EUROPEAN JOINT PROGRAMME
FOR THE INTEGRATION OF
RADIATION PROTECTION RESEARCH

CONCERT

onfidence

Coping with uncertainty for improved modelling
and decision making in nuclear emergencies

EJP-CONCERT

European Joint Programme for the Integration of Radiation Protection Research

H2020 – 662287

D9.33 - Indicators for robust decision making

Lead Author(s): Wolfgang Raskob, Tim Müller (KIT),
Ahmed Nagy, Catrinel Turcanu (SCK•CEN)

Review(s): ST9.1.6 partners and CONFIDENCE all
and CONCERT coordination team

Work package / Task	WP 9 T: 9.1	ST 9.1.1
Deliverable nature:	Report	
Dissemination level: (Confidentiality)	Public	
Contractual delivery date:	CONCERT M37	
Actual delivery date:	29.06.2018	
Version:	1	
Total number of pages:	32	
Keywords:	Robust decision making, visualization of uncertainties	
Approved by the coordinator:	M38	
Submitted to EC by the coordinator:	M38	

Abstract

Decision making in nuclear emergencies heavily relies on the use of results from simulation models, either part of European decision support systems or specialized models that might be used for a particular purpose. So far, uncertainty handling is hardly supported in these mathematical models. Therefore, work package 6 of CONFIDENCE performed research on indicators that allow decision makers to better understand what the performance of different countermeasure strategies is and whether model results are appropriate to use under the given circumstances. Our work focused on two main areas, one related to "robustness" of countermeasure strategies, the other dealing with indicators quantifying the uncertainty in and reliability of results from simulation models in general. The first topic mainly focused on the concepts of robust decision options and robust recommendations in a Multi-Criteria Analysis framework. The second one focused on results from decision support systems and was based on a review performed by work package 1 of CONFIDENCE.

Table of contents

1. General introduction

The attempt to address in a pluralistic way the complexity of nuclear accident preparedness and response management, has prompted multi-criteria decision aid (MCDA) as a promising methodological framework (French, 1996; Hämäläinen et al, 2000a and 2000b; Gelderman et al, 2005), particularly due to its added value in achieving a common understanding and clarification of the complex issues involved. As an appraisal tool, MCDA is deemed capable of handling complex multi-factorial decision problems that affect several stakeholders where an equitable, inclusive and transparent decision process is sought (Gamper and Turcanu, 2015).

Following Stewart (2008), errors and biases occur at any of the different stages of MCDA. He points out that human judgement, data analysis and mathematical/computational processes, and the process as a whole needs to be robust, with respect to:

- External uncertainties, for instance uncertainties in the performance of a given alternative with respect to the different criteria;
- Internal uncertainties, related to the MCDA model parameters, e.g. such as weights or tradeoffs;
- Choice of the preference model, considering for instance the effect of particular value functions chosen to explicate the performance of decision alternatives on particular criteria.;
- Proper identification of stakeholders, criteria and alternatives (see also Dey et al 2016).

External uncertainties derived from the other work packages of CONFIDENCE may encompass various types of uncertainties (French et al, 2017).

Presently, CONFIDENCE WP1 provides an ensemble of weather and source term combinations, each with a certain probability of occurrence. The corresponding impacts of possible countermeasure strategies with respect to different technical criteria (e.g. averted dose, waste generated) is then assessed using the model chains of the JRODOS system for each of these weather and source term combinations. In addition, decisions have to take into account also non-technical factors, that are also affected by uncertainties, but which are modelled outside the JRODOS system (e.g. stakeholder preferences).

This report is structured in two sections.

The **first section** of this document summarises different conceptualizations of robustness in the field of multi-criteria decision aid (MCDA), with the aim of proposing methodologies that can help evaluate the performance of different countermeasure strategies in the presence of uncertainty. This analysis is carried out in view of proposing indicators for the robustness of different countermeasure strategies, to be used in conjunction with the MCDA tool developed under CONFIDENCE WP6.

We explore robustness analysis approaches and concepts dealing with the various type of uncertainties outlined above, that could be used to define robustness indicators. Bouyssou et al (2006) point out the difference between stability and robustness of decisions in multi-criteria analysis. While both types of analyses are recommended, we focus in this report on the former. We follow here Bouyssou et al (2006) who define a robust decision alternative as one that is *"relatively good for all (or almost all) the plausible sets of data and which does not imply too much risk"* (pp. 367). Stability of a decision alternative implies that it maintains a similar level of performance in the presence of some perturbations of the data and/or model parameters (the "most plausible" ones) which were used to

assess the initial performance. Checks for stability usually involve a posteriori sensitivity analysis. For instance, Rios-Insua and French (1991) developed distance based tools for sensitivity analysis, which can be used to determine the smallest changes in model parameters necessary in order to change the ranking of the current optimal solution.

In contrast, the robustness approach will generally seek to incorporate uncertainty in an *a priori* manner such as to identify decision alternatives that are "good" for all or almost all admissible values of data and parameters. A robustness indicator could then be used to express how well a countermeasure strategy performs taking into account the various sources of uncertainty.

The **second section** of the report deals with indicators to support decision makers with information whether various data items (e.g. model or monitoring results) are appropriate to use in the various phases.

So far, uncertainty handling is hardly supported in the mathematical models involved in radiological consequence assessment. Therefore, work package 6 of CONFIDENCE performed research on indicators that allows decision makers to better understand the performance of models and if results are appropriate to use under the given circumstances.

These concepts will be further discussed in the frame of CONFIDENCE and developed depending on the review in workshops and stakeholder panels; the most promising concepts will be realized.

2. Evaluating the robustness of countermeasure strategies

2.1 Introduction

In optimisation and decision aid, the notion of robustness can be operationalised in different ways. Comprehensive overviews of the concept may be found in Dias (2007), Hites et al (2006), Bouyssou et al (2006).

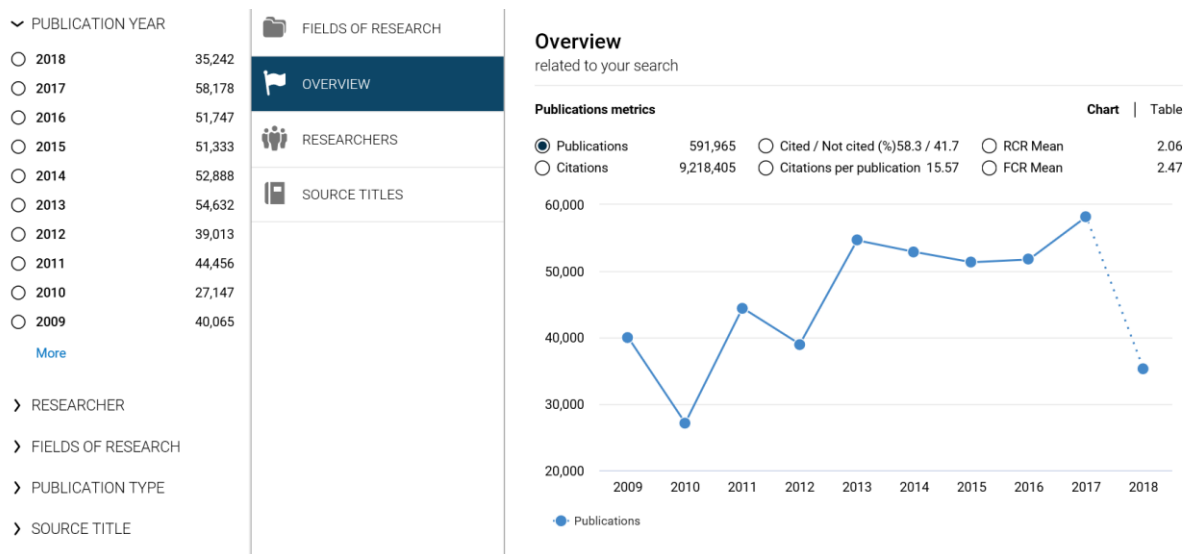In recent years, the concept of robustness in multi-criteria decision-aid has received increasing attention.

Figure 1: Search engine: app.dimensions.ai, search terms: "multi-criteria" and "robustness"

Lempert and Collins (2007) argue that "when uncertainty is well characterized, the cause-effect relationships well understood, and the values clear, the optimum expected utility approach demonstrably yields the best answer", but "these conditions may not hold". In such situations, there is a need for robustness analysis.

Bouyssou et al (2006) highlight however that robustness is not an objective concept since it is strongly dependent on the preferences of the decision-maker. A robust decision alternative may for instance be defined as one that is feasible in "most" (e.g. 95%) of the problem versions, "very good" (e.g. within 5% of the optimal value) in "many" versions and "not too bad" (e.g. among the 10% best feasible solutions) in the others (Bouyssou et al, 2006, pp. 374). Naturally, all these concepts may have different interpretations.

In strategic decisions involving sequential decision-making (Rosenhead et al, 1972), the robustness of a decision is a measure of flexibility, expressing the potential of a decision taken at a given moment to allow for achieving near-optimal states in the future.

Kouvelis and Yu (1997) defined robust solutions for discrete optimisation problems. An "absolute robust" solution exhibit the best "worst-case" behaviour, a "deviation robust" solution minimises the deviation from the best achievable performance relative to each problem version, while a "relative robust" solution minimises the % deviation from the best achievable performance in each problem version.

After Vincke (1999), a multi-criteria decision aid method is robust if the solutions derived from different admissible method-specific parameter sets do not contradict each other. In turn, a robust solution is one that is always near, or does not contradict solutions corresponding to other admissible parameter instances.

Roy and Bouyssou (1993) used the term "robust" to denote a result or conclusion that is not "clearly invalidated" for any parameter instance belonging to the domain of possible values for a decision model's parameters (e.g. weights). Robustness analysis is then defined as the process of elaborating recommendations founded on robust conclusions (e.g. "alternative *a* has a better rank than alternative *b*" or "alternative *a* has the best rank").

Certain types of sensitivity analysis can also be used to tackle robustness issues (French 2003). For instance, in probabilistic sensitivity analysis, uncertainty is modelled through probability distributions for model data and parameters (e.g. performance scores) and is propagated through the MCDA model with Monte Carlo simulations in order to derive rank distributions for the different alternatives (Broekhuizen et al, 2015).

Concerning robustness indicators, recent reviews from McPhail et al (2018) and Giulliani and Castelleti (2016) summarised a number of robustness metrics used in multi-criteria decision aid under deep uncertainty. Such metrics are based on expected level of performance (expected value) across a set of scenarios[1]; the variation in the level of performance (variance, skew), deviation from scenario best (regret-base metrics); worst or best case performance (e.g. maximin); and the range of scenarios with acceptable performance (satisficing metrics). McPhail et al (2018) analyse in detail 11 such approaches and classify them depending on:

- the performance transformation value (e.g. none, regret from best decision alternative, Starr's domain criterion);
- scenario subset (e.g. all, best case, worst case, worst half, worst case and best case, 90th percentile);
- robustness metric applied on the (possibly transformed) performance value (e.g. mean, weighted mean, kurtosis) .

The former study points out that the choice of a robustness metric should depend on the decision-context (suitability of using absolute performance as a reference point), the decision-maker's level of risk aversion and her/his "preference toward maximizing performance, minimizing variance, or some higher-order moment" (McPhail et al, 2018). Most rules analysed assumed high to moderate levels or risk aversion. In a nuclear accident decision context, the decision maker is likely to be inclined towards a prudent attitude, taking into account, to the extent possible, the negative implications of a potential decision. However, Giulliani and Castelleti (2016) emphasize that the risk attitude of decision-makers may evolve in time, depending on favourable or adverse events.

French (1998, pp. 46) analysed in detail four decision rules: Wald's maximin return criterion (Wald, 1950), Hurwickz's optimism-pessimism rule (Hurwicz, 1953), Savage's minimax regret criterion (Savage, 1951) and Laplace's principle of insufficient reason, pointing out that any such rules violates at least one of the principles for consistent choice. For instance, the maximin criterion, which associates to any decision alternative its worst-case performance, does not satisfy the principle of

---

[1] Throughout this document a scenario will denote a problem version, characterised by a particular instance of model data and parameters (Hites et al, 2006).

independence with respect to an improvement (adding a constant term) in the performance of all alternatives in a given scenario.

The next sections in chapter 2 explore recent applications of the concept of robustness in MCDA academic literature, in order to propose definitions of robustness indicators that could help decision-makers to better understand the performance of various countermeasures in the presence of uncertainty in nuclear emergency management. Subsection 2.2 and 2.4 explain the review methodology, objectives and research questions, and the data collection, while subsection 2.4 summarizes the findings. Section 2.5 proposes some approaches to evaluate robustness that might be suitable for CONFIDENCE. Section 2.6 presents our conclusion and future work.

2.2. Review Methodology

An exploratory literature review was carried out following the first three steps of a systematic protocol (Van Solingen et al., 2002): 1) defining the research objectives and research questions, 2) identifying the search process (search terms, keywords and resources, 3) outlining the study selection procedure and the inclusion/ exclusion criteria, 4) outlining the study quality assessment, 5) defining the data extraction approach and 6) defining the synthesis of the extracted data.

This study aimed at exploring possible definitions of robustness that lend themselves to application in a nuclear emergency context, particularly the MCDA approach in CONFIDENCE. For this purpose we analysed academic literature reporting on recent applications of multi-criteria analysis. The work had two objectives:

- **Objective 1:** Exploring terms and definitions related to robustness in recent applications in the MCDA field.
- **Objective 2:** Identifying robustness approaches used in recent MCDA applications that can be useful for emergency planning and response and are compatible with the CONFIDENCE WP6 MCDA approach.

The research questions were formulated using a goal question metric approach (Kitchenham et al., 2009):
RQ$_1$: Which are the most common terms used in the context of decision robustness?
RQ$_2$: What are the different conceptualisations of robustness in recent applications of MCDA?
RQ$_3$: Which robustness approaches can be used to define robustness indicators that could be applied for emergency planning and response in CONFIDENCE?

2.3. Data Collection

An exploratory study was carried out of academic literature. In a first step, articles were collected in August 2017 using the online search platform LIMO of the library of the KU Leuven (Camps et al, 2017) and the following keywords: "Multi-Criteria Analysis" OR "Multi-Criteria Decision" AND "robustness". While this might have led to omission of studies referring to "multi-attribute" or "multi-objective" and not include "multi-criteria", it is sufficiently exhaustive to cover a broad range of applications. The search space included articles written in English language and published between the years 2007 and 2017. This search resulted in 43 articles for the keyword "Multi-Criteria Analysis" and 176 articles for "Multi-Criteria Decision", resulting in a total of 219 articles. As the focus of this study was on identifying potential definitions or uses of the concept of robustness accounting for uncertainty in an *a priori*

rather than *a posteriori* manner, the articles that provided only local sensitivity analysis or did not formulate or refer to a conceptualisation of robustness were not considered for further analysis. This filtering process yielded a number of 51 remaining articles, which give a more or less clear definition on how robustness was interpreted and modelled. For each of the 51 articles we searched for: the MCDA method used, the approach to robustness, the software used to perform the analysis, and the application field of the study (see Table 1 for examples).

Subfield	Definition	Method	Uncertainty Elicitation Method
Disaster management	A robust alternative is an alternative that is simultaneously good in most scenarios and not too risky in any single scenario (Comes et al, 2013)	MAUT	Stakeholders
Scenario planning	A robust alternative has a small relative regret (compared to the best achievable performance for each scenario) across a wide range of plausible futures (Durbach and Stewart, 2012b).	Scenario based approaches	Decision maker specified Probabilities and variances
Aircraft selection	Robustness is the ability to withstand uncontrollable variations in production and usage, yet minimize potential losses due to uncertainties (Sun et al, 2011).	TOPSIS	Analyst specified probabilities
Environmental resources and water management	A robust decision is identified by aggregating the expected value and variance across a number of scenarios (Zarghami and Szidarovszky, 2009).	Stochastic and fuzzy approaches	Stakeholders

Table 1: Examples of the use of the concepts of robustness in recent literature on MCDA applications

For every study/article we extracted the following data items:

[A1]Author
[A2] Year
[A3] Title
[A4] Source
[A5] Venue
[A6] Keywords: {MCDA robustness, decision support robustness, robustness definition, robustness quantification, robust decision, robust emergency management}
[A7] Focus of the robustness research {the method, the solution, the conclusion}
[A8] Robustness definition

[A9] Ways of measuring robustness

[A10] Primary references referring to robustness

[A11] Multi-criteria (decision) analysis / aid methods: {TOPSIS, AHP, ELECTRE, MAVT, etc…}

[A12] Software used

[A13] Application field

[A14] Type of uncertainty taken into consideration: {model parameter, data uncertainties, probability distributions, intervals, scenarios}

[A15] Elicitation of uncertainty information {analyst, stakeholders, expert, combination}

Research question number	Relevant variables
RQ$_1$	[A6], [A7], [A14], [A15]
RQ$_2$	[A8], [A9], [A10]
RQ$_3$	[A8], [A9], [A11], [A13], [A15]

Table 2: The mapping between variables and research questions.

At the next subsequent stage, the articles that referred to methods other than the CONFIDENCE WP6 MCDA approach (MAUT/MAVT) were not analysed further, unless they provided conceptualisations of robustness that can be adapted to MAUT/MAVT.

In a second step, we extended the review with a number of articles identified through a search in the academic literature (journal articles) collected in May-June 2018 using the online search platform SCOPUS and the following keywords: ("Multi-Criteria" OR "Multi-Objective" OR "Multi-attribute") AND ("robustness") AND ("uncertainty"). Articles had to be published in 2007 or later and written in English language. The search space included articles written in English language and published between the years 2007 and 2018 in one of the following fields: Environment, Earth Sciences, Mathematics, Decision Sciences. This search resulted in 50 articles for the keyword "Multi-Criteria", 6 articles for "Multi-Attribute" and 136 for "Multi-Objective". Articles collected in this second stage were filtered based on their abstracts and the following criteria:

- The article addresses only internal uncertainties;
- The article used the term "robust" without indication of a proper development or for a very specific application that is not relevant to disaster management;
- The article uses concepts that are not foreseen in CONFIDENCE WP6, e.g. modelling of decision-maker's preferences with fuzzy sets, or addressing a classification rather than ranking problem.

Instead of an exhaustive analysis of all the remaining, newly collected articles, we identified a number that bring something new compared to the previous literature search; these were considered for further analysis.

2.4 Review Summary

The articles reporting on conceptualisation or application of robustness analysis address one or both of the following issues: i) how to generate or select the relevant problem versions; and ii) how to evaluate the robustness of decision alternatives against the problem versions selected or generate.

i) Identifying robust decision alternatives

Agusdinata et al (2009) define robust policies through a rule similar to the minimisation of the maximum "regret", whereby regret is calculated as the distance between a given policy and the maximum performance achievable in a particular circumstance. The farther the performance of a selected policy is from the maximum performance, the higher the regret values are (see also section 2.5.1).

Santos et al (2017) propose the use of semi-deviation below a benchmark to assess robustness of decision strategies, as this avoids labelling an alternative as low-risk when presenting low variability, and it allows specifying the a threshold below which the decision maker considers risk as unacceptable.

Comes et al (2011) present an approach to construct scenarios and develop Decision Maps to facilitate identification of robust decision options, understood as performing on average sufficiently well for a set of scenarios, or guaranteeing a minimum performance for all scenarios. A visual representation is used to display a graph of the performance of the different alternatives in the context of the scenario leading to worst, medium and best performance scenarios for each alternative. This presents the information to the decision maker in a cognitively smooth manner. Following Vincke (1999) and Hites et al (2006), Comes et al (2013) utilise in a more recent study a MAVT approach to identify robust alternatives as those that perform well under a number of scenarios used to model 'what could go wrong?'. The robustness value of a decision alternative in MCDA is assessed by means of linear aggregation of performances corresponding to different scenarios methods. The approach is useful to shed light on different situations that need to be taken into consideration whether short term or long term.

Tom et al. (2018) identifies robust strategies among the Pareto optimal solutions by calculating the ratio between the scenario combinations for which the solution meets certain constraints and the total number of scenarios considered.

Durbach and Stewart (2012a) point out that when using scenarios with multi-criteria analysis"is how (or even whether) to compare and aggregate results from different scenarios". They recommend that aggregation should use "swing" weights rather than scenario probabilities, since the set of scenarios does not constitute a complete probability space, or likelihoods, "because scenarios are incomplete descriptions". They also argue that while the use of swing weights or probabilities has been criticized, the frequently encountered alternatives in scenario planning (e.g. worst performance or regret based) can in some cases be misleading.

Scholten et al (2014, 2015, 2017) use the term robust to refer to alternatives that perform consistently well over a number of scenarios developed in close consultation with stakeholders. In Scholten et al (2015) expected utility values are calculated based on probability distributions for attribute scores. In addition, they employed a comprehensive uncertainty analysis on preferences parameters (aggregation function, marginal value function curvature, utility function curvature, criteria weights) through a probabilistic description, in order to derive rank probability distributions for each decision alternative in each of a number of scenarios.

A large number of studies use uncertainty analysis on model parameters with Monte Carlo simulations to derive and / or visualise the probabilistic ranking of decision alternatives (Mendecka et al, 2017; Troldborg et al, 2014) and range of variation for each decision alternative (El Hanandeh et al, 2010). To avoid computationally intensive Monte Carlo simulations, Zucca et el (2008) assessed robustness with respect to criteria weights using a selected set of weights representing the points of view of different stakeholders.

Dalalah et al (2011) use a visual representation of the overall performance of different strategies depending on the values of a specific model parameter. This can help the decision-maker to assess the robustness of a solution in a qualitative way, but is however only applicable to varying one parameter at a time.

Zarghami and Szidarovszky (2009) assesses the robustness of an alternative, in the presence of uncertainty in the model data and parameters, as the difference between the expected value of the alternative and the product between the variance of its outcome and a coefficient modelling the risk attitude of the decision maker (see also 2.5.2). Instead of combing the two measures, Ligmann-Zielinska, A. and Jankowski (2014) evaluate as robust the options with high average outcome and low variance. Based on Vincke (1999), Montibeller et al (2006) define robust options as those close to the ideal performance in all scenarios. They use the deviation from the ideal performance as a measure of inter-scenario robustness. An interesting remark is that in their case study the decision makers preferred different weights for the two scenarios considered. Montibeller and Franco (2011) further define also inter-scenario risk, which represents the spread of values across different scenarios. They suggest that in practice, the most helpful way of supporting decision-makers' choices has been a visual inspection of the performances of the different decision alternatives and the spreads of their performance, with a focus on inter-scenario robustness and inter-scenario risk.

The work of Zhou et al. (2018) follows a Gaussian processes for constrained regression to design optimization for uncertainty quantification in engineering products. The model is expressed by identifying the expected values for the different parameters and calculating the covariance of the different variables of interest, then modelling the observations through Gaussian regression. The solutions are categorised in three sets 1) feasible robust 2) feasible deterministic 3) design under consideration. The robustness of the solution is quantified using two indicators 1) objective robustness and 2) feasibility robustness where the authors define them as follows. Objective robustness requires that the variation is within the acceptable objective variation range, pre-specified by decision makers. Feasibility robustness requires that the constraints are not violated due to perturbation of the uncertain variables, even under a worst-scenario situation.

Lempert and Collins(2007) use three robustness approaches and compare the results with the optimal utility and the precautionary approach for a problem characterised by uncertainty probability distribution associated to potential outcomes. The first robustness approaches (Lempert et al, 2006) assumes trading of some optimal performance (expected regret for the distribution leading to smallest expected regret) for less sensitivity to assumptions (expected regret for the distribution leading to largest expected regret), in a manner similar to the Hurwicz optimism-pessimism rule (Hurwicz, 1953). The second approach looks at satisficing alternatives over a wide range of futures, whereas the third is inspired by Rosenhead's idea of "keeping the options open". Their study concludes that robust decision-making should be preferred over the optimal approach when their uncertainty is sufficiently deep and the decision alternatives sufficiently rich. They also argue that the robust decision-making approach captures the underlying concept of the precautionary principle.

Kassab (2011) employs the information gap theory to assess the robustness of decision alternatives in the presence of uncertainty in the goals and preferences of the stakeholders. The question of robustness is reformulated as: *how wrong can the models and data be, without jeopardizing the quality of the outcome*? (Regan et al, 2005) (see section 2.5.3).

Exploratory model based approaches are highly suitable for supporting robust planning with staked/ deep levels of uncertainty (Weaver et al. 2013). The key design principles of robust decisions under adaptive policies is to develop flexible plans that can be adapted over time in response to the variability

of the uncertainty levels as a part of long term decision making. The inclusion of uncertainty in long term planning has resulted in the development of new model based approaches which include Dynamic Adaptive Policy pathways (Haasnoot et. al., 2013), Many-Objective Robust Decision Making (Hadka et al 2013, 2015) and robust decision making (Groves and Lempert 2007).

Korteling et al (2013) explored Info-Gap decision theory (see section 2.5.4) to evaluate all the uncertain parameters with the two trajectories that quantify robustness and opportuness, instead of randomly running a plethora of possible futures. The most robust decision alternative is the defined as one that delivers the same performance, (equal or better than the critical reward criteria), as other alternatives at higher levels of uncertainty. The most opportune alternative is the one that delivers the same performance, (equal or better than the minimum reward criteria), as other alternatives at lower levels of uncertainty.

Finally, Tarun et al (2011) use robustness expressed as insensitivity to the variation of the model inputs as a criterion in the MCDA analysis.

ii) Formulating robust recommendations

Angelo et al (2017) and Diaby and Goeree (2014) use the approach of Dias and Climaco (2000) for robustness analysis in order to calculate the maximal regret and the range of values that a decision alternative can take, in the presence of uncertainty in the criteria weights. This approach assumes that the set of acceptable criteria weights is described by a number of given linear constraints. Robust conclusions are identified as those compatible with this incomplete information on the weights provided by the decision makers (e.g. which options are dominated).

Schuwirth et al (2012) investigate the robustness of MCDA modelling assumptions by checking whether rank reversals occur in the presence of uncertainty in e.g. criteria weights or the shape of the utility function (risk neutral vs. risk averse), among others.

iii) Robust methods and models

Durbach and Stewart (2012b) compared the classical MAUT model with simplified models, e.g. quantile models (using the 5%, 50% and 95% quantiles) and scenario models (with different percentages of coverage). They refer to robustness in terms of accuracy (loss of utility) in the presence of assessment errors. Their results suggest that if analysts lack the resources to implement MAUT, a quantile model could be used instead, where the weight on the median is 0.63 and the remainder is shared between the extreme quantiles. They also draw attention that the use of scenarios has to done carefully in order to avoid important omissions, and that a scenario model, even if correctly applied, will lead to an outcome that is more different to MAUT than other simplified models.

Ali et al (2017) defined robustness as a subjective measure used to make "correct" predictions based on noisy data that might have missing values.

Norese and Carbone (2014) refer to 'robust' models that are able to "include and synthesize different knowledge elements, to be analysed, validated and improved both during the development process and in the future, and to propose reliable interpretations of the examined situation " .

2.5. Robustness indicators with potential application for CONFIDENCE WP6

2.5.1. Maximin indicator

The *maximin* indicator or Wald metric (Wald 1950) associates to any decision alternative *a* its worst-case performance.

$$R(a)= \min \{ f(a, s) \mid s \text{ scenario}\},$$

where f(*a*, *s*) is the performance of alternative *a* under scenario s. A scenario s denotes a plausible combination of model data and parameters.

This metric is associated with a pessimistic point of view as it assumes that the worst will happen. The decision option maximising R corresponds to the absolute robust solution in the sense of Kouvelis and Yu (1997).

2.5.2 Regret-based indicator

If instead of maximising the worst-case performance, we aim at minimising the deviation from the best achievable performance under each scenario, we obtain the deviation robust solution (Kouvelis and Yu, 1997). The robustness of an alternative *a* would then be assessed with Savage's rule (Savage, 1951):

$$R(a) = \max_s \{ \max_b f(b,s) - f(a,s) \mid s \text{ scenario} \}$$

This represents a more realistic and less conservative point of view of the decision-maker, compared to maximin criterion.

2.5.3. Expected value based indicators

When external uncertainties are modelled in a probabilistic way, the robustness of a decision alternative *a* could be also assessed on the basis of its expected value:

$$R(a) = E(a) = \int_s f(a,s) \cdot p(s) ds$$

where *p(s)* is the probability of scenario *s*.

When the distribution p is unknown, the arithmetic average could be used instead (Laplace's principle of insufficient reason).

To take into account the risk attitude of the decision maker, Zarghami and Szidarovszky (2009) propose to include also the variance of performance:

$$R(a)= E(a) - \beta*Var(a)$$

where β is a positive weight showing the "importance of decreasing risk (variance) in comparison to maximizing the expected payoff". The higher the risk aversion of the decision maker, the larger should be selected the value of β.

Walsh et al (2013) extending the work of Gluck et al. (2012) apply the expected value on the "functionality" of a decision alternative, instead of its performance, in order to quantify robustness. Functionality refers to the property whereby certain goals can be achieved in an acceptable time frame:

Functionality (*a, s*) = *Success*(*a, s*) − *Failure*(*a,s*)/T, where

Success (*a, s)* = probability (or rate) of success under scenario s, e.g. collective dose is kept below a certain threshold

Failure (*a*, s) = probability (or rate) of failure under scenario s

T= tolerance (parameter set by decision-maker)

A negative functionality indicates that the system exceeds allowable risk, while a positive value of functionality indicates that the system remains safely within allowable risk.

The robustness of decision alternative *a* can then be assessed as (Walsh et al, 2013):

$$R(a) = E(\textit{Functionality (}a)) = \int_{s} Functionality(a,s) \cdot p(s)ds$$

2.5.4. Info-Gap based indicator

Info Gap Decision Theory (IGDT) (Ben-Haim, 2006) was introduced to assist decision-making in the presence of severe knowledge gaps, when probabilistic models of uncertainty are unreliable, inappropriate, or unavailable.

Assume the expected utility model:

$$E(a_j) = \sum p_i \cdot v_{ij} ,$$

assigning to each alternative a_j its expected utility, calculated on the basis of its performance v_{ij} under scenario s_i, and the probability p_i of scenario s_i.

When uncertainty in the probabilities and utilities may be represented by intervals of unknown size around each alternative, the information gap is defined as the fractional deviation from the nominal value. For instance, the information-gap model for utility uncertainty is the family of nested intervals given by:

$$\frac{|v_{ij} - \widetilde{v_{iJ}}|}{\widetilde{v_{iJ}}} \leq \alpha$$

where $\alpha \geq 0$ is called the uncertainty horizon. In this model, the performance v_{ij} of alternative a_j varies from its nominal value $\widetilde{v_{iJ}}$ by no more than a fraction α.

The info-gap model for uncertain utilities is defined as (Regan et al, 2005):

$$U_v(\alpha, \tilde{v}) = \{v \mid \max(0, (1 - \alpha) \cdot \widetilde{v_{iJ}}) \leq v_{ij} \leq \min(1, (1 + \alpha) \cdot \widetilde{v_{iJ}}), \forall i = 1,..n; \forall j = 1..m\},$$

where n is the number of scenarios and m is the number of decision alternatives.

The same conceptualization can also be used to model uncertainties in the probability values p_i and the info-gap model for probabilities, taking into account the additional constraint that the sum of probability values adds up to 1.

The robustness of a decision alternative a_j is defined as:

$$R(a_j, E_c) = max \left\{ \alpha \mid \min_{\substack{v \in U_v(\alpha, \tilde{v}) \\ p \in U_p(\alpha, \tilde{p})}} E(a_j) \geq E_c \right\}$$

where E_c is a minimal acceptable level for the expected utility.

2.6. Conclusion for robustness indicators

Under deep uncertainty the decision makers in a radiological emergency face multiple challenges due to the absence of solid reliable information. As a result, there is a need to adopt policies and decision that can perform satisfactorily under a wide range of futures. Evaluating the level of robustness of a policy or decision alternative can support the decision maker in this task.

We proposed multiple approaches to quantify the robustness of decision options, that can adapted for use in CONFIDENCE and tested in stakeholder panels. The final aim of using robustness indicators is not to replace the MAUT with another decision rule, but to provide additional information allowing the decision maker to understand the potential impact of various decision options. It can be envisaged to use one or more robustness indicators, or even to add one of these as an additional criterion to the MAUT analysis.

Though robustness measures offer insight at a snapshot in time, there may be a need for approaches that are flexible enough and can evolve coherently as information becomes available and deep uncertainty declines. In that sense, the adaptive policy methodologies may offer a suitable approach (Kwakkel et al, 2016; Hasnoot et al, 2013). sThe current work focused on identifying robustness quantification approaches and the construction of robust policies. Adaptive visual policy maps (Hasnoot et al, 2013) can complement the information needed to understand the robustness of different decision options and to unload the decision makers from the cognitive task associated with memorizing the different consequences and time frames for the different policies. Applying visual quantitative approaches can offer a double check to ensure flexible robust policies especially under deep uncertainty.

Combining visual decision aids and quantitative robustness measures can help decision makers navigate multiple levels of deep uncertainty. Visualization of dynamic adaptive strategies can for instance offer an effective approach to identify and switch among robust policies to meet identified objectives. For our future work, visualization of uncertainty will be tackled as part of CONFIDENCE WP 6.2 where multiple cartographic techniques and approaches will be investigated to identify effective uncertainty visualization techniques. In addition, robustness calculation can be complemented with cartographic visualization approaches in order to simplify robustness and uncertainty communication to the decision makers. Future research can bring cartographic approaches to uncertainty visualization and robustness under radiological emergency scenarios. Our future work focuses on demonstrating the usage of robustness calculation based on the methods presented, in addition to exploring the use adaptive dynamic policies to radiological emergency scenarios.

3. Visualization of model results

3.1 Introduction

Decision making heavily relies on the use of results from simulation models, either part of a model chain in European decision support systems or specialized models that might be used for a particular purpose. So far, uncertainty handling is hardly supported in these mathematical models. As part of WP1 of CONFIDENCE, the uncertainties in early phase models, in particular atmospheric dispersion and dose models were analyzed. In addition, ensembles of meteorological forecast data will be used to describe the variability of the weather.

However, ensemble results might be difficult to interpret by the decision maker. Therefore, our aim is to develop indicators to categorise results of simulation models in decision support systems (JRODOS will be used as example) as appropriate, or not, for decision making in an evolving exposure situation. The indicator system is a visualization scheme that can be added to each result. For instance, dose assessments based on source term estimations in the very early phase are very uncertain but they become more reliable after days/weeks. In this respect, appropriateness considers mainly uncertainty of model endpoints in relation to uncertain input and not necessarily the quality of the model. However, the quality of the model as such has to be considered when defining indictors.

This indicators system should provide guidance on the use of results in the various phases of an emergency (see Figure 2).
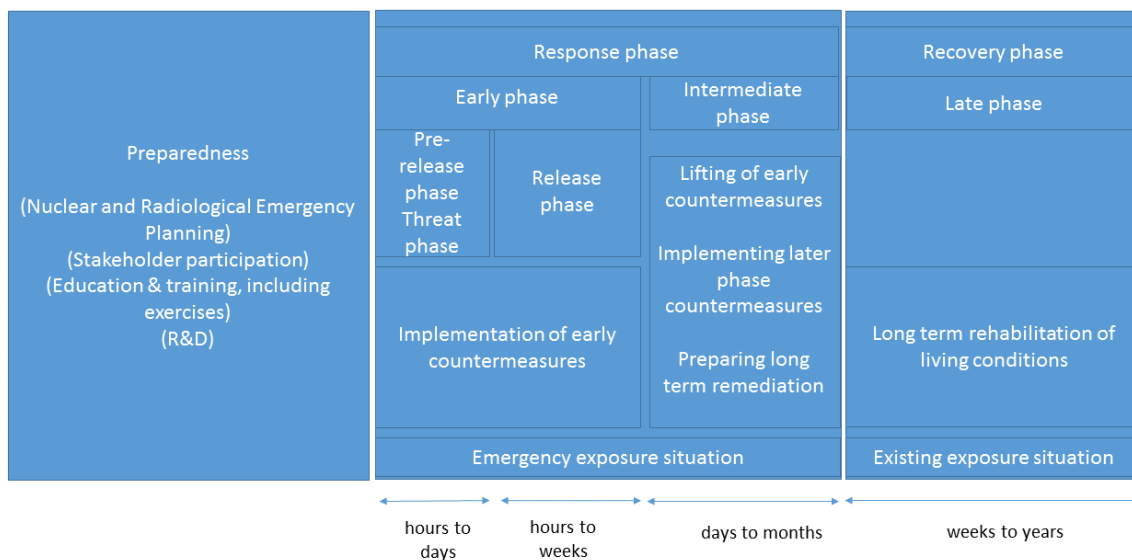

Figure 2 Phases of an emergency (NERIS SRA 2017)

In the Urgency phase – often referred to as "Early" phase, decisions are based on consequence modelling using simulation tools and prognostic information about source term and weather data. However, decisions about urgent protective actions such as sheltering, evacuation and iodine prophylaxis are taken before the arrival of the plume to achieve the best possible efficiency (doses are minimal if evacuation is completed before plume arrival).

The Post-accident phase can be subdivided into a phase of transition to recovery and a long-term phase. In the transition phase, the measures initiated in the release phase will be revised and possibly lifting announced. In this phase, the use of monitoring is dominating over model calculations and prepares the basis for long term decisions.

The long-term recovery phase aims to return social and economic life on affected territories to normal conditions. Measures are based mainly on monitoring, however, predictions on the time dependent concentrations in feed- and foodstuffs are still important and require modelling. This is also true for decontamination as also here simulation models help to define best possible remediation strategies.

The visual representation of reliability (uncertainty) has to be self-explaining. In this respect, the "traffic light" colors that are widely used in risk communication (see e.g. Aven and Renn, 2010) might be a possible candidate for the display on screen and in the various result sheets (e.g. in risk communication: red signals intolerable, yellow tolerable and green acceptable). Cultural differences and colour blindness should be taken into account in further work.

Further possibility is to grey out results that might be not used for decision making in that particular phase. However, a general issue is the use of a particular endpoint in various phases of an emergency.

On the other hand, a good indicator might be derived on the basis of the simulation results. For example concentrations in feed- and foodstuffs based on monitoring information might be more reliable compared to results based on initial concentrations from atmospheric dispersion calculations. All this is discussed in the following chapters of this report.

The indicator system proposed will be implemented into the JRODOS system as well as in other simulation and representations systems and demonstrated in at least one of the stakeholder sessions of CONFIDENCE WP6. Based on the feedback, the approach will be revised and finally adapted.

This report concentrates on the early phase as here decisions on early phase countermeasures have to be taken with highest uncertainty.

3.2 Key uncertain parameters and associated endpoints

In deliverable D9.1 of the CONFIDENCE project, guidelines describing atmospheric dispersion modelling uncertainties were developed (Mathieu and Korsakissok, 2018). From these guidelines, the key parameters related to atmospheric dispersion. Ensembles of meteorological forecast information as well as of source terms may also play an important role in generating indicators. For example the spread of the weather data might be an indicator about the quality of the prognosis and the stability of the weather conditions.

The indicator system proposed will not account for sensitivities of internal model parameters as this will exceed the intention of this indicator system. More important is the indication if a particular model is suitable for a given purpose.

3.2.1 Atmospheric dispersion and deposition

Mathieu and Korsakissok (2018) rank the following parameters for atmospheric dispersion models as most important.

Parameter	Uncertainty/Effect	Rank
Wind direction	• Direction changes can result in large impacts on the dispersion results as a moderate change in direction can significantly alter air concentrations. • The timing of the wind direction change is important when combined with the timing of changes in the release rate. • The importance of wind direction to the response may depend on the phase of the response or the countermeasures taken. For example, if the countermeasures involve evacuating everybody within 5km of the nuclear power plant then wind direction will not impact this countermeasure.	1
Stability	The stability may be very important for certain types of release (e.g. an elevated release) and in the vicinity of the source. Stability may also be more important for some model configurations than others.	1 (early phase)
Precipitation		2
Wind speed	Wind speed is moderately important to dispersion calculations	3
Mixed layer depth	The importance of the mixed layer depth depends on the situation and the meteorology. For example, if the plume is entirely within the mixed-layer then the mixed-layer depth is not important but if the plume is likely to be fumigated into the mixed-layer then the timing of the mixed-layer depth change can be very important.	4

Table 3 Rank of the sensitivity of dispersion models to input meteorological variables from Mathieu and Korsakissok, 2018

Based on Table 3, results linked to wind direction (all dispersion calculations) and precipitation (wet deposition) have to be considered.

As mentioned before, ensembles are used to represent uncertainties in the atmospheric dispersion and deposition models. Therefore, the widespread of ensemble results might be a further indication if results are suitable for decision making at that stage.

Source terms in the threat or release phase are highly uncertain if not monitored via the stack. As long as the release is monitored via the stack, source term information can be directly transferred to simulation models. However, releases via stack are typically much lower compared to releases from buildings resulting from a core melt. Assessments in Germany clearly indicated that areas for early countermeasures for venting or filtered venting source terms are much smaller than those from releases following a core melt (Walter et al. 2015).

Dispersion models of different complexity, e.g. Gaussian and Lagrangian particle type, have different modelling capabilities, however, it is difficult to judge which results are preferable. There is one clear exception and this is changing wind direction with height and complex terrain. Under these conditions, Lagrangian particle models are favored against Gaussian type approaches.

In general, usage of simulation models has intrinsic uncertainties but just listing them does not help decision makers to find appropriate answers. Therefore, other model parameters such as Precipitation, scavenging coefficients, dry deposition velocity, surface resistance, particle size distribution, vertical diffusion parameters, horizontal diffusion parameters were not taken into account for the indicator system.

3.2.2 Dose estimation and countermeasure areas

Doses are typically calculated by multiplying a concentration value (either from simulation or monitoring) by a dose conversion factor (DCF) and further parameters such as breathing rate for the inhalation dose or shielding factors when looking in particular for external doses.

Dose estimations contain a second layer of uncertainty with their internal parameters such as the DCF and other input parameters such as the shielding and breathing rates. In particular the shielding parameter is highly uncertain and is typically derived from population density data (JRODOS) or set to mean values representing a particular environment (e.g. rural or urban)

Areas for countermeasures are one of the most important endpoints for decision making in the early phase. Investigations with different dispersion models in JRODOS (Gaussian and Lagrangian particle type) indicated that even if concentrations differ, areas for countermeasures differed less. This might be the result of a yes or no decision if a particular dose limit is exceeded or not.

3.2.3 Food chain modelling

Parametrisations of food chain models are not discussed here. We also no include particular diets in our assessment, even if they are available in the models. In general, dietary information is difficult to judge as people might change them following an accident. However, key radionuclides such as iodine, cesium and strontium are surely investigated in more details compared to those that were not so prominent in past accidents. Therefore, we propose only to indicate if the results are based on input from simulation models or based on monitoring information. If the latter is the case, results should be indicated as more reliable.

3.2.4 Inhabited area modelling

Similar to foodchain modelling, we only consider the input data as important and distinguish between user-defined input, input based on monitoring and input from atmospheric dispersion simulations; the latter one has to be characterized as most uncertain.

3.2.5 Aquatic modelling

Aquatic modelling is less advanced compared to atmospheric dispersion. Even if after Fukushima a lot of effort was devoted to improve modelling, in particular in the marine environment (e.g. PREPARE and COMET projects), gaps still exist (J. Vives i Batlle et al., 2018 and T. Duranova et al., 2016).

In general, surface run-off is the most demanding task and results are most uncertain for particular events. In the longer term when interested in averages over months or years, results become more reliable (e.g. MOIRA suite for long-term behavior in catchments). Based on experience with the Hydrological Model chain (HDM) of JRODOS, uncertainty ranking is still difficult and requires further exploration.

3.3 Important endpoints for decision making in the various phases

3.3.1 Threat and release phase

Decisions are typically based on doses (particular criterion is exceeded in areas for protective measures) and concentration levels (activity in foodstuffs and cloud for identifying the arrival of the main activity).

At present, source term reconstruction based on monitoring information and dispersion simulation (part of PREPARE project, see Duranova et al, 2016) and assimilation of concentration in foodchain (CONFIDENCE) has been studied or the work is under way. Results based on these two approaches might also result in better performance, however not necessarily of similar quality as based on monitoring.

3.3.2 Transition and recovery phases

Monitoring becomes more and more important in the longer-term assessments. In this respect, as soon as results of simulations, e.g. foodchain, are based on monitoring, their performance should be higher compared to results based on dispersion models. Data assimilation at this stage helps to improve monitoring in areas where no monitoring information exists but only modelling, and thus should be indicated with the same quality factor as monitoring alone.

3.4 Indicators for the adequacy of particular data as a basis for decision-making

The indicators proposed in the following are a first attempt and have to be discussed in the panels of CONFIDENCE WP6 and decision makers as well as experts in future. One should also be aware that a too detailed subdivision of indicators might be hard to argue as this would require details about all simulation models and their parameters used that are difficult to obtain. Nevertheless, this might be a topic for further research. One could imagine also five colors with green/yellow mark indicating the transition between green and yellow and yellow/red mark indicating the transition between yellow and red. Both schemes are used for the following tables. The red color should indicate that the results should be taken with care and only used if no other alternatives with yellow or green are available. In general, result are preferred in this order of colors green -→ yellow -→ red. The colors themselves might change, but a particular order is important.

Endpoint	Early phase (pre-release and release)	Early phase based on ensemble modelling	Early phase based on data assimilation (food and source term)	Transition phase	Long-term recovery phase
Dose maps	red	yellow	yellow	yellow	green
Dose rate maps	red	yellow	yellow	yellow	green
Countermeasure areas	red	yellow	yellow	yellow	green
Plume arrival time	red	yellow	yellow	n.a.	n.a.
Concentration in feed and foodstuffs	red	yellow	yellow	yellow	green
Concentration in rivers from run-off	red	n.a.	n.a.		
Concentration in rivers from direct release	red	n.a.	n.a.	yellow	yellow
Concentration in lakes and reservoirs	red	n.a.	n.a.	yellow	yellow
Concentration in marine food products	red	n.a.	n.a.	yellow	yellow
Inhabited area countermeasures	red	yellow	yellow	yellow	green
Food countermeasures	red	yellow	yellow	yellow	green

Table 4: indicator system based on three colours

Endpoint	Early phase (pre-release and release)	Early phase based on ensemble modelling*	Early phase based on data assimilation (food and source term) **	Transition phase	Long-term recovery phase
Dose maps	red	yellow	Yellow/red	yellow	green
Dose rate maps	red	yellow	Yellow/red	yellow	green
Countermeasure areas	Yellow/red	yellow	Yellow/red	yellow	green
Plume arrival time	Yellow/red	yellow	n.a.	n.a.	n.a.
Concentration in feed and foodstuffs	red	yellow	Yellow/green	yellow	green
Concentration in rivers from run-off	red	n.a.	n.a.	yellow	
Concentration in rivers from direct release	Yellow/red	n.a.	n.a.	yellow	yellow
Concentration in lakes and reservoirs	red	n.a.	n.a.	yellow	yellow
Concentration in marine food products	red	n.a.	n.a.	yellow	yellow
Inhabited area countermeasures	red	yellow	yellow	yellow	green
Food countermeasures	red	yellow	yellow	yellow	green

Table 5: indicator system based on five colours (* indicates that performance should be increased if ensembles are very close; ** as long as release is ongoing assimilation of source term is uncertain)

Indicators will be integrated into example results of JRODOS and other simulation models. In each of the relevant map results, an additional icon will be introduced. For example for the "yellow/red" indicator, the following might be used. ⬤⬤◯ Possibly a heading indicating the purpose of the characterization is also necessary. Once these indicators are implemented, they will be used in CONFIDENCE WP6 panels to be discussed with experts and decision makers. At present, two panels are planned for this. Following the discussion in the consortium, the color code will be further analysed. In addition, dyschromatopsia (red-green problem) and black and white issues will be further investigated. However, once the technical basis is defined, the color code can be changed easily according to the request of the end users.

3.5 Conclusions for visualization

The review on uncertainty indicators demonstrated, that such a characterization of results from simulation models is not straightforward. Detailed and specialized characterization of endpoints is difficult as the uncertainties of input data and internal model parameters are large. Therefore, only the characterization of result classes in the time frame of an ongoing emergency is proposed. Generally, one can state that performance of models might be better in the later phase as input parameters are more or less validated. Further to this, also the issue of ensemble simulation might show the range of applicability of model results and might receive therefore a higher rank.

Important is now to demonstrate these approaches in real applications to ask for advice from experts, in particular from the various panels of CONFIDENCE. This will be done on one hand side within JRODOS but also in other GIS-based tools.

4. References

Agusdinata, D. B., van Der Pas, J. W. G. M., Walker, W. E., & Marchau, V. A. W. J. (2009). Multi-criteria analysis for evaluating the impacts of intelligent speed adaptation. Journal of advanced transportation, 43(4), 413-454.

Ali, R., Lee, S., & Chung, T. C. (2017). Accurate multi-criteria decision making methodology for recommending machine learning algorithm. Expert Systems with Applications, 71, 257-278.

Angelo, A. C. M., Saraiva, A. B., Clímaco, J. C. N., Infante, C. E., & Valle, R. (2017). Life Cycle Assessment and Multi-criteria Decision Analysis: Selection of a strategy for domestic food waste management in Rio de Janeiro. Journal of Cleaner Production, 143, 744-756.

Aven T. and Renn O. (Eds.) (2010), Risk Management and Governance: Concepts, Guidelines and Applications, Springer Heidelberg, DOI 10.1007/978-3-642-13926-0

Ben-Haim Y. (2006). Info-gap decision theory: decisions under severe uncertainty. 2nd ed. Oxford, UK: Academic Press.

Bouyssou, D., Marchant, T., Pirlot, M., Tsoukiàs, A. and Vincke, P., 2006. Evaluation and decision models with multiple criteria: Stepping stones for the analyst (Vol. 86). Springer Science & Business Media.

Broekhuizen, H., Groothuis-Oudshoorn, C. G., van Til, J. A., Hummel, J. M., & IJzerman, M. J. (2015). A review and classification of approaches for dealing with uncertainty in multi-criteria decision analysis for healthcare decisions. Pharmacoeconomics, 33(5), 445-455.

Camps A., Turcanu C., Nagy A. (2017). The Use Of The Concept Of Robustness In Multi-Criteria Decision Analysis: A Review Of Recent Literature. Internal report. Belgian Nuclear Research Centre SCK•CEN: Mol.

Comes, T., Hiete, M., & Schultmann, F. (2013). An Approach to Multi-Criteria Decision Problems Under Severe Uncertainty. Journal of Multi-Criteria Decision Analysis, 20(1-2), 29-48.

Comes, T., Hiete, M., Wijngaards, N., & Schultmann, F. (2011). Decision maps: A framework for multi-criteria decision support under severe uncertainty. Decision Support Systems, 52(1), 108-118.

Dalalah, D., Hayajneh, M., & Batieha, F. (2011). A fuzzy multi-criteria decision making model for supplier selection. Expert systems with applications, 38(7), 8384-8391.

Dey, B., Bairagi, B., Sarkar, B., & Sanyal, S. K. (2016). Multi objective performance analysis: A novel multi-criteria decision making approach for a supply chain. Computers & Industrial Engineering, 94, 105-124.

Diaby, V., & Goeree, R. (2014). How to use multi-criteria decision analysis methods for reimbursement decision-making in healthcare: a step-by-step guide. Expert review of pharmacoeconomics & outcomes research, 14(1), 81-99.

Dias, L., Clímaco, J. (2000). Additive aggregation with interdependent parameters: the VIP-analysis software. J. Oper. Res. Soc. 51, 1070 - 1082

Dias, L.C. (2007). A note on the role of robustness analysis in decision aiding processes. Working paper. Institute of Systems Engineering and Computers INESC-Coimbra, Portugal. www.inesc.pt

Dubey, S. K., & Sharma, D. (2015). Software Quality Appraisal Using Multi-Criteria Decision Approach. International Journal of Information Engineering and Electronic Business, 7(2), 8.

Duranova, T., Raskob, W. and Schneider, T. (Eds.) (2016), Innovative integrated tools and platforms for radiological emergency preparedness and post-accident response in Europe. Key results of the PREPARE European research project, in special issue of Radioprotection, Volume 51

Durbach, I. N., & Stewart, T. J. (2012a). Modeling uncertainty in multi-criteria decision analysis. *European Journal of Operational Research*, *223*(1), 1-14.

Durbach, I. N., & Stewart, T. J. (2012b). A comparison of simplified value function approaches for treating uncertainty in multi-criteria decision analysis. Omega, 40(4), 456-464.

El Hanandeh, A., & El-Zein, A. (2010). The development and application of multi-criteria decision-making tool with consideration of uncertainty: The selection of a management strategy for the bio-degradable fraction in the municipal solid waste. Bioresource technology, 101(2), 555-561.

Feizizadeh, B., Jankowski, P., & Blaschke, T. (2014). A GIS based spatially-explicit sensitivity and uncertainty analysis approach for multi-criteria decision analysis. Computers & geosciences, 64, 81-95.

Fraedrich D. and Gover R. (2001). Statistical uncertainty analysis as a tool for model validation. In: Prado R. and Bolado R. (eds), SAMO 2001: Third international Symposium on Sensitivity Analysis of Model Output. CIEMAT, Madrid, Spain, 165–168.

French S (1988) Decision theory: an introduction to the mathematics of rationality. Halsted Press

French S (1996). Multi-attribute decision support in the event of a nuclear accident. J Multi-Crit Dec Anal 5: 39-57.

French S. (1998). Sensitivity analysis and uncertainty modelling in decision support systems. In: Chan K., Tarantola S. and Campolongo F. (1998) SAMO: Second International Symposium on Sensitivity Analysis in Model Output. EU Joint research Centre, European Commission, Luxembourg, 117–120.

French S., Haywood S., Oughton D., Smith J.Q., Turcanu C. (2017). The Various Meanings of Uncertainty. Report for the CONFIDENCE project. https://portal.iket.kit.edu/CONFIDENCE/index.php?action=confidence&title=objectives

French, S (2003) 'Modelling, making inferences and making decisions: the roles of sensitivity analysis' TOP, 11, 229-252.

Gamper, C. D., & Turcanu, C. (2015). Multi-criteria analysis: A tool for going beyond monetization. *The tools of policy formulation: Actors, capacities, venues and effects. Cheltenham (UK): Edward Elgar Publishing. p*, 121-141.

Gelderman J, Treitz M, Bertsch V and Rentz O (2005). Moderated Decision Support and Countermeasure Planning for Off-site Emergency Management. In: Loulou R, Waaub J-P and Zaccour G (eds). Energy and Environment: Modelling and Analysis. Springer, pp 63-80.

Giuliani, M., & Castelletti, A. (2016). Is robustness really robust? How different definitions of robustness impact decision-making under climate change. Climatic Change, 1–16.

Gluck, K.A., McNamara, J.M., Brighton, H.J., Dayan, P., Kareev, Y., Krause, J., Kurzban, R., Selten, R., Stevens, J.R., Voelkl, B. and Wimsatt, W.C. (2012). Robustness in a variable environment. In J.R.

Stevens, P. Hammerstein (Eds.) *Evolution and the mechanisms of decision making* (pp. 195-214). Strüngmann forum report, vol. 11,MIT Press. Cambridge, MA.

Groves, D.G. , Lempert R.J. (2007). A new analytic method for finding policy-relevant scenarios Global Environmental Change 17: 76-85.

Haasnoot, M., Kwakkel, J. H., Walker, W. E., & ter Maat, J. (2013). Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global environmental change*, *23*(2), 485-498.

Hadka, D., & Reed, P. (2013). Borg: An auto-adaptive many-objective evolutionary computing framework. *Evolutionary computation*, *21*(2), 231-259.

Hadka, D., Herman, J., Reed, P., & Keller, K. (2015). An open source framework for many-objective robust decision making. *Environmental Modelling & Software*, *74*, 114-129.

Hämäläinen RP, Sinkko K, Lindstedt MRK, Amman M and Salo A (2000a). Decision analysis interviews on protective actions in Finland supported by the RODOS system. STUK-A173. Radiation and Nuclear Safety Authority: Helsinki, Finland.

Hämäläinen RP, Lindstedt MRK and Sinkko K (2000b). Multi-Attribute Risk Analysis in Nuclear Emergency Management. Risk Analysis 20: 455-467.

Hämäläinen RP, Sinkko K, Lindstedt MRK, Amman M and Salo A (1998). RODOS and decision conferencing on early phase protective actions in Finland. STUK-A159, Radiation and Nuclear Safety Authority: Helsinki, Finland, ISBN 951-712-238-7.

Hanan, D., Burnley, S., & Cooke, D. (2013). A multi-criteria decision analysis assessment of waste paper management options. Waste management, 33(3), 566-573.

Hites, R., De Smet, Y., Risse, N., Salazar-Neumann, M., & Vincke, P. (2006). About the applicability of MCDA to some robustness problems. *European Journal of Operational Research*, *174*(1), 322-332.

Hurwicz, L. (1953). Optimality criterion for decision making under ignorance. In Uncertain. Expect. Econ. Essays Honour GLS Shackle.

Kassab, M., Hipel, K., & Hegazy, T. (2011). Multi-criteria decision analysis for infrastructure privatisation using conflict resolution. Structure and Infrastructure Engineering, 7(9), 661-671.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology*, *51*(1), 7-15. https://doi.org/10.1016/j.infsof.2008.09.009.

Korteling, B., Dessai, S., & Kapelan, Z. (2013). Using information-gap decision theory for water resources planning under severe uncertainty. *Water resources management*, *27*(4), 1149-1172.

Kouvelis P., Yu, G. (1997). Robust discrete optimization and its applications. Kluwer Academic Publishers.

Kwakkel J.H., Haasnoot M., Walker, W.E.. Comparing Robust Decision-Making and Dynamic Adaptive Policy Pathways for model-based decision support under deep uncertainty, Environmental Modelling & Software, Volume 86, 2016, Pages 168-183.

Lempert, R. J., Groves, D. G., Popper, S. W., & Bankes, S. C. (2006). A general, analytic method for generating robust strategies and narrative scenarios. Management science, 52(4), 514-528.

Lempert, R. J., & Collins, M. T. (2007). Managing the risk of uncertain threshold responses: Comparison of robust, optimum, and precautionary approaches. Risk Analysis, 27(4): 1009–1026.

Ligmann-Zielinska, A. and Jankowski, P. (2014). Spatially-explicit integrated uncertainty and sensitivity analysis of criteria weights in multicriteria land suitability evaluation. *Environmental Modelling & Software*, *57*, 235-247.

Mathieu, A., and Korsakissok, I. (2018), Guidelines describing atmospheric dispersion modelling uncertainties, CONCERT deliverable D9.1, [WWW Document]. URL http://www.concert-h2020.eu/en/Publications

McPhail, C., Maier, H.R., Kwakkel, J.H., Giuliani, M., Castelletti, A. and Westra, S. (2018). Robustness Metrics: How Are They Calculated, When Should They Be Used and Why Do They Give Different Results?. Earth's Future 6(2):.169-191.

Mendecka, B., Lombardi, L., and Kozioł, J. (2017). Probabilistic multi-criteria analysis for evaluation of biodiesel production technologies from used cooking oil. Renewable Energy.

Montibeller, G., & Franco, L. A. (2011). Raising the bar: strategic multi-criteria decision analysis. Journal of the Operational Research Society, 62(5), 855-867.

Montibeller, G., Gummer, H. and Tumidei, D. (2006). Combining scenario planning and multi-criteria decision analysis in practice. Journal of Multi-Criteria Decision Analysis, 14(1-3), 5-20.

NERIS, Updated Version of the NERIS SRA (2017), https://eu-neris.net/library/sra/156-updated-version-of-the-neris-sra.html

Norese, M. F. and Carbone, V. (2014). An application of ELECTRE Tri to support innovation. Journal of Multi-Criteria Decision Analysis, 21(1-2), 77-93.

Prado R. and Bolado R. (2001). SAMO 2001: Third international Symposium on Sensitivity Analysis of Model Output. CIEMAT, Spain.

Regan, H. M., Ben-Haim, Y., Langford, B., Wilson, W. G., Lundberg, P., Andelman, S. J. and Burgman, M. A. (2005). Robust decision-making under severe uncertainty for conservation management. Ecological Applications, 15(4), 1471-1477

Rios Insua, D. and French, S. (1991). A framework for sensitivity analysis in discrete multi-objective decision-making. *European journal of operational research*, *54*(2), 176-190.

Rosenhead, J., Elton, M., & Gupta, S. K. (1972). Robustness and optimality as criteria for strategic decisions. Operational Research Quarterly 23(4): 413-431.

Rosenhead, J. (2001). Robustness analysis: Keeping your options open. In J. Rosenhead & J. Mingers (Eds.), Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty, and Conflict. Chichester, UK: Wiley.

Roy, B. (2013). Robustness for Operations Research and Decision Aiding. Wiley Encyclopedia of Operations Research and Management Science.

Roy, B. and Boyssou, D., 1993. Aide Multicritère à la Décision: Méthodes et Cas. Economica, Paris, 695 p. (in French).

Savage, L. (1951.) The theory of statistical decision. J Am Stat Assoc 46(253):55–67.

Santos SM, Gaspar AT, Schiozer DJ. (2017). Risk management in petroleum development projects: Technical and economic indicators to define a robust production strategy. Journal of Petroleum Science and Engineering, 151, 116-27.

Scholten, L., Maurer, M., & Lienert, J. (2017). Comparing multi-criteria decision analysis and integrated assessment to support long-term water supply planning. PloS one, 12(5), e0176663.

Scholten, L., Scheidegger, A., Reichert, P., Mauer, M., & Lienert, J. (2014). Strategic rehabilitation planning of piped water networks using multi-criteria decision analysis. Water research, 49, 124-143.

Scholten, L., Schuwirth, N., Reichert, P., & Lienert, J. (2015). Tackling uncertainty in multi-criteria decision analysis–An application to water supply infrastructure planning. European Journal of Operational Research, 242(1), 243-260.

Schuwirth, N., Reichert, P., & Lienert, J. (2012). Methodological aspects of multi-criteria decision analysis for policy support: A case study on pharmaceutical removal from hospital wastewater. European Journal of Operational Research, 220(2), 472-483.

Stewart T.J. (2008). Robustness Analysis and MCDA. Newsletter of the European Working Group "Multiple Criteria Decision Aiding". Series 3, nº 18, Fall 2008

Sun, X., Gollnick, V., and Stumpf, E. (2011). Robustness Consideration in Multi-Criteria Decision Making to an Aircraft Selection Problem. Journal of Multi-Criteria Decision Analysis, 18(1-2), 55-64.

Tarun, P. K., Chen, V. C., Corley, H. W., & Jiang, F. (2011). Optimizing selection of technologies in a multiple stage, multiple objective wastewater treatment system. Journal of Multi-Criteria Decision Analysis, 18(1-2), 115-142.

Tom, R., Kapelan, Z. and Ralph, L. (2018). A resilience-based methodology for improved water resources adaptation planning under deep uncertainty with real world application. Water Resources Management, 32 (6), 2013-2031 " (2018).

Troldborg, M., Heslop, S., & Hough, R. L. (2014). Assessing the sustainability of renewable energy technologies using multi-criteria analysis: Suitability of approach for national-scale assessments and associated uncertainties. Renewable and Sustainable Energy Reviews, 39, 1173-1184.

Turcanu, C., Carlé B., Hardeman F. On the Constructive Role of Multi-Criteria Analysis in Nuclear Emergency Management. In: D.Ruan, F.Hardeman and K.van der Meer (Eds.), Intelligent Decision and Policy Making Support Systems, Springer-Verlag, Berlin, pp. 297-313, 2008.

Van Solingen, R. , Basili, V. , Caldiera, G. and Rombach, H. D. (2002). Goal Question Metric (GQM) Approach. In Encyclopedia of Software Engineering, J. J. Marciniak (Ed.).

Vincke, P. (1999). Robust solutions and methods in decision-aid. Journal of multicriteria decision analysis, 8(3), 181.

Vives i Batlle J., Aoyamab M., Bradshaw C., Brown J., Buesseler O ., Casacuberta N., Christl M., Duffa C., Impens N.R.E.N., Iosjpe M., Masqué P., Nishikawak J.. (2018), Marine radioecology after the

Fukushima Dai-ichi nuclear accident: Are we better positioned to understand the impact of radionuclides in marine ecosystems? Sci Total Environ. March 2018, Vol. 618:80-92

Wald, A. (1950). Statistical Decision Functions. Wiley, New York, NY

Walsh, M. M., Einstein, E. H., & Gluck, K. A. (2013). A quantification of robustness. *Journal of Applied Research in Memory and Cognition*, *2*(3), 137-148.

Walter H, Gering F, Arnold K, Gerich B, Heinrich G, Welte U., (2015), Simulation potentieller Unfallszenarien für den Notfallschutz in der Umgebung von Kernkraftwerken mit RODOS, Hrsg.: Bundesamt für Strahlenschutz, BfS-SCHR-55/14, 23.02.2015, urn:nbn:de:0221-2015021712440

Weaver, C. P., Lempert, R. J., Brown, C., Hall, J. A., Revell, D., & Sarewitz, D. (2013). Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks. *Wiley Interdisciplinary Reviews: Climate Change*, *4*(1), 39-60.

Zarghami, M., and Szidarovszky, F. (2009). Stochastic-fuzzy multi criteria decision making for robust water resources management. Stochastic Environmental Research and Risk Assessment, 23(3), 329-339.

Zhou, Q., Jiang, P., Huang, X., Zhang, F., & Zhou, T. (2018). A multi-objective robust optimization approach based on Gaussian process model. Structural and Multidisciplinary Optimization, *57*(1), 213-233.

Zucca, A., Sharifi, A. M., and Fabbri, A. G. (2008). Application of spatial multi-criteria analysis to site selection for a local park: A case study in the Bergamo Province, Italy. Journal of Environmental Management, 88(4), 752-769.