Integrated Risk-Capability Analysis under Deep Uncertainty: an E***SD***MA Approach

Erik Pruyt[1] and Jan H. Kwakkel

Delft University of Technology, Faculty of Technology, Policy and Management

Abstract

Integrated risk-capability analysis methodologies for dealing with increasing degrees of complexity and deep uncertainty are urgently needed in an ever more complex and uncertain world. Although scenario approaches, risk assessment methods, and capability analysis methods are used, few organizations and nations use truly integrated risk-capability approaches, and almost none use integrated risk-capability approaches that take dynamic complexity and deep uncertainty seriously into account. This paper presents and illustrates a novel integrated risk-capability analysis approach for dealing with deeply uncertain dynamically complex risks, and discusses near future developments related to integrated risk-capability analysis for such issues. This approach combines System Dynamics Modeling for dealing with dynamic complexity and Exploratory Modeling and Analysis for dealing with deep uncertainty. This multi-method approach is illustrated here using an acute and a chronic public health risk: a new flu and Lyme disease.

Keywords: Integrated Risk-Capability Analysis, National Risk Assessment, Capability Analysis, Exploratory Modeling and Analysis, System Dynamics

1. Introduction

1.1. Integrated Risk-Capability Analysis

Many governments use national risk assessment (NRA) methods as well as capability-based planning or capability analysis (CA) approaches to be prepared for major threats. Many recently developed NRAs are all-hazard, scenario-based, and multi-dimensional: scenarios of accidental, natural, and man-made risks are developed by experts, and made comparable in terms of their multi-dimensional impact and likelihood by means of multi-criteria (MCDA) approaches and a common set of criteria. All-hazard CA methods subsequently help –in some cases– to identify the most important capabilities, both generic (applicable to all or most risks) and specific (only for a particular risk or very few

[1]Corresponding Author – E-mail: E.Pruyt@tudelft.nl – Telephone: + 31 15 2787468

specific risks), that need to be reinforced to be able to prevent, protect against, respond to, and/or recover from major events (see `http://cdn-cbp.org` or `https://www.rkb.us/hspd8.cfm` for lists and detailed examples of such capabilities). A multi-method consistently combining risk assessment (RA) and CA in an integrated way is called here an Integrated Risk Capability Analysis (IRCA).

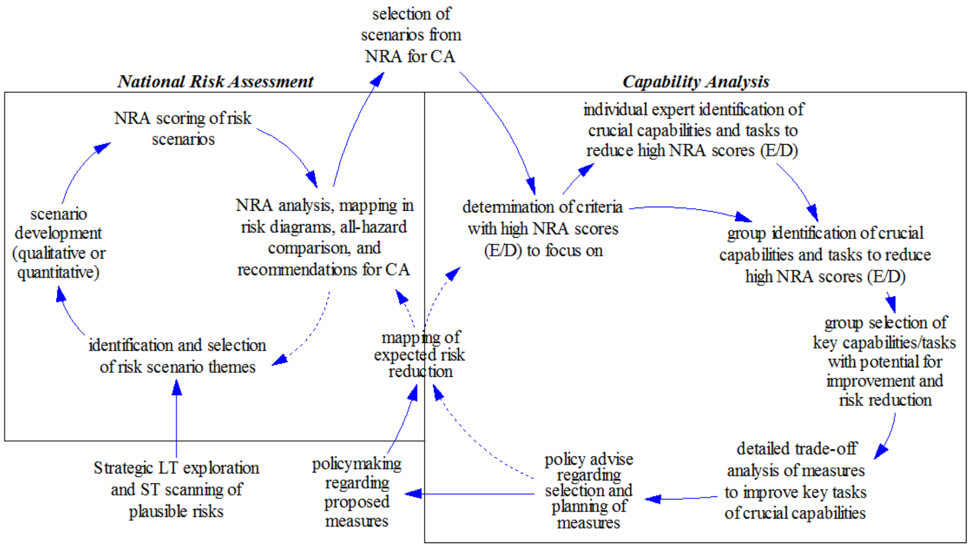


Figure 1: Iterative process of the state of the art Integrated Risk Assessment and Capability Analysis ***not*** under deep uncertainty

The Netherlands, for example, developed an IRCA consisting of an all-hazard NRA MCDA approach embedded in a recurring process to develop, assess, plot and compare all-hazard risk scenarios, and an expert-based CA approach to identify and select capabilities requiring further improvement or investments (see Figure 1 and (Pruyt and Wijnmalen, 2010; Bergmans et al., 2009a; Pruyt and Kwakkel, accepted)). All kinds of risks are scored and assessed in the NRA on the following 10 criteria:

C1.1 Infringement of the Dutch territorial integrity

C1.2 Infringement of the integrity of the Dutch international position

C2.1 Number of fatalities

C2.2 Number of seriously injured and chronically ill

C2.3 Physical suffering (lack of fulfilment of basic needs)

C3.1 Economic costs (repair costs regarding sustained damage, loss of income)

C4.1 Long-term damage to the environment

C5.1 Disruption of everyday life (schools, work public transport, etc.)

C5.2 Violation of the democratic system (political, financial, etc.)

C5.3 Psychological impact (public outrage and anxiety/fear)

The same qualitative labels are used to score all criteria according to criterion-specific scales: E stands for 'catastrophic', D for 'very serious', C for 'serious', B for 'substantial', A for 'limited', and '–' for 'not applicable'. These qualitative labels are the same for all criteria, but have criterion-specific underlying scales (e.g. $C2.2()$ = number of injured and chronically ill: $C2.2(risk) < 10 \Rightarrow A$; $10 \leq C2.2(risk) < 100 \Rightarrow B$; $100 \leq C2.2(risk) < 1000 \Rightarrow C$; $1000 \leq C2.2(risk) < 10000 \Rightarrow D$; $C2.2(risk) \geq 10000 \Rightarrow E$). The likelihood is expressed with A-E labels too: E stands for 'very likely', D for 'likely', C for 'possible', B for 'unlikely', and A for 'very unlikely'. Table 1 displays the scores of some of the scenarios published in the 2010 NRA Report (MinVenJ, 2010).

Scenario	LH	1.1	1.2	2.1	2.2	2.3	3.1	4.1	5.1	5.2	5.3
Flu pandemic mild	**D**	-	-	**D**	C	A	**D**	-	B	-	**E**
Flu pandemic serious	**D**	-	-	**E**	**D**	**E**	**D**	-	**E**	C	**E**
Right-wing extremism	**D**	-	A	A	A	-	A	-	B	B	C
Left-wing extremism	C	-	A	A	A	-	B	-	A	A	A
Animal rights activism	C	-	A	A	A	-	A	-	A	-	**D**
Animal rights extremism	C	-	A	A	A	-	A	-	**D**	B	**E**
...	...	...	...	...	...	...	...	...	...	...	...

Table 1: A-E scores of some 2010 NRA risks – Source: MinVenJ (2010)

These criteria and qualitative labels are assumed to be commensurable and comparable. A 3^{rd} order exponential scale is used –for the basic calculations and standard risk diagram– to transform the qualitative label scores into numbers which are multiplied by the corresponding weights of the criteria. The sums of these products are the total multi-criteria scores –more specifically the MAVT scores (Belton and Stewart, 2002)– which are plotted on the logarithmic Y-axis of the risk diagram (see Figure 2). All scenarios with an aggregated impact score between 0.33 and 1 are –given the 3^{rd} order exponential scale– *catastrophic*; all scenarios with an aggregated impact score between 0.11 and 0.33 are *very serious*; et cetera. Risks are then selected, based among else on the mapping of risks in this risk diagram, as inputs of a capability-based planning process.

It should be clear from the above that the Dutch IRCA –like most IRCAs– was first and foremost developed for relatively well-known and relatively simple incident-type risks – not for deeply uncertain and/or dynamically complex risks. The Dutch approach was complemented several years after its initial development with two approaches for developing and assessing slumbering/latent process-type risks since these risks may be at least as important as incident risks. Risks of the latter type are characterized particularly by dynamic complexity and/or deep uncertainty.

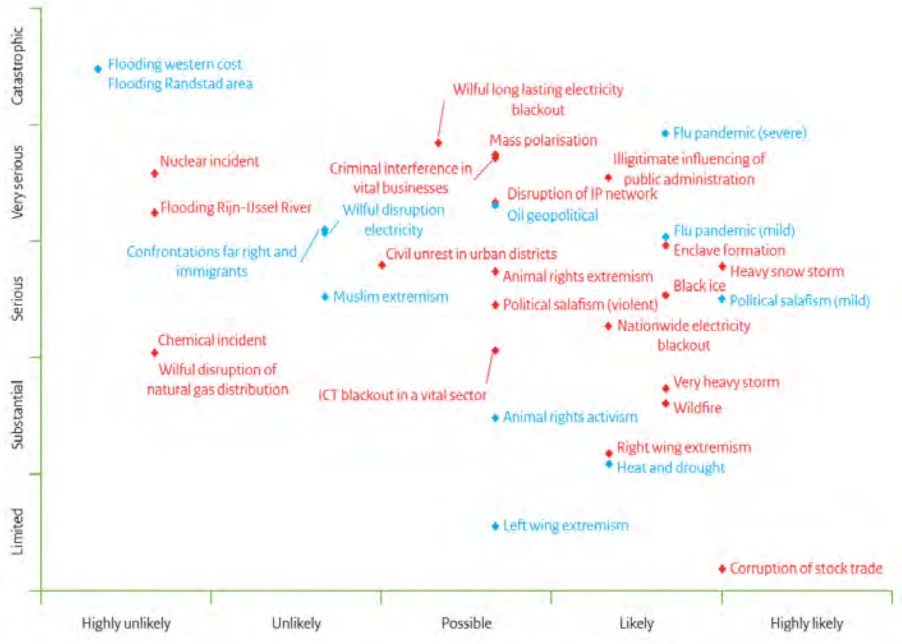


Figure 2: Expected values of risks mapped in the logarithmic risk diagram (source: (Bergmans et al., 2009b, p63))

1.2. IRCA for Deeply Uncertain Dynamically Complex Risks

A risk is dynamically complex if it is characterized by subtle cause and effect relations, if its time evolutionary behavior matters, and/or if the mediating effects of interventions on its dynamics are subtle too. System Dynamics[2] (SD) is a modeling and simulation method for dealing with dynamically complex issues. One of the approaches added to the Dutch IRCA for developing and assessing slumbering/latent process-type risks is based on System Dynamics modeling and simulation. The other approach developed for the Dutch IRCA is somewhat similar to a morphological analysis[3] and allows dealing with risks of –at most– medium uncertainty and some dynamic complexity.

Risks are deeply uncertain if *(i)* it is uncertain which of many plausible underlying mechanisms will generate the real-world dynamics, *(ii)* it is uncertain which probabilities may be attached to plausible outcomes, and *(iii)* different

[2]See for a start (Forrester, 1961; Richardson and Pugh III, 1981; Ford, 1999; Sterman, 2000).

[3]Morphological analysis could be used as a systemic scenario development method for dealing with a limited degree of combinatorial complexity. It allows to combine more driving forces and uncertainties than in traditional scenario development methods. These driving forces and uncertainties do not need to be quantifiable. The driving forces and uncertainties are then set out as axes of a box/grid. The cells of the grid contain all combinations of driving forces and uncertainties.

experts may disagree about the acceptability of the outcomes (Lempert et al., 2003). Deeply uncertain risks are thus risks that could be modeled, although the models, structures, and parameter values used are at most plausible, and so are their outcomes. But the occurrence of these plausible scenarios cannot be captured by means of just one or very few scenarios with ordinal or cardinal likelihoods, nor by means of probability distributions.

Exploratory Modeling and Analysis (EMA) is a model-based methodology for dealing with deep uncertainty. EMA is in fact a computational methodology for systematically exploring issues under deep uncertainty, and for testing and comparing the robustness of policies under deep uncertainty (Agusdinata, 2008; Bankes, 1993; Lempert et al., 2003; Kwakkel and Pruyt, 2012a). It consists of using exploratory models to generate tens of thousands of scenarios –called an 'ensemble of scenarios'– in view of exploring and analyzing this ensemble of plausible futures, and testing the robustness of policy options across the ensemble of scenarios, in other words, testing whether the outcomes are acceptable for all scenarios generated by sweeping the entire multi-dimensional uncertainty space. However, in order to perform the EMA methodology, computational models are required. If risks are dynamically complex and deeply uncertain then models for dealing with dynamic complexity are required.

The combination of SD and EMA –which we call Exploratory System Dynamics Modeling and Analysis (ESDMA)– could indeed be used for dealing with deep uncertainty and dynamic complexity. Since SD models are very useful for simulating dynamically complex issues, they are used in ESDMA as EMA scenario generators. Or stated differently, since EMA allows to extend the use of SD models to deep uncertainty, EMA is used to make and use SD modeling for deeply uncertain issues. Risks (e.g. societal aging, acute and chronic pandemics, radicalization, et cetera) could then be captured by means of different exploratory SD models which are subsequently used to generate tens of thousands of plausible scenarios for each risk. Scenario discovery techniques could then be applied to each of the risks to select a small representative set of scenarios for each of these risks. These representative scenarios could subsequently be used in a CA under deep uncertainty, based on a separate generic CA model in order to assess the overall all-hazard effectiveness and robustness of different capability policies. Finally, different capability investment and activation policies could be assessed and compared, and one may try to improve the most promising policies even further using robust optimization techniques.

1.3. Goal and Organization of the Paper

The main goal of this paper is to explain and illustrate a new IRCA approach based on EMA and SD modeling and simulation for dealing with all sorts of risks, including deeply uncertain dynamically complex risks. The methodology is introduced in section 2: the phases of the approach are briefly explained in subsection 2.1, followed by a general discussion of SD-based NRA in subsection 2.2, and a more in-depth discussion of the SD-based CA in subsection 2.3. The approach is illustrated in section 3. Concluding remarks are formulated in section 4.

This paper overlaps to a large extent with (Pruyt, 2012). The reason for this overlap is due to a lack of progress between the state of this line of work at the time of presenting (Pruyt, 2012) and the state of our work at the time of presenting the current paper. We planned on extending the former paper with robust optimization over multiple risks in the current paper, but were unable to do so due to (i) several technical problems related to robust optimization in our EMA workbench, (ii) methodological and computational problems related to multi-hazard / multi-model robust optimization, and (iii) lack of time. Hence, the main differences between the current paper and (Pruyt, 2012) are the emphasis, less versus more emphasis on System Dynamics modeling, and the size of the paper. This paper also overlaps to a lesser extent with two other papers:

- One of the two illustrations, Lyme Disease, is elaborated in more detail in (Pruyt and Coumou, 2012), hence the overlap in diagrams and graphs.
- The other illustration is used in (Pruyt and Kwakkel, accepted). Some figures used in this paper therefore also overlap with figures used there.

2. Methodology

2.1. Overview: Phases of the IRCA

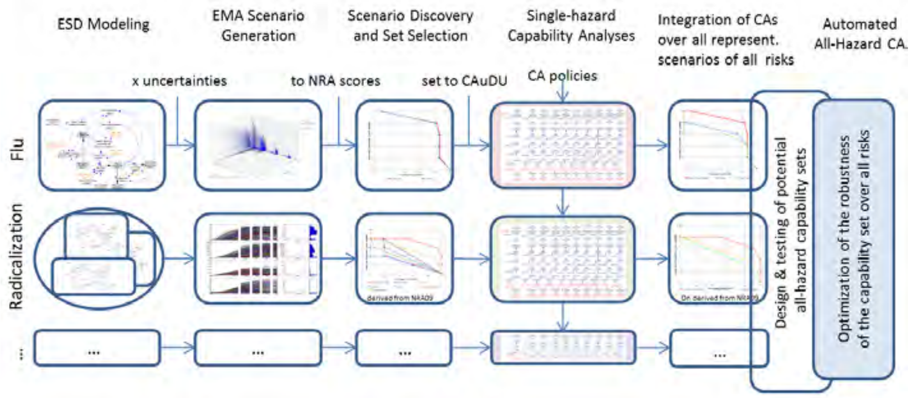


Figure 3: Integrated All-hazard Risk-Capability under deep uncertainty

The computational IRCA approach presented here is an integrated quantitative model-based approach for dealing with all sorts of risks, from well-known to uncertain, and from simple to dynamically complex. It consists of the six phases displayed in Figure 3:

1. SD Modeling: The first step in multidimensional risk assessment under deep uncertainty is the development and/or use of one or more simulation models to generate an ensemble of scenarios for each risk, e.g. an ensemble of plausible flu scenarios. Certainties and uncertainties about the risk need

to be identified in order to build plausible simulation models for generating the widest variety of plausible risk scenarios.

2. EMA Scenario Generation phase: These models and the widest possible uncertainty ranges are combined using EMA in order to generate tens of thousands of plausible time-evolutionary scenarios for each of the risks.
3. Scenario Discovery and Set Selection phase: Then, a much smaller set of scenarios that are representative (in terms of impacts, time-evolutionary behavior, and origin in the multi-dimensional uncertainty space) for the full ensemble of scenarios of a particular risk needs to be identified, for example using scenario discovery methods. In this paper, scenarios are clustered and selected based on their time-evolutionary behavior, overall impact, and –to some extent– the origin in the multi-dimensional uncertainty space; the exemplars of these clusters are then used in the ensuing CA. Envelope diagrams are useful at this stage to verify the representativeness of the selected ensemble with the large ensemble in terms of overall impact.
4. Single-hazard Capability Analysis phase: The deeply uncertain mediating effects of different capability policies on scenarios are then calculated for all representative scenarios of all risks using risk-specific versions of a CA model under deep uncertainty (CAuDU), i.e. a rather generic and open SD model, resulting in thousands of simulations per representative risk scenario per risk, or tens of thousands of runs per risk. Each risk requires different settings of the CAuDU model, which is why risks need to be treated separately until after the CAuDU.
5. Integration of CAs over all representative scenarios of all risks: The deeply uncertain mediating effects of different capability policies are then calculated for, and compared over, all risks. A countable set of potential all-hazard capability policy mixes could then be designed, tested, and compared, e.g. with MCDA methods, in order to provide insight into the appropriateness of different sets of capabilities over all risks. Robust optimization for particular risks may also be useful at this stage to search for policy mixes increase the robustness of these.
6. Automated All-Hazard CA: Robust optimization could also be used to obtain the most robust capability set given a particular investment level, starting from promising sets, for all representative scenarios of all risks simultaneously. Ideally, an all-hazard capability analysis consists of loading all representative sets of scenarios for all risks as cases for an all-hazard CA, followed by automated robustness testing (varying all combinations of capabilities, maximizing the robustness at particular costs or minimizing costs for a particular acceptability level). The most robust capabilities policy mix over all deeply uncertain risks could then be identified and chosen.

2.2. In Depth: SD-based Risk Assessment and Scenario Discovery and Selection

Dealing with risks is dealing with uncertainties. In the model-based Risk Assessment discussed and illustrated in this paper, uncertainties could be parametric (continuous or integer) but also categorical (e.g. for switches), and could relate to uncertain parameters, functions, exogenous time series, scenarios, model structures, models, and even methods. Some of these uncertainties may require additional model structures. Additional structures are also needed to translate values on continuous key performance indicators in these SD models to the discrete NRA classes underlying the NRA criteria (see the C11…C53 structures in Figures 5 and 14 (in orange)).

A sampling plan generated in our Python workbench combining uncertainties and model(s) and executed by Vensim DSS then generates the ensemble of scenarios. Key performance indicators for each run are saved to a data file, which could be used and handled at a later stage, e.g. to visualize and analyze the artificial data, or to discover and select exemplars representative for many scenarios.

Applying scenario discovery and selection algorithms is particularly useful at this point *(i)* if a very small number of interesting scenarios is needed as is the case in the traditional Dutch IRCA, or *(ii)* if a subset of scenarios representative for the larger set is desirable as is the case here for computational reasons (hundred uncertain CA runs are generated for each of the 100 NRA scenario entering the CAuDU). Depending on the goal and the issue, different exemplars are desirable: ideally, exemplars are representative for many other scenarios in terms of multi-dimensional impacts, time-evolutionary behavior (i.e. the dynamics), and the origins in the multi-dimensional uncertainty space. Although selecting exemplars based on all three of these aspects is rather complicated, selecting exemplars based on one of these aspects is straightforward and could be supported by:

- data set splitting to select exemplars that are representative in terms of multi-dimensional impacts.
- using a time-series clusterer algorithm –in our case a more advanced version of the one proposed by Yucel and Barlas (2011) with a metric proposed by Yucel (2012)– with dendrogram and cluster plots to cluster time series data based on attributes and select the similarity level at which to classify/plot clusters. Using this clusterer and visualizing exemplars for selected clusters is a powerful scenario discovery and selection approach if time-evolutionary behavior (i.e. the dynamics) is important.
- using a new version of PRIM or Patient Rule Induction Method (Friedman and Fisher, 1999) –one that can deal with categorical and continuous uncertainties– with a binary classification function and PRIM box plotting, to identify uncertainty space boxes with high concentrations of runs that perform below/above a particular threshold, and hence, to identify exemplars that represent regions in the multi-dimensional uncertainty space with high concentrations of highly undesirable or desirable outcomes (e.g. catastrophic subspaces) or of outcomes with particular characteristics.

2.3. In Depth: SD-based Capability Analysis

A CA on the outcomes of the model-based NRA phase could be performed in at least two ways: either CA structures are built into each NRA model, or all risks are pushed through the same CA model. The major disadvantages of the former approach are the additional modeling required in simulation models related to each risk, and the inability to perform multi-risk CA on these rather different models with very different sets of uncertainties. Hence, we chose the latter approach.

The current state of the art of CAuDU corresponds to manual, repetitive, single-hazard CA under deep uncertainty. For doing so, one needs to identify first of all the certainties and uncertainties related to plausible mediating effects of capabilities on risks. In other words, one or more plausible capability model(s) are required, as well as an assessment of plausible occurrence and strength of mediating effects of capabilities on particular risks.

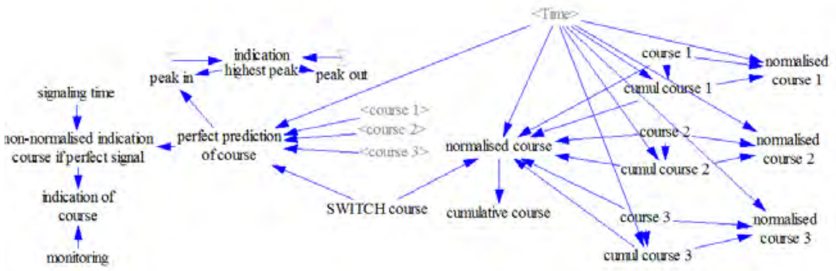
Figure 4 shows structures of the SD-based CA model used here: it consists of *(a)* structures to simulate multiple evolutions over time, *(b)* structures related to different types of capabilities (prevention, protection, sensing, response, recovering), (dis)investment in capabilities, and the (maximum) plausible mediating effects of capabilities on impact criteria (e.g. lkp1 C4, lkp2 C4, and lkp3 C4 are different lookup functions of plausible mediating effects of the recovering capability C4), and *(c)* structures to calculate the successive mediating effects under deep uncertainty of those capabilities on the risks at hand and their possible interaction effects. In this CA model, many choices could –but do not have to– be made: uncertainty rules unless it is replaced by less uncertain or more certain information.

The input of the model-based CA consist of *(i)* a set of terminal scenario values ranges, i.e. ranges for the possible impact of a risk scenario on each of the 10 NRA criteria, *(ii)* a set of maximum reducible NRA impact scores, *(iii)* a selection of possible evolutions over time[4], and *(iv)* information regarding the mediating effects of capabilities on the different impact dimensions for a particular risk. A sampling plan is then generated and executed. Robust optimization could be used, starting from a promising capability policy to further improve the capability policy over a large ensemble of plausible scenarios.

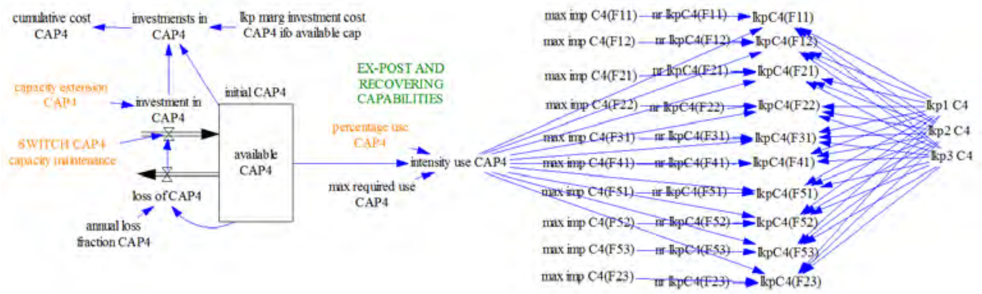
Ongoing work will soon result in automatic robust optimization over all risks using robust optimization of of capability mixes. The most appropriate capability strategy under uncertainty could then be selected. Currently, following lighter form of full-hazard capability analysis under deep uncertainty could already be performed:

1. Identification of promising sets of measures to improve capabilities for all scenarios;

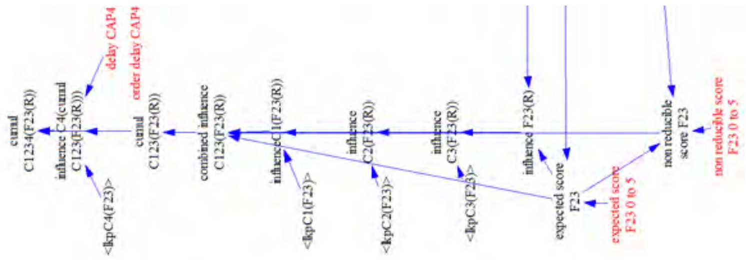
[4]The time evolutionary scenarios generated with the SD-based NRA models could be used as well. The reason for combining impacts and a selection of evolutions over time instead of directly using the time evolutionary scenarios, is that this approach also allows using impact scores not generated with dynamic models.



(a) structures to simulate multiple evolutions over time



(b) structures related to capabilities, (dis)investment in capabilities, and plausible mediating effects on impact criteria



(c) structures for each criterion to calculate mediating effects under deep uncertainty of capabilities on risks

Figure 4: Structures of the Risk-Capabilities model used in this paper

2. Application of these promising sets of measures to improve capabilities for all scenarios to each of the risks;
3. Comparison of the outcomes of promising sets of measures to improve capabilities for each of the risks;
4. Use of a (multi-criteria) method to select the most appropriate promising set of measures to improve capabilities for all scenarios of all risks.

2.4. In Practice – The Process, People, and Programs

Risk-specific experts are required to provide insights in plausible underlying mechanisms and uncertainties, for example in a risk-related GMB workshop. Based on these insights, modelers build for each risk a simulation model or a set of simulation models.

We currently build our exploratory SD models in Vensim DSS, then use a shell written in Python to generate experimental designs and force Vensim DSS to execute tens of thousands of experiments (i.e. combinations of uncertainties and models) to generate tens of thousands of transient scenarios (simulation runs). Python stores the data when generated. We then use a library of algorithms coded in Python, C, and C++ to analyze the ensemble of scenarios, and (interactively) visualize the most interesting findings, select specific scenarios, etc.

Risk-capability experts are required to provide insights in plausible mediating effects, and policy makers or their representatives are required to provide insights in politically acceptable sets of capabilities in risk-capability workshops in which the consequences of hypotheses and capability sets can be simulated on the spot.

Again, using Python and Vensim DSS, an experimental design combining all remaining uncertainties is forced upon the input (a particular risk scenario or a set of risk scenarios), the CAuDU model, and a capabilities strategy contained in or enforced upon the model, in order to generate the widest variety of plausible risk-capability scenarios.

The condition *sine qua non* for real-world implementation of this integrated risk capability analysis under deep uncertainty is the existence of software to generate large ensembles of scenarios, of software to select representative ensembles of scenarios, risk-specific models, generic risk-capabilities models that can be adapted to different risks, and software to perform robust optimization robustness of capability sets simultaneously over many risks – all of which are indeed available or under development. Hence, all ingredients will soon be there to turn the current state of science into a new state of art.

3. Illustration: IRCA under Deep Uncertainty for The Netherlands

3.1. RA and Partial CA for a Plethora of Risks

3.1.1. Flu Pandemic

The SD model displayed in Figure 5 was used to generate plausible flu scenarios. The model was developed and used in 2009 by Pruyt and Hamarat (2010)

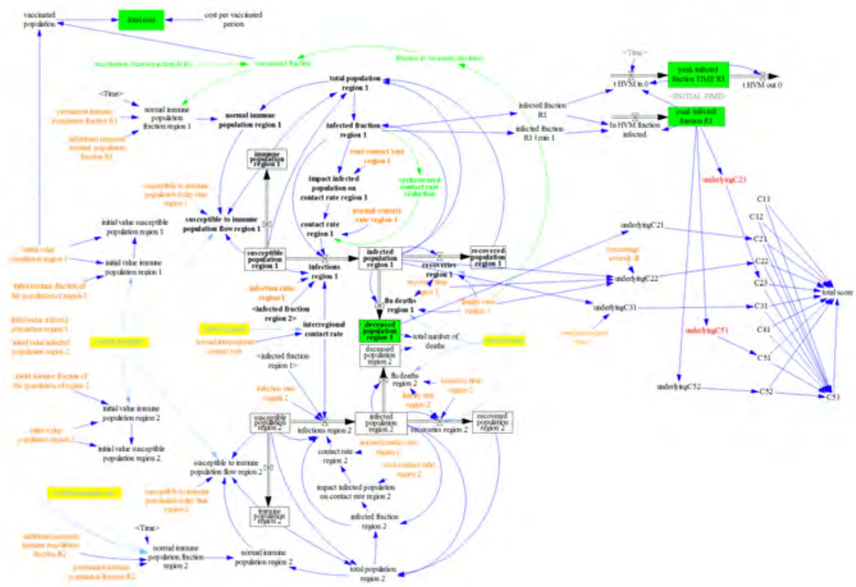


Figure 5: Exploratory System Dynamics model for simulating flu scenarios

during the H1N1 pandemic. For this paper, the model was slightly adapted (populations and initial numbers of infected individuals) in order to make it useful for generating plausible scenarios of a new type of flu in the Netherlands. Figure 6 shows a lines plot of the infected fraction for 1000 plausible outbreaks in the Netherlands of a new flu virus over a period of 48 months. Note in the bottommost graph of Figure 6 that catastrophic –in NRA terms, i.e. above 0.33– are almost all extremely infectious and happen fast, and that most new flu outbreaks are 'very serious' but not catastrophic in NRA terms and happen slightly slower.

Figure 7 shows a selection of 100 representative scenarios selected from the larger ensemble of 1000 runs. This set of representative scenarios is composed of 2 exemplars from each of 16 different time-series clusters displayed in Figure 8(a), supplemented with 68 hand-picked exemplars from the largest clusters (proportional to the size of the clusters). Figure 8(b) displays a 'risk envelope diagram', which could be used to plot deeply uncertain risks as well as other risks without any knowledge about real probabilities. A risk envelope diagram is in fact a risk diagram in which the cumulative relative number of runs in each of the total impact classes starting with the highest impact class are plotted. In other words, 20% of the 1000 risk analysis runs have a catastrophic NRA impact, about 83% of these 1000 runs have at least a very serious impact, and 98% have at least a limited impact. This risk envelope diagram also shows that the small 'representative ensemble' (blue line) is largely –but not entirely– representative of the larger ensemble (red line) – the total impact of the smaller

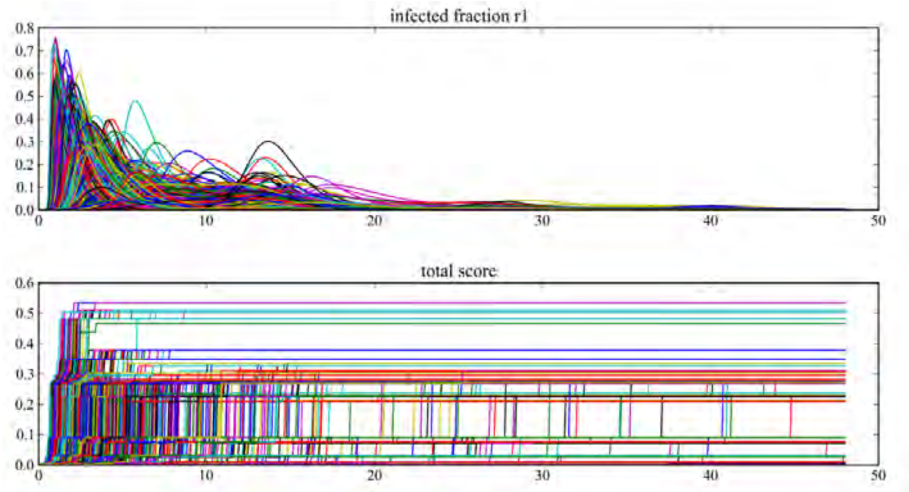


Figure 6: Top: Infected fraction of the Dutch population for 1000 plausible flu scenarios; Bottom: Total NRA scores of the 1000 plausible flu scenarios

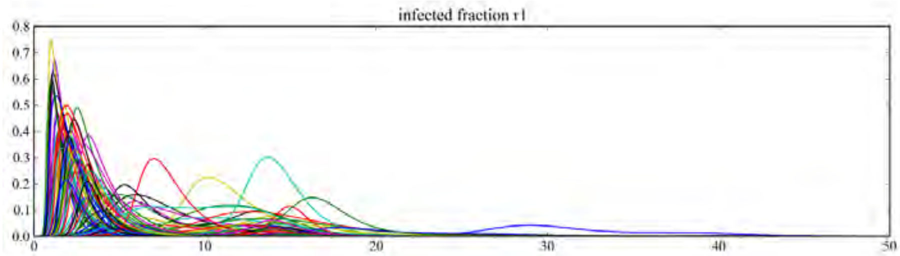
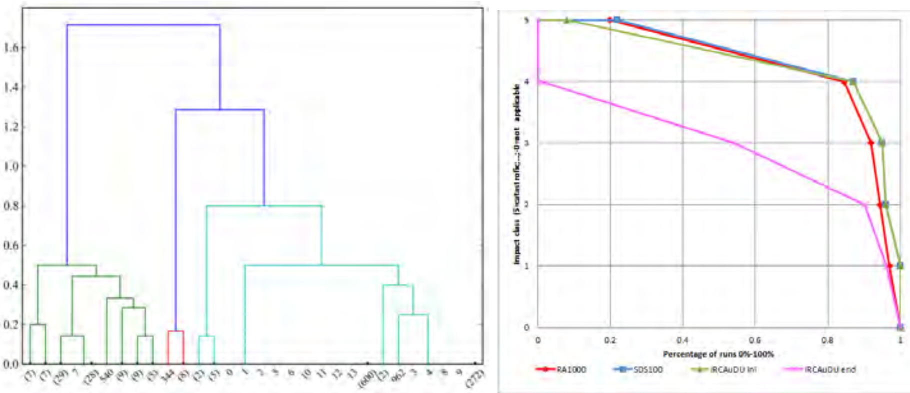


Figure 7: Small representative ensemble of 100 flu scenarios

ensemble is slightly worse in NRA terms than the larger ensemble.

Figure 8(b) also shows the overall impact of an arbitrary capabilities policy with uncertain mediating effects on the small ensemble: first, all NRA criterion scores of the small ensemble are combined under uncertainty with selected evolutions resulting in the 'IRCAuDU ini' ensemble of 10000 runs, i.e. without any capabilities policy, which are also subjected to the capabilities policy, resulting in the 'IRCAuDU end' ensemble of 10000 runs. These two envelopes summarize the information displayed in Figure 9 which shows the simulated evolutions in the CA model in terms of NRA scores and NRA classification with and without capabilities policy.

Figure 9 displays in other words the reduction of the impact range of the two national risk analysis flu scenarios by a particular capabilities strategy with uncertain mediating effects. The cumulative evolutions over time of these flu pandemic scenarios are different from the ones simulated with the flu model: a selected set of plausible evolutions and ranges of total impact scores on the



(a) Time series clusterer dendrogram (flu)

(b) Risk envelopes diagram (flu)

Figure 8: Dendrogram (left) and Risk Envelopes Diagram (right) for the flu case

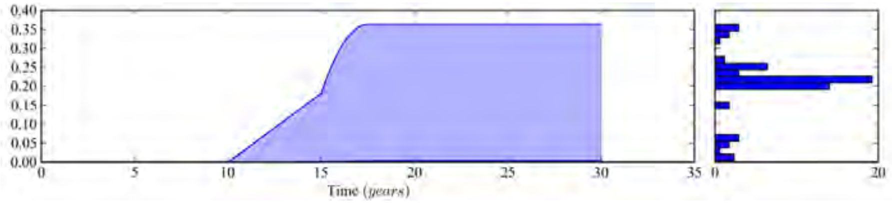
criteria are combined in the CA model instead of directly importing the evolutions of the flu scenarios from the flu model. This choice was made in order to treat all scenarios –also those NRA scenarios which were not generated with simulation models– entering the CA model the same.

3.1.2. Lyme disease

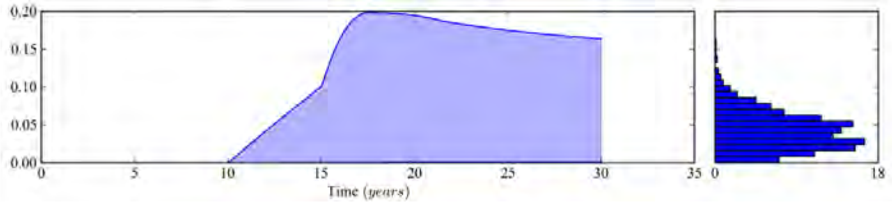
Lyme disease poses an uncertain dynamic threat to the Dutch population and the Dutch health care system. On the total Dutch population of over 16 million, 800000 are chronically ill. The extent of Lyme disease, chronic Lyme disease, and post lyme syndrome, their contribution to the large number of chronically ill, and their future development are unknown. The SD simulation models displayed in Figure 14 on page 19 was therefore developed and used by Pruyt and Coumou (2012) to generate thousands of plausible evolutions of lyme disease in order to assess the future risk posed by it.

The topmost graph in Figure 10 displays 1000 strongly smoothed (3rd order smoothing with a 24 month smoothing time) evolutions of the number of known and unknown sick caused by lyme disease in the Netherlands, i.e. all those that were diagnosed and are treated, and all those that were not diagnosed and remain untreated. The unit of time is months. The bottommost graph in Figure 10 shows the risk posed by lyme disease using the Dutch NRA framework adapted to deeply uncertain dynamically complex risks. Note that none of these scenarios is catastrophic (>0.33).

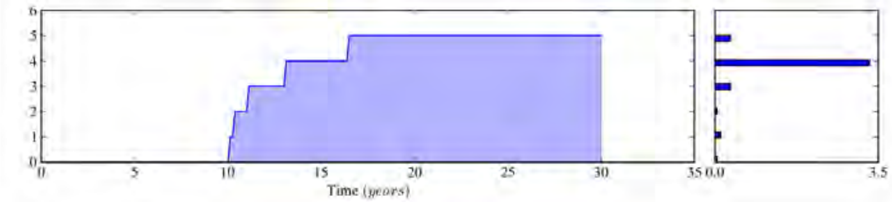
Figure 11 displays an ensemble of 100 scenarios discovered in, and selected from, the larger ensemble using time series clustering (see dendrogram in Figure 12(a)). Figure 12(b) shows that this small ensemble represents the larger ensemble well in terms of the total NRA score (compare the red and blue envelopes). The green and pink envelopes in Figure 12(b) summarize the information contained in Figure 13 regarding the 10000 scenarios without capabilities policy



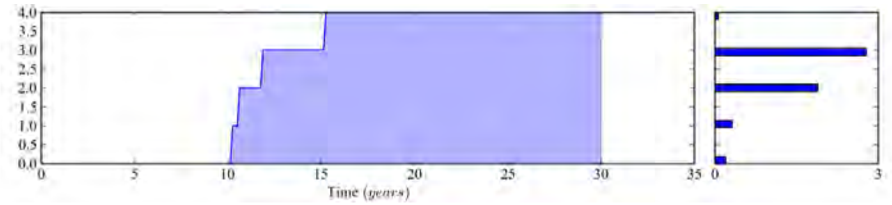
(a) NRA scores without capability policy



(b) NRA scores with capability policy



(c) with capability policy in NRA classes (5 = catastrophic; 4 = very serious; 3 = serious; 2 = substantial; 1 = limited; 0 = not)



(d) with capability policy in NRA classes (5 = catastrophic; 4 = very serious; 3 = serious; 2 = substantial; 1 = limited; 0 = not)

Figure 9: Flu: CA without a capabilities policy and with a capabilities policy, expressed in terms of NRA scores ($1 \leq$ catastrophic $\leq 0.33 > \ldots$) and NRA classes (5 = catastrophic; 4 = very serious;...)

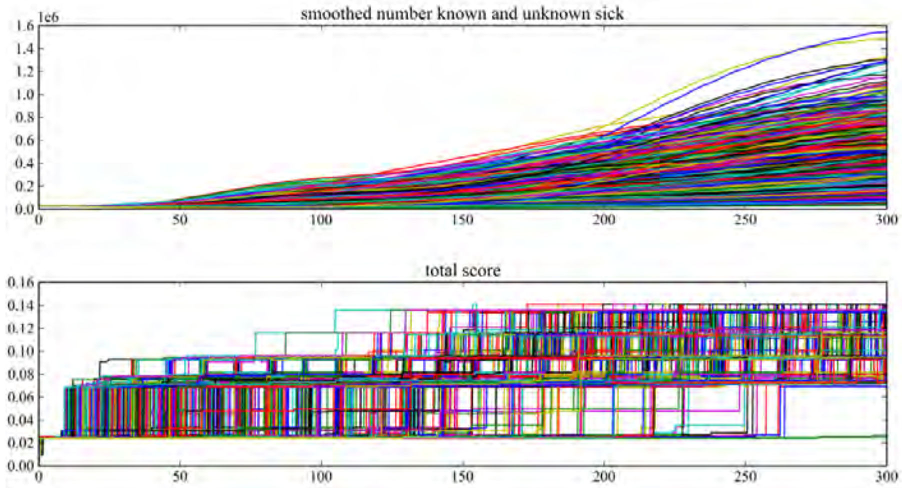


Figure 10: Top: smoothed number of known and unknown sick due to lyme disease; Bottom: Total NRA scores of the 1000 plausible lyme scenarios

and the 10000 scenarios with the same (arbitrary) capabilities policy with the same settings as in the flu case, assuming for the sake of illustration that the same capabilities policy impacts flu and lyme disease similarly.

3.2. Comparison of CA Policies over Multiple Risks

Risk envelopes diagrams could also be used to visualize the overall mediating effect of a policy on multiple risks or to visualize the overall mediating effect of different policies on one or more risks as in Figure 15. Figure 15 shows the mediating effects of two policies on the flu ensemble (in red) and the lyme disease ensemble (in blue). The 'no policy' cases are displayed with full lines, the policy 1 envelopes with dashed lines, and the policy 2 envelopes with dotted lines. Policy 1, a hypothetical capabilities mix with the same mediating effects on flu and lyme disease, results in a significant improvement of both envelopes. Policy 2, another hypothetical capabilities mix with the same mediating effects on flu and lyme disease, results in a less significant improvement of both envelopes.

Risk envelopes diagrams with multiple policies for multiple risks, quickly become hard to read, unless very few risks and very few policies with clear and unambiguous effects are plotted. Else, multi-criteria analysis or robust optimization may be required.

4. Conclusions

In spite of the fact that most IRCAs attempt to be fully and all-hazard integrated, few really are in practice, because they were not designed as such, and because they are appropriate only for a subset of risks, i.e. not for dynamically

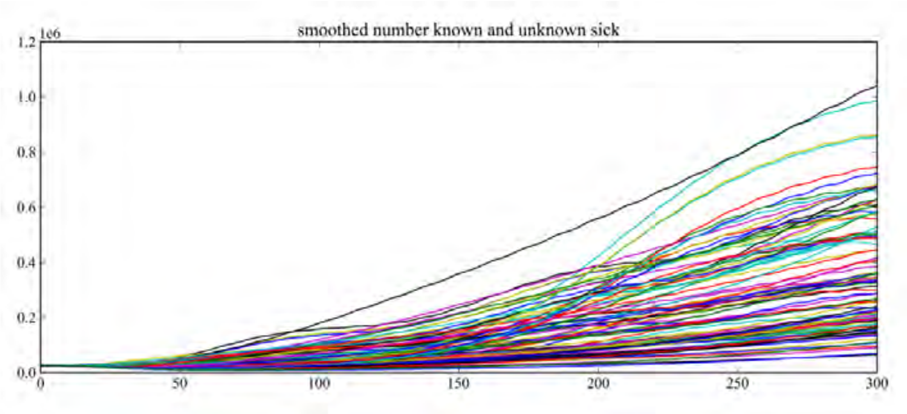
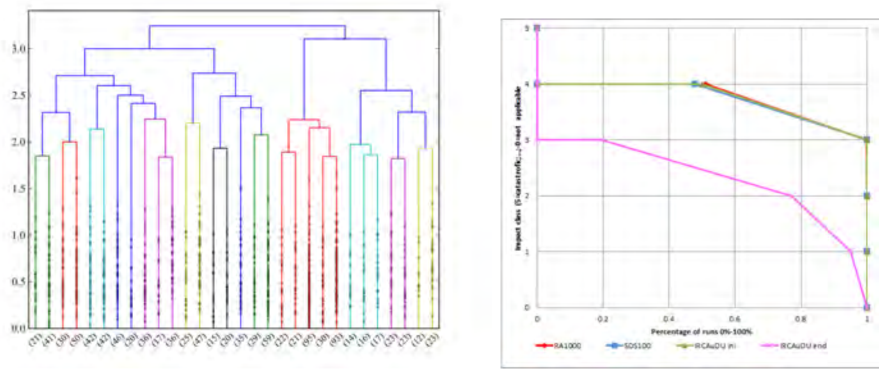


Figure 11: Small representative ensemble of 100 lyme disease scenarios



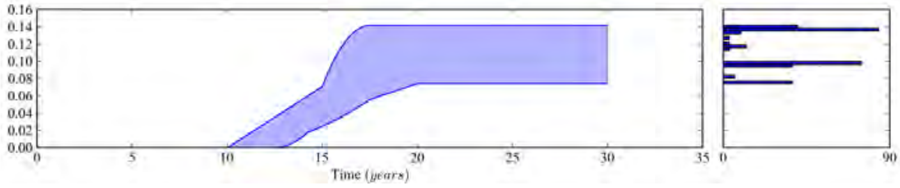
(a) Time series clusterer dendrogram Lyme (b) Risk Envelopes Diagram Lyme

Figure 12: Lyme: Dendrogram (left) and Risk Envelopes Diagram (right)

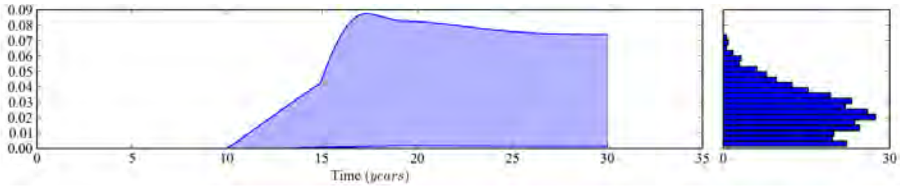
complex and/or deeply uncertain risks. Truly all-hazard and fully integrated IRCAs that would allow dealing with increasing degrees of complexity and deep uncertainty are nevertheless needed more than ever: being able to deal with such risks is ever more important, because most grand challenges and major risks in our ever more complex and ever more uncertain world are dynamically complex and deeply uncertain (Kwakkel and Pruyt, 2012b).

Quantitative as well as qualitative risk assessments, capability analyses, and IRCA approaches are currently used and should be used in the future. We nevertheless argue that the current approaches are not sufficiently integrated and do not appropriately deal with deep uncertainty. The latter critique is addressed in the current paper.

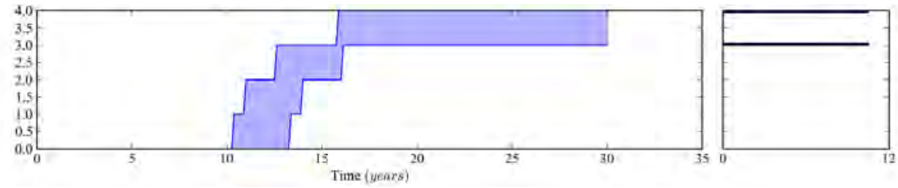
When confronted with deeply uncertain risks, and if scenario variants could be generated easily and cheaply, then it makes more sense to treat these risks



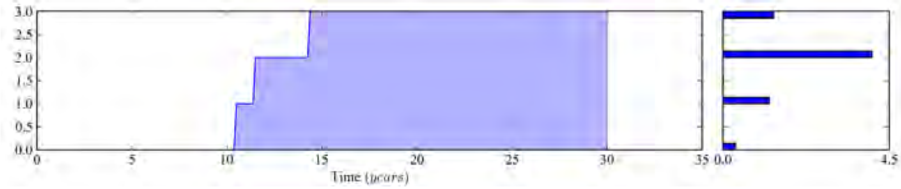
(a) NRA scores Lyme without capability policy



(b) NRA scores Lyme with capability policy



(c) Lyme with capability policy in NRA classes (5 = catastrophic; 4 = very serious; 3 = serious; 2 = substantial; 1 = limited; 0 = not)



(d) Lyme with capability policy in NRA classes (5 = catastrophic; 4 = very serious; 3 = serious; 2 = substantial; 1 = limited; 0 = not)

Figure 13: Lyme: CA without a capabilities policy and with a capabilities policy, expressed in terms of NRA scores ($1 \leq$ catastrophic $\leq 0.33 > \ldots$) and NRA classes (5 = catastrophic; 4 = very serious;...)

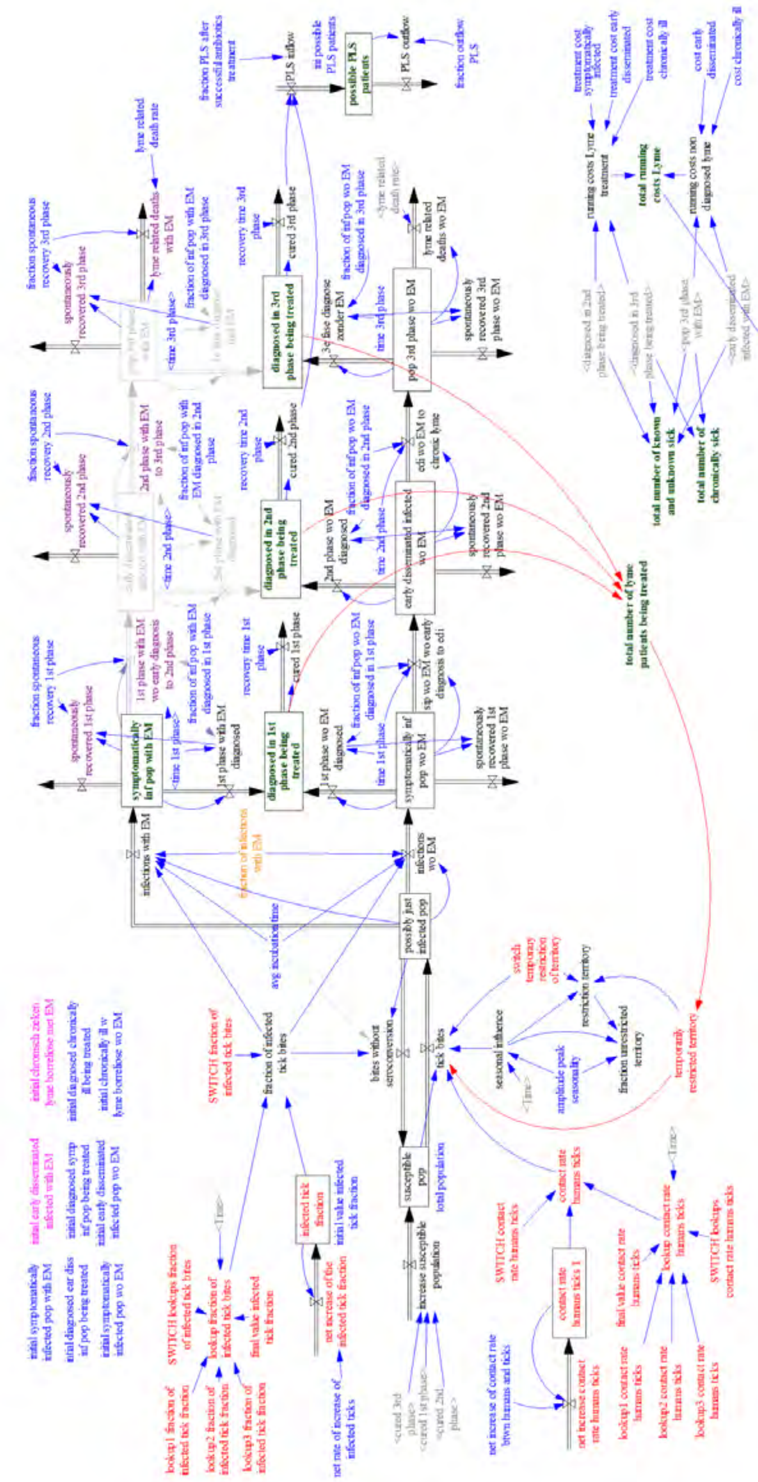


Figure 14: SD simulation model used to generate plausible scenarios wrt lyme disease

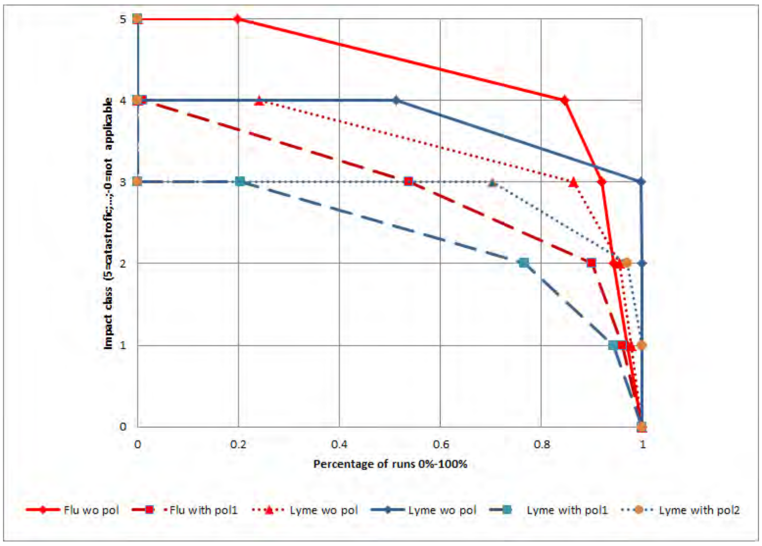


Figure 15: Multi-risk multi-policy risk envelopes diagram: No policy, policy mix 1, and policy mix 2 for flu and lyme disease

as deeply uncertain risks, i.e. to generate the largest plausible set of scenarios and test policies related to investments in capabilities over all these scenarios, or over the smallest set of different scenarios that are genuinely representative of the entire ensemble. The same is true for the mediating effects of investments in capabilities on risks: if these mediating effects are deeply uncertain, then, instead of treating them as unknowable or known, they should be exploited and explored. In this paper, both risks and the mediating effects of capabilities on risks are treated as deeply uncertain. In the near future will it be possible to test and optimize capability sets with uncertain impacts across multiple deeply uncertain risks.

Compared to qualitative approaches, this purely quantitative approach allows dealing with deep uncertainty as well as with dynamic complexity, and allows distilling the most robust sets of capabilities over a plethora of risks. As such, it is complementary to innovative qualitative approaches.

The main difference between the state of science and the state of science as discussed in this paper is thus the explicit versus the lack of consideration of deep uncertainty in scenario development, risk assessment, and capability analysis.

One of the major critiques to be expected against this approach is that it is more time consuming. That is not necessarily true. It is true though that this approach requires EMA software, algorithms, models and modelers. Preferably excellent System Dynamics modelers.

Acknowledgements

This paper overlaps to a large extent with (Pruyt, 2012). The reason for this overlap is due to a lack of progress between the state of this line of work at the time of writing of (Pruyt, 2012) and the state of our work at the time of writing of the current paper. We wanted to include robust optimization over multiple risks in the current paper, but were unable to due to several technical problems related to robust optimization in our workbench, and multi-hazard / multi-model robust optimization. This paper also overlaps to a much smaller extent with two other papers:

- One of the two illustrations, Lyme Disease, is elaborated in more detail in (Pruyt and Coumou, 2012), hence the overlap in diagrams and graphs.
- The other illustration is used in (Pruyt and Kwakkel, accepted). Some figures used in this paper also overlap with figures used in the latter paper.

References

Agusdinata, D., 2008. Exploratory modeling and analysis. a promising method to deal with deep uncertainty. Phd dissertation, Delft University of Technology, Delft, http://www.nextgenerationinfrastructures.eu/index.php?pageID=17&itemID=433706.

Bankes, S., May-June 1993. Exploratory modeling for policy analysis. Operations Research 41 (3), 435–449.

Belton, V., Stewart, T., 2002. Multiple Criteria Decision Analysis: an Integrated Approach. Kluwer Academic Press, Boston, MA.

Bergmans, H., van der Horst, J., Janssen, L., Pruyt, E., et al., October 2009a. Working with scenarios, risk assessment and capabilities in the national safety and security strategy of the netherlands. Tech. rep., Ministerie van Binnenlandse Zaken en Koninkrijkrelaties, The Hague, http://www.minbzk.nl/aspx/download.aspx?file=/contents/pages/102316/nlmethodforriskassessmentandcapabilityanalysing2009.pdf.

Bergmans, H., et al., October 2009b. Working with scenarios, risk assessment and capabilities in the national safety and security strategy of the netherlands. Tech. rep., Ministry of the Interior and Kingdom Relations, The Hague.

Ford, A., 1999. Modeling the environment: an introduction to system dynamics models of environmental systems. Island Press, Washington (D.C.).

Forrester, J., 1961. Industrial Dynamics. MIT Press, Cambridge, MA.

Friedman, J. H., Fisher, N. I., 1999. Bump hunting in high-dimensional data. Statistics and Computing 9, 123–143, 10.1023/A:1008894516817.
URL http://dx.doi.org/10.1023/A:1008894516817

Kwakkel, J., Pruyt, E., 2012a. Exploratory modelling and analysis, an approach for model-based foresight under deep uncertainty. Technological Forecasting and Social Change (accepted for publication).

Kwakkel, J., Pruyt, E., 2012b. Using system dynamics for grand challenges: The ESDMA approach. System Dynamics Review (submitted for publication).

Lempert, R., et al., 2003. Shaping the next one hundred years: New methods for quantitative, long-term policy analysis. RAND report MR-1626, The RAND Pardee Center, Santa Monica, CA, http://www.rand.org/pubs/monograph_reports/2007/MR1626.pdf.

MinVenJ, November 2010. Nationale risicobeoordeling bevindingenrapportage 2010. Tech. rep., Ministerie van Veiligheid en Justitie, The Hague, publicly available since 24 February 2011 on https://zoek.officielebekendmakingen.nl/blg-101376.pdf.

Pruyt, E., 2012. State of art versus state of science of Integrated Risk-Capability Analysis under deep uncertainty. In: Proceedings of the 3rd CESUN Conference. CESUN, Delft, NL, cesun2012.tudelft.nl/images/2/26/Pruyt.pdf.

Pruyt, E., Coumou, J., July 2012. Developing scenarios for deeply uncertain dynamically complex risks: Exploring plausible futures of lyme disease. In: Proceedings of the 30th International Conference of the System Dynamics Society. International System Dynamics Society, St.-Gallen, CH.

Pruyt, E., Hamarat, C., 2010. The Influenza A(H1N1)v Pandemic: An Exploratory System Dynamics Approach. In: Proceedings of the 28th International Conference of the System Dynamics Society. International System Dynamics Society, Seoul, Korea, http://systemdynamics.org/conferences/2010/proceed/papers/P1389.pdf.

Pruyt, E., Kwakkel, J., accepted. Integrated risk-capability analysis in national safety and security: State of the art versus the state of science. Accepted for publication in: Technological Forecasting and Social Change.

Pruyt, E., Wijnmalen, D., 2010. The Dutch National Risk Assessment: An integrated multi-criteria application. Lecture Notes in Economics and Mathematical Systems: Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems 634, 133–143, http://www.springerlink.com/content/l60t423960560140/.

Richardson, G., Pugh III, A., 1981. Introduction to System Dynamics Modeling. Productivity Press: Portland, Previously published by MIT Press.

Sterman, J., 2000. Business dynamics: systems thinking and modeling for a complex world. Irwin/McGraw-Hill: Boston.

Yucel, G., July 2012. A novel way to measure (dis)similarity between model behaviors based on dynamic pattern features. In: Proceedings of the 30th International Conference of the System Dynamics Society. International System Dynamics Society, St.-Gallen, CH.

Yucel, G., Barlas, Y., April/June 2011. Automated parameter specification in dynamic feedback models based on behavior pattern features. System Dynamics Review 27 (2), 195–215, dOI: 10.1002/sdr.457.