) *Asset Management: the state of Art in Europe from a Life Cycle Perspective.* Dordrecht: Springer.

Herder, P. M. & Wijnia, Y. 2012. A Systems View on Infrastructure Asset Management. *In:* Van Der Lei, T., Herder, P. & Wijnia, Y. (eds.) *Asset Management.* Springer Netherlands.

Knops, H. P. A., Ajodha, V. S. & Wijnia, Y. C. 2006. Quality Regulation of Electricity Distribution Companies: Is Everything under Control? *In:* Roggenkamp, M. & Hammer, U. (eds.) *European Energy Law Report III.* Antwerpem: Intersentia.

Masera, M., Wijnia, Y. C., de Vries, L. J., Kuenzi, C., Sajeva, M. & Weijnen, M. P. C. 2006. Governing Risks in the European Critical Electricity Infrastrcuture. *In:* Gheorghe, A. V., Masera, M., De Vries, L. J. & Weijnen, M. P. C. (eds.) *Critical Infrastructures at Risk, Securing the European Electric Power System.* Dordrecht: Springer.

Van der Lei, T., Wijnia, Y. C. & Herder, P. M. 2010. TOWARDS AN INTEGRAL ASSET MANAGEMENT FRAMEWORK USING ASSET CHARACTERISTICS, ASSET ENVIRONMENT AND LIFECYCLE PHASES AS LEADING PRINCIPLES. *the 5th World conference on engineering asset management* Brisbane, Australia: Springer.

Wijnia, Y. 2015. Dealing with Uncertainty in the Asset Replacement Decision. *In:* Tse, P. W., Mathew, J., Wong, K., Lam, R. & Ko, C. N. (eds.) *Engineering Asset Management - Systems, Professional Practices and Certification.* Springer International Publishing.

Wijnia, Y. & de Croon, J. 2015. The asset management process reference model for infrastructures *In:* Amadi-Echend, J. E. & Mathew, J. (eds.) *9th World conference on engineering asset management.* Pretoria, South africa: Springer.

Wijnia, Y. & de Croon, J. 2015. Designing an Asset Management Guideline for the Dutch Wastewater Industry. *In:* Lee, W. B., Choi, B., Ma, L. & Mathew, J. (eds.) *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012).* Springer International Publishing.

Wijnia, Y., de Croon, J. & Liyanage, J. P. 2014. Application of a Unified Reference Model Across Asset Types: Comparative Cases. *In:* Lee, J., Ni, J., Sarangapani, J. & Mathew, J. (eds.) *Engineering Asset Management 2011.* Springer London.

Wijnia, Y., de Croon, J. & Liyanage, J. P. 2014. Towards an Asset Management Reference Model: Basis for a Unified Approach. *In:* Lee, J., Ni, J., Sarangapani, J. & Mathew, J. (eds.) *Engineering Asset Management 2011.* Springer London.

Wijnia, Y. C. 2004. The Challenges of a risk based asset management organisation. *Strategic asset management***,** 832-837.

Wijnia, Y. C. 2007. The Human factor in Asset Management Process development. *2nd World Conference on Engineering Asset Management (WCEAM 2007).* Harrogate, United Kingdom.

Wijnia, Y. C. 2009. Asset Management for Infrastructures in Fast Developing Countries. *Infrastructure systems 2009: Developing 21st century Infrastructure Networks.* Chennai.

Wijnia, Y. 2012. Asset Risk Management: Issues in the Design and Use of the Risk Matrix. *In*: Mathew, J., Ma, L., Tan, A., Weijnen, M. & Lee, J. (eds.) *Engineering Asset Management and Infrastructure Sustainability*. Springer London.

Wijnia, Y. C. 2013. Dealing with uncertainty in the asset replacement decision. *In:* Mathew, J. & Tse, P. (eds.) *8th World conference on engineering asset management.* Hong Kong: Springer.

Wijnia, Y. C. 2013. Grid planning for renewable energy in urban deltas. *In:* Brown, K. (ed.) *International Conference on Strategic Infrastructure Asset Management for Deltas 2013.* Sydney, Australia.

Wijnia, Y. C. 2014. Tactical Risk Management *CESUN, 4th International Engineering Systems Symposium.* Hoboken, New Jersey.

Wijnia, Y. C. 2015 Towards quantification of optimality in asset management (to be publsihed). *WCEAM2015.*

Tampere, Finland.

Wijnia, Y. C., Croon, J. d. & Meerman, W. t. 2003. Ketenomkering bij Essent Netwerk Noord. *Energie&Techniek*.

Wijnia, Y. C., Flonk, J. & Hermkens, R. J. M. 2006. Reliability in the gas distribution infrastructure. *Seminar Reliability*. Delft.

Wijnia, Y. C., Hakvoort, R. & de Jong, H. M. 2007. Financing the replacements of energy infrastructures. *30th IAEE Annual International Energy conference 2006: From Restructuring to Sustainability: Energy Policies for the 21st Century*. Wellington, New Zealand.

Wijnia, Y. C. & Herder, P. M. 2004. Modeling Interdependencies in electricity infrastructure risk. *1st Annual CZAEE International Conference "Critical Infrastructure in the energy sector: Vulnerabilities and protection"*. Prague.

Wijnia, Y. C. & Herder, P. M. 2005. Options for real options: Dealing with uncertainty in investment decisions for electricity networks. *International conference on Systems, Man and Cybernetics*. Hawaii.

Wijnia, Y. C. & Herder, P. M. 2009. The State of Asset Management in the Netherlands. *World Conference on Engineering Asset Management*. Athens: Springer.

Wijnia, Y. C. & Herder, P. M. 2013. Flipping a coin in infrastructure decisions. *In:* Brown, K. (ed.) *International Conference on Strategic Infrastructure Asset Management for Deltas 2013*. Sydney, Australia.

Wijnia, Y. C., Herder, P. M., Korn, M. S., Poorts, M. & Veldman, E. 2008. Long term infrastructure risk management. *WCEAM IMS 2008*. Beijing.

Wijnia, Y. C. & Hermkens, R. J. M. 2006. Measuring safety in gas distribution networks. *Proceedings of the 1st World Congress on Engineering Asset Management*, WCEAM 2006.

Wijnia, Y. C., Hermkens, R. J. M. & Flonk, J. 2010. The Safety Indicator: Measuring Safety in Gas Distribution Networks. *In*: Amadi-Echendu, J. E., Brown, K., Willett, R. & Mathew, J. (eds.) *Definitions, Concepts and Scope of Engineering Asset Management*. Springer London.

Wijnia, Y. C. & Huisma, J. 2007. Changing behavior for effective Asset Management. *2nd World Conference on Engineering Asset Management (WCEAM 2007)*. Harrogate, United Kingdom.

Wijnia, Y. C., Korn, M. S. & De Jager, S. Y. 2008. Replacing infrastructure assets: Assumptions used in the long term optimization. Essent Netwerk B.V.

Wijnia, Y. C., Korn, M. S., de Jager, S. Y. & Herder, P. M. 2006. Long Term optimization of asset replacement in energy infrastructures. *2006 IEEE Conference on Systems, Man, and Cybernetics*. Taipei, Taiwan.

Wijnia, Y. C. & Nikolic, I. 2007. Assessing Business Continuity Risk in IT. *2007 IEEE Conference on Systems, Man, and Cybernetics*. Montreal.

Wijnia, Y. C. & Peters, J. C. F. M. 2008. Integrating sustainability into risk based asset management. *International conference on infrastructure systems: Building networks for a brighter future*. Rotterdam.

Wijnia, Y. C. & Warners, J. P. 2006. Asset Management is a strange Business. *Strategic asset management*, 2.

Wijnia, Y. C. & Warners, J. P. 2006. Portfolio planning part 1. *Strategic asset management*, 3-4.

Wijnia, Y. C. & Warners, J. P. 2006. Portfolio planning part 2. *Strategic asset management*, 11-12.

Wijnia, Y. C. & Warners, J. P. 2006. Portfolio planning part 3. *Strategic asset management*, 19-20.

Wijnia, Y. C. & Warners, J. P. 2006. Prioritizing investment. The value of portfolio decisions in electricity infrastructure management. *29th IAEE Annual International Energy Conference 2006: 'Securing Energy in Insecure Times'*. Potsdam.

CURRICULUM VITAE

Ype Cornelis Wijnia was born on March 8th 1972 in Alphen aan den Rijn, the Netherlands. He finished his pre-university education at the Sint Maartens College in Groningen in 1990. In the same year Ype started with applied physics at Delft University of Technology. In 1994, Ype switched to the new faculty of System Engineering and Policy Analysis, from which he received his MSc degree in 1998.

After his graduation, Ype started working for EDON Netwerk (which in a series of mergers and name changes developed via Essent Netwerk (Noord) into Enexis) as investment appraisal engineer. In 1999, he became project leader for reorganizing decision making on investments, which turned out to be the start of asset management. For his work in this reorganization Ype was finalist in the Utilities Young management award of Accenture in 2001. Within the asset management department, Ype was appointed risk manager in 2001. In 2004, Ype introduced PAS 55 to Essent Netwerk. This resulted in the first certification against PAS 55 in the Netherlands and the second certification worldwide. In 2005 Ype was appointed risk manager for enterprise risk management as well. In 2006, Essent Netwerk and Ype were finalist in the Asset Management Innovation Award of the UK based Institute of Asset Management on incorporating external stakeholders into asset management. Ype continued his position as risk manager until 2008, when he left Essent to start his own business on asset management consultancy. This is still his current position. Ype is currently mostly working on the implementation of asset management for local governments, concerning assets for wastewater, roads, engineering structures and the green space.

Ype returned to Delft University of Technology in 2004, when Essent Netwerk and Next Generation Infrastructures agreed on a part time PhD project for risk management in asset management. During his research Ype published more than 50 contributions to the field of asset management. Since the first conference in 2006 Ype is part of the scientific asset management community which annually convenes in the World Conference on Engineering Asset Management. Since 2014 Ype is a fellow of the International Society for Engineering Asset Management. In 2015 Ype was asked to join the board of this society.

Processing Risk in Asset Management

In the liberalized energy market Distribution Network Operators (DNOs) are confronted with income reductions by the regulator. The common response to this challenge is the implementation of asset management, which can be regarded as systematically applying Cost Benefit Analysis (CBA) to the risks in the networks. In short, this is Risk Based Optimization (RBO). However, application of RBO is mostly limited to interventions on individual assets like upgrades, replacements and maintenance. Whether RBO is feasible for higher levels of aggregation like the portfolio of interventions or even the whole system was not clear. The unavoidable subjectivity and uncertainty associated with risk decision making could threaten the acceptance of decision outcomes.

The experiments conducted in this research reveal that there are no fundamental barriers to risk based optimization of the whole system. Embracing uncertainty and subjectivity allows for relatively simple tools, as the tools do not need to be more accurate than our knowledge of the future. The condition for this to work is that the rational RBO decisions are embedded in a well-designed sociotechnical process. A systematic implementation of RBO on all levels (individual assets, portfolios of interventions and the whole system) results in a reduction of the total costs of the system (expenditure plus residual risk) of about 20%.