



**HAL**  
open science

## Dealing with uncertainty in risk analysis : combining safety and security

Houssein Abdo

► **To cite this version:**

Houssein Abdo. Dealing with uncertainty in risk analysis : combining safety and security. Automatic. Université Grenoble Alpes, 2017. English. NNT : 2017GREAT113 . tel-01829574

**HAL Id: tel-01829574**

**<https://theses.hal.science/tel-01829574>**

Submitted on 4 Jul 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

### **DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES**

Spécialité : **Automatique productive**

Arrêté ministériel : 25 mai 2016

Présentée par

**Houssein ABDO**

Thèse dirigée par **Jean-Marie FLAUS, Professeur, G-SCOP**, et  
codirigée par **François MASSE, Ingénieur, INERIS**

financée par l'**INERIS (Institut national de l'environnement  
industriel et des risques)**

préparée au sein du **Laboratoire G-SCOP**  
dans l'**École Doctorale EEATS**

## **Dealing with uncertainties in risk analysis: combining safety and cybersecurity**

Thèse soutenue publiquement le **12 Décembre 2017**,  
devant le jury composé de :

**M. Emmanuel GARBOLINO**

Maître de conférences, HDR, CRC, MINES ParisTech, Rapporteur

**M. Laurent PERRIN**

Professeur, Université de Lorraine ENSIC, Rapporteur

**M. Jean BIGEON**

Directeur de recherches, CNRS, Examineur

**M. Jean DEZERT**

Maître de recherches, Office National d'Etudes et de Recherches  
Aérospatiales, Examineur

**M. Youssef LAAROUCHI**

Ingénieur de recherches, Électricité de France - EDF, Examineur

**Mme. Maria DI MASCOLO**

Directrice de recherches, CNRS, Présidente

**M. François MASSE**

Ingénieur, INERIS, Encadrant

**M. Jean-Marie FLAUS**

Professeur, Université Grenoble Alpes, Directeur de thèse



Houssein ABDO : *Dealing with uncertainties in risk analysis : combining safety and cybersecurity* Thèse de doctorat, 12 Décembre 2017







*À ma famille,  
À mes amis,  
À mes rencontres de la vie, d'aujourd'hui et d'hier,  
À tous ceux qui m'aiment,  
À tous ceux que j'aime,*



---

## Acknowledgments/Remerciements

Cette thèse est l'aboutissement de plusieurs années de travail et n'aurait jamais pu voir le jour sans le soutien et l'aide de nombreuses personnes qui m'ont accompagnée tout au long de cette épopée! Je tiens donc ici à les remercier et leur témoigner de ma reconnaissance.

Je voudrais tout d'abord exprimer mes plus profonds remerciements à mon directeur de thèse, Jean-Marie FLAUS, pour toute sa disponibilité, sa gentillesse et ses directives si précieuses. Merci pour sa patience, ses nombreux conseils, ses idées, ses nombreuses relectures, et surtout pour m'avoir fait confiance dans cette aventure. Un grand merci pour cet encadrement de qualité intellectuellement et humainement.

Je tiens également à remercier chaleureusement à mon encadrant, François MASSE, pour ses idées et sa collaboration dans la réalisation de la thèse. Merci à L'INERIS pour financier cette thèse. Merci à tout les membres de l'INERIS pour ces collaborations, je remercie particulièrement Franck PRATS, Sylvain CHAUMETTE et Valérie DE DIANOUS.

Mes remerciements les plus sincères aux membres du jury : Maria Di-MASCOLO, qui m'a fait l'honneur de présider le jury ; Emmanuel GARBOLINO et Laurent PERRIN, qui ont accepté d'être rapporteurs de ce mémoire, merci pour leurs conseils et améliorations qu'ils ont suggérées, ainsi que Jean BIGEON, Jean DEZERT, Youssef LAAROUCHE d'avoir examiné ma thèse.

Je remercie aussi tous les enseignants et les chercheurs du laboratoire G-SCOP. Ces années à G-SCOP m'ont donné l'occasion de passer de l'autre côté du miroir en travaillant avec ceux qui étaient auparavant mes enseignants.

Je souhaite aussi saluer l'ensemble des ingénieurs, stagiaires et doctorants du laboratoire G-SCOP, sans qui ces quelques années auraient paru bien plus mornes. Je pense spécialement à Mohamad, François-Xavier, Tian, Kléber, Mohamad-Houssein, Abdullah,

---

Tamara, Ahmed, Aiman, Amine, Abed et Hadi.

Enfin, je remercie mes parents Ali et Mariam pour m'avoir soutenu durant toutes mes études, mes soeurs Khouloud et Zeina, mes frères Haydar et Hassan, qui ont largement contribué à l'aboutissement de ce projet de thèse.

Une page se tourne, une autre s'ouvre, merci à tous ceux et celles qui y ont contribué et m'ont amené à faire les bons choix.

---

# Contents

<b>List of Figures</b>	<b>xi</b>
<b>List of Tables</b>	<b>xv</b>
<b>Résumé</b>	<b>1</b>
<b>Abstract</b>	<b>3</b>
<b>1 Introduction</b>	<b>7</b>
1.1 Background . . . . .	8
1.2 Motivations and objectives . . . . .	10
1.3 Contributions and Methodology . . . . .	12
1.4 Thesis outline . . . . .	14
<b>2 Context &amp; problematic</b>	<b>17</b>
2.1 Introduction . . . . .	19
2.2 Understanding hazard analysis for critical facilities in France - The INERIS methodology . . . . .	20
2.2.1 Representation of risk scenarios . . . . .	21
2.2.2 Probability analysis using bow-tie based on the interval semi-quantitative approach . . . . .	23
2.2.2.1 characterizing input data . . . . .	24
2.2.2.2 propagating input data . . . . .	26
2.2.2.3 Final probability . . . . .	27
2.2.3 Severity analysis . . . . .	28
2.2.4 Decision-Risk matrix . . . . .	29

2.3	Introducing uncertainty in industrial risk analysis: the different sources of uncertainty . . . . .	31
2.3.1	Parameter uncertainty . . . . .	31
2.3.1.1	Types of parameter uncertainty . . . . .	32
2.3.1.2	Causes of parameter uncertainty . . . . .	32
2.3.2	Model uncertainty . . . . .	36
2.3.3	Completeness uncertainty . . . . .	36
2.4	Why uncertainty represents a major problem and should be considered? . .	37
2.4.1	Red-River flood - Why risk analysts should be explicit about uncertainty? . . . . .	37
2.4.2	The ASSURANCE benchmark project: Assessment of Uncertainties in Risk Analysis of Chemical Establishments . . . . .	39
2.5	Addressing uncertainty during risk assessment . . . . .	41
2.5.1	Steps to address parameter uncertainty . . . . .	41
2.5.2	Addressing completeness uncertainty . . . . .	41
2.6	Where uncertainties affect the INERIS risk assessment process . . . . .	42
2.6.1	Uncertainty in likelihood analysis: drawbacks of the interval semi-quantitative approach . . . . .	43
2.6.2	Uncertainty in effect analysis . . . . .	44
2.6.3	Completeness uncertainty in the risk identification process: No-considering of cyber-security threats . . . . .	47
2.7	Conclusion . . . . .	48
<b>3</b>	<b>Literature review</b>	<b>51</b>
3.1	State of art on approaches for parameter uncertainty analysis . . . . .	53
3.1.1	Interval analysis . . . . .	53
3.1.2	Probabilistic approach . . . . .	54
3.1.3	Fuzzy approach . . . . .	55
3.1.4	Evidence approach . . . . .	58
3.1.4.1	Independent input parameters . . . . .	59
3.1.4.2	Dependent input parameters . . . . .	60
3.1.4.3	Why evidence theory is the best to treat ignorance . . . .	62
3.2	Industrial Automation and Control System - IACS . . . . .	63
3.3	Cyber-security for industrial control systems . . . . .	64
3.4	The security risk analysis process for industrial control systems . . . . .	66
3.5	A review of cyber-security risk analysis approaches for Industrial Control Systems . . . . .	66
3.5.1	Attack-tree-based approaches . . . . .	67

3.5.1.1	Attack trees for assessing vulnerabilities in SCADA [28]	68
3.5.1.2	Through the Description of Attacks: a Multidimensional View [67]	69
3.5.1.3	Attack countermeasure trees (ACT): towards unifying the constructs of attack and defense trees [135]	69
3.5.2	Security Modeling with BDMP: From Theory to Implementation [126]	70
3.5.3	Evaluating the risk of cyber attacks on SCADA systems via Petri net analysis with application to hazardous liquid loading operations [80]	71
3.5.4	Network Vulnerability Assessment using Bayesian Networks [108]	72
3.5.5	Discussion	72
3.6	Existing Approaches that combine safety and security for industrial control systems	74
3.6.1	Integrating cyber attacks within fault trees [68]	74
3.6.2	Modeling safety and security inter-dependencies with BDMP (Boolean logic driven Markov processes) [125]	75
3.6.3	Studying interrelationships of safety and security for software assurance in cyber-physical systems: Approach based on Bayesian belief networks [98]	75
3.6.4	Modeling and analysis of safety-critical cyber physical systems using state/event fault trees [134]	76
3.6.5	Discussion	76
3.7	Conclusion	78
<b>4</b>	<b>Treatment of uncertainty in probability analysis: a fuzzy semi-quantitative approach</b>	<b>81</b>
4.1	Introduction	83
4.2	Proposed methodology: fuzzy semi-quantitative approach	84
4.2.1	Characterizing inputs for probability analysis using the fuzzy semi-quantitative approach	84
4.2.1.1	Define basic event frequencies using a fuzzy scale	84
4.2.1.2	Define secondary events probability of occurrence using a fuzzy scale	86
4.2.1.3	Define risk barriers confidence levels using a fuzzy scale	87
4.2.2	Representing (fuzzifying) input data based on the proposed fuzzy scales	87
4.2.3	Propagating fuzzy frequencies through the Bow-Tie	88



4.2.3.1	Treatment of OR gate . . . . .	89
4.2.3.2	Treatment of AND gate . . . . .	90
4.2.3.3	Treatment of secondary events . . . . .	90
4.2.3.4	Treatment of security barriers . . . . .	91
4.2.4	Decision making under fuzzy environment . . . . .	92
4.2.5	Handling the existing limitations of the interval approach . . . . .	93
4.3	Case study . . . . .	95
4.4	Conclusion . . . . .	97

**5 Treatment of aleatory and epistemic uncertainties in analyzing the effect of risks 101**

5.1	Introduction . . . . .	104
5.2	A new uncertainty analysis approach with randomness and fuzzy theory to deal with variability and imprecision . . . . .	105
5.2.1	Introduction . . . . .	105
5.2.2	Problem statement . . . . .	105
5.2.3	Proposed fuzzy-probabilistic approach . . . . .	107
5.2.3.1	Difference between probability and possibility . . . . .	107
5.2.3.2	Why a different new hybrid approach is needed . . . . .	108
5.2.4	Uncertainty representation . . . . .	110
5.2.5	Uncertainty propagation . . . . .	111
5.2.5.1	Monte Carlo to propagate Random Variables . . . . .	112
5.2.5.2	Monte Carlo to propagate uncertainty described by fuzzy numbers . . . . .	112
5.2.5.3	2D Monte Carlo to propagate Fuzzy Random Variables . . . . .	113
5.2.6	Case study - Application of the proposed hybrid approach to effect analysis with considering of parameter uncertainties . . . . .	116
5.2.7	Description . . . . .	116
5.2.7.1	Reactor model . . . . .	117
5.2.7.2	Computation of exposed area . . . . .	118
5.2.7.3	Uncertainty modeling . . . . .	119
5.2.7.4	Simulation conditions . . . . .	120
5.3	Comparison of the proposed hybrid approach with the existing approaches: applying uncertainty analysis to a Loss Of Containment scenario . . . . .	131
5.3.1	Models used in the analysis . . . . .	132
5.3.2	Calculation of the concentration without considering uncertainty . . . . .	133
5.3.3	Uncertainty modeling . . . . .	135
5.3.4	Uncertainty analysis using interval analysis . . . . .	136

5.3.5	Uncertainty analysis using the fuzzy approach . . . . .	136
5.3.6	Uncertainty analysis using the probabilistic approach . . . . .	139
5.3.7	Uncertainty analysis using Evidence theory . . . . .	140
5.3.8	Uncertainty analysis using the probabilistic-fuzzy approach . . . . .	142
5.3.9	Comparison of all approaches . . . . .	144
5.4	A global approach to treat all types and causes of uncertainty in effect analysis: the ALI-Aggregated Likelihood Index . . . . .	148
5.4.1	Uncertainty representation: characterizing each uncertain parameter based on the available information . . . . .	149
5.4.1.1	Evidence theory with fuzzy focal elements . . . . .	150
5.4.1.2	Construction of the Fuzzy belief and plausibility for a BOE with fuzzy focal elements . . . . .	151
5.4.2	Uncertainty propagation . . . . .	151
5.4.3	Decision Making under uncertain environment . . . . .	153
5.4.4	Case study - Applying of the developed approach to a loss of containment scenario . . . . .	154
5.4.4.1	Introducing fuzzy logic to analyze parameter uncertainty in calculating $\sigma_y$ and $\sigma_z$ . . . . .	154
5.4.4.2	Calculation of the concentration based on the proposed approach . . . . .	156
5.4.4.3	Comparing the proposed approach with the pure probabilistic approach . . . . .	157
5.5	Conclusion . . . . .	159
<b>6</b>	<b>Handling one aspect of completeness uncertainty: introducing cybersecurity within industrial risk analysis</b>	<b>171</b>
6.1	Introduction . . . . .	174
6.2	Global definition of industrial risk . . . . .	174
6.2.1	Definition of risks related to safety . . . . .	175
6.2.2	Definition of risks related to security . . . . .	175
6.2.3	Definition of risks related to safety and security . . . . .	175
6.3	Methodology for combined safety/security risk analysis . . . . .	176
6.3.1	Introduction behind the global idea . . . . .	176
6.3.2	Step-1: representation of a risk scenario . . . . .	178
6.3.2.1	Security risk analysis using a new extended Attack Tree . . . . .	178
6.3.2.2	Combined ATBT analysis . . . . .	181
6.3.3	Step-2: likelihood evaluation . . . . .	182
6.3.3.1	Determining minimal cut sets . . . . .	183

6.3.3.2	Characterizing likelihoods of occurrence of input events . .	184
6.3.3.3	Calculating the likelihoods of MCs . . . . .	186
6.4	Case study . . . . .	188
6.4.1	Description . . . . .	188
6.4.2	Application . . . . .	189
6.4.2.1	Step-1: Constructing ATBT for safety/security analysis . .	189
6.4.2.2	Step-2: Likelihood evaluation . . . . .	190
6.4.3	Discussion and improvement . . . . .	191
6.5	Modeling Stuxnet using the ATBT . . . . .	192
6.6	Limitations and future work . . . . .	195
6.7	Conclusion . . . . .	196
<b>Global conclusion and perspectives</b>		<b>207</b>
<b>Bibliography</b>		<b>213</b>
<b>Annexe A: Publications list</b>		<b>227</b>
H	Journal articles . . . . .	227
I	International conferences . . . . .	227
J	National conferences . . . . .	228
K	Under submission . . . . .	228

---

# List of Figures

1.1	Toulouse chemical factory explosion consequences. . . . .	8
1.2	Research overview and main questions. . . . .	11
1.3	The research methodology this work provides. . . . .	13
1.4	The most suitable theories to represent uncertain input parameters regarding the causes of parameter uncertainty that affect each parameter. . . . .	14
2.1	The hazard analysis process for high hazard critical facilities. . . . .	21
2.2	Elements of a “Bow-Tie” diagram. . . . .	23
2.3	Types of parameter uncertainty. . . . .	33
2.4	Uncertainty in using expert elicitation to characterizing likelihood of inputs. . . . .	34
2.5	Classification of parameter uncertainty. . . . .	35
2.6	Examples of known and unknown causes of the completeness uncertainty. . . . .	38
2.7	Grand Forks after levee over-topped. . . . .	39
2.8	Variation of results for the consequence assessment of the reference scenarios. . . . .	40
2.9	Steps to address parameter uncertainty during risk assessment. . . . .	42
2.10	Where uncertainties affect the risk assessment process. . . . .	43
2.11	Limitations presented in the traditional semi-quantitative approach. . . . .	44
2.12	Parameter uncertainty in effect analysis. . . . .	46
3.1	Monte Carlo propagation of probability distributions. . . . .	56
3.2	Triangular possibility distribution on the interval [4, 8], where the most possible value equals 6. . . . .	57
3.3	The triangular distributions of $X_1$ , $X_2$ and $Y = X_1 + X_2$ are shown in the top right, top left and bottom of the figure respectively. . . . .	58

3.4	Upper and lower CDFs of X, i.e. $Pl_X(-\infty, x)$ and $Bel_X(-\infty, x)$ in the top left; upper and lower CDFs of Y, i.e. $Pl_Y(-\infty, y)$ and $Bel_Y(-\infty, y)$ in the top right; upper and lower CDFs of $Z = X + Y$ , i.e. $Pl_Z(-\infty, z)$ and $Bel_X(-\infty, z)$ in the bottom middle. . . . .	61
3.5	Propagation of uncertainty represented in terms of bodies of evidences when input parameters are considered to be independent (left) and dependent (right). . . . .	61
3.6	Body of evidence for the output $Z = X + Y$ when total uncertainty between X and Y is considered. . . . .	62
3.7	Components and architecture of IAS. . . . .	64
3.8	Attack tree structure using an example as developed by [140]. . . . .	67
3.9	Attack tree example for the MODBUS-based SCADA system given in [28].	68
3.10	Example of the attack tree proposed in [67]. . . . .	70
3.11	ATC with one attack and multiple pairs of detection and mitigation events [135]. . . . .	71
3.12	A Simple Example of Bayesian Attack Graph [108]. . . . .	72
4.1	Framework for estimating the probability of accidents in Bow-Tie analysis based on the fuzzy semi-quantitative approach. . . . .	85
4.2	Mapping event frequencies on fuzzy scale. . . . .	86
4.3	Fuzzy frequency class $FX$ . . . . .	87
4.4	Mapping probability of secondary events on fuzzy scale. . . . .	87
4.5	Mapping confidence levels on fuzzy scale. . . . .	88
4.6	OR gate example. . . . .	89
4.7	AND gate example. . . . .	90
4.8	Consideration of an ES within the fuzzy semi quantitative approach. . .	91
4.9	Consideration of a security barrier within the fuzzy semi quantitative approach. . . . .	92
4.10	Decision making under fuzzy environment. . . . .	93
4.11	How the fuzzy semi-quantitative approach deals with the limitations mentioned in Section 2.6.1. . . . .	94
4.12	A deviation will not lead to different result that affects the decision. . . .	94
4.13	The bow-tie diagram of the LOC scenario under study. . . . .	99
5.1	Uncertainty presented in the analysis and its effects. . . . .	106
5.2	Two bottles of liquid with their uncertainty representations. . . . .	108
5.3	Addition of A and B with the consideration of uncertainty. . . . .	109
5.4	Representation of a fuzzy random variable. . . . .	111

---

5.5	Representation and propagation of uncertainty. . . . .	112
5.6	Use of Monte Carlo to estimate the output probability distribution. . . . .	112
5.7	Use of Monte Carlo to estimate the fuzzy result. . . . .	113
5.8	Use of a 2D Monte Carlo simulation to handle aleatory, epistemic and mixed uncertainties in risk assessment for dynamic systems. . . . .	115
5.9	Modeling simulations of the system using fuzzy input. . . . .	116
5.10	Polymerization reactor with its cooling unit and bursting disk. . . . .	117
5.11	Simulation under normal conditions with normal numbers. . . . .	122
5.12	Simulation under abnormal conditions with normal numbers. . . . .	123
5.13	CDF of mass released after consideration of aleatory uncertainty. . . . .	124
5.14	Masses released after the simulations. . . . .	126
5.15	CDF of results after taking into consideration all uncertainty types. . . . .	128
5.16	Triangular fuzzy number of the response time. . . . .	128
5.17	Estimation of the fuzzy mass released for $TTF = 752.5$ and different reponse times. . . . .	129
5.18	Loss of containment scenario . . . . .	131
5.19	Uncertain and precise input parameters used to calculate the concentration at a specific end-point . . . . .	135
5.20	Fuzzy concentration at grid point $C(500, 0, 0)$ . . . . .	137
5.21	Cumulative necessity function (CNF), and cumulative possibility function (CPoF) at grid point $C(500, 0, 0)$ . . . . .	138
5.22	How distribution shapes can contribute to decision making. . . . .	139
5.23	The pdf for the concentration at grid point $C(500, 0, 0)$ using the Monte Carlo simulation . . . . .	141
5.24	The CDF for the concentration at grid point $C(500, 0, 0)$ . . . . .	142
5.25	The CBF and CPF for the concentration at grid point $C(500, 0, 0)$ . . . . .	143
5.26	Difference between the fuzzy approach and the evidence approach based on total dependence . . . . .	144
5.27	Lower, most likely and upper distributions for the concentration at grid point $C(500, 0, 0)$ . . . . .	145
5.28	Comparison of the results generated from the different approaches. . . . .	146
5.29	Summary of the characteristics of the approaches highlighted in this comparison . . . . .	147
5.30	The ALI framework to deal with uncertainties . . . . .	162
5.31	Flow chart for representing parameter uncertainty based on the available information . . . . .	163
5.32	Aggregating multiple sources of data in the same formalism. . . . .	164

5.33 Representing and propagating of uncertainty within the ALI methodology 165

5.34 Decision tree for selecting the more appropriate uncertainty propagation technique regarding the risk model and uncertain input data . . . . . 166

5.35 Aggregating the likelihood in a single likelihood index . . . . . 167

5.36 Uncertain and precise input parameters used to calculate the concentration at a specific end-point . . . . . 167

5.37 The proposed fuzzy scales for the wind speed and the day time isolation 168

5.38 The ALI distribution for the concentration at grid point  $C(500, 0, 0)$  using the 2-stages MC simulation . . . . . 168

5.39 The cumulative ALI distribution vs the cumulative pure probabilistic distribution for the concentration at grid point  $C(500, 0, 0)$  . . . . . 169

6.1 Global definition of risk. . . . . 177

6.2 Framework of the proposed approach for safety/security risk analysis. . . 179

6.3 How attackers exploit system vulnerabilities in order to cause damages. . 181

6.4 Example of the structure of the proposed attack tree. . . . . 182

6.5 The different form of a security basic event. . . . . 182

6.6 Modeling WannaCry ransomware attack using the proposed AT. . . . . 183

6.7 Example of how we attach an AT to its corresponding event in BT. . . . 183

6.8 Characterizing the likelihood of security basic events. . . . . 198

6.9 Example of how calculating the likelihood of an MC. . . . . 199

6.10 The chemical reactor with its control system structure. . . . . 199

6.11 Combined ATBT of the scenario under study. . . . . 200

6.12 AT for the goal: gain unauthorized access to SCADA. . . . . 201

6.13 Calculating the likelihood of MC number 19. . . . . 202

6.14 The top objective of Stuxnet. . . . . 203

6.15 Attack tree of the “spreading of Stuxnet”. . . . . 204

6.16 Attack tree of “Stuxnet self installation”. . . . . 205

6.17 A global vision of the thesis structure. . . . . 207

---

# List of Tables

2.1	Abbreviations, significations and definitions of elements listed in the Bow-Tie diagram. . . . .	24
2.2	Determining the frequency classes based on the semi-quantitative approach.	25
2.3	Representative values of the probability of failure of ESs. . . . .	25
2.4	Confidence Level: probability of failure on demand. . . . .	26
2.5	Probability of occurrence scale extracted from the French ministerial order of 29/09/2005 related to the evaluation of risk. . . . .	27
2.6	Transforming of frequency classes into probability levels. . . . .	27
2.7	French end-point values for the intensity of thermal radiation, toxic and over-pressure effects. . . . .	28
2.8	French scale for the classification of the severity of a potential accident. .	29
2.9	French risk matrix. . . . .	30
2.10	Deviation between frequencies of the top events of the common scenarios analyzed by the partners (events per year). . . . .	40
2.11	Difference between safety and security. . . . .	48
3.1	Body of evidence for the output $Z = X + Y$ obtained by means of the Cartesian product. . . . .	60
3.2	Description of events used for representing an attack scenario. . . . .	69
3.3	List of approaches combining safety and security for ICS. . . . .	73
4.1	Estimated results obtained using quantitative, interval semi-quantitative and fuzzy semi quantitative approaches. . . . .	97
5.1	Model equations and output variable description for the system. . . . .	119
5.2	Lamda using the TM5 1300 abacus. . . . .	120



5.3	Fuzzy numbers for uncertain parameters. . . . .	120
5.4	Fuzzy distance and fuzzy mass released for the simulations. . . . .	125
5.5	Results for fixed fuzzy time response (4.5, 6.2, 10.3) . . . . .	129
5.6	Result at failure time(752.5) . . . . .	130
5.7	Result at time of failure equal to 752.5. . . . .	130
5.8	Pasquill stability classes . . . . .	133
5.9	Briggs' Fitting Constants for $\sigma_y$ and $\sigma_z$ . . . . .	134
5.10	Input values used to calculate the concentration at a grid point without considering uncertainty . . . . .	134
5.11	Types of uncertainty and sources of information about the uncertain parameters . . . . .	135
5.12	Uncertain variables represented as intervals . . . . .	136
5.13	Uncertain variables specified as triangular fuzzy numbers . . . . .	137
5.14	Uncertain variables represented using probability distributions. . . . .	140
5.15	Uncertain variables specified as spaces of evidence . . . . .	141
5.16	Uncertain variables represented using probability distributions . . . . .	143
5.17	90% confidence intervals for each of the approaches used. . . . .	146
5.18	Determining the stability class based on the proposed fuzzy scale . . . . .	155
5.19	Example of determining the stability classes based on the fuzzy scheme . . . . .	156
5.20	Uncertain variables represented using either probability distributions, fuzzy numbers, FRV and body of evidence . . . . .	157
5.21	Uncertain variables represented using probability distributions, Fuzzy numbers and FRVs . . . . .	158
5.22	Uncertain variables represented using probability distributions . . . . .	158
5.23	90% confidence intervals for the ALI-approach vs the pure probabilistic approach. . . . .	159
6.1	Description of events used for representing an attack scenario. . . . .	180
6.2	Qualitative scale to characterize the frequency of input safety events. . . . .	185
6.3	Qualitative scale to characterize the likelihood of input security events. . . . .	186
6.4	Analysis scale - Overall likelihood. . . . .	187
6.5	The identified MCs for the scenario under study. . . . .	192
6.6	The re-identified MCs after the added improvement. . . . .	193
6.7	The identified MCs for the Stuxnet attack scenario. . . . .	195

---

# Résumé

L'analyse des risques est un élément essentiel pour la prise de décision réglementaire liée aux industries à haut risques. Une analyse systématique des risques se compose de trois étapes: (i) l'identification des scénarios indésirables de risque. (ii) l'estimation de la probabilité d'occurrence des scénarios des risques. (iii) le calcul d'effet des conséquences des scénarios de risque identifiés. L'analyse de la vraisemblance et de la gravité s'effectue à l'aide de modèles qui dépendent de plusieurs paramètres d'entrée.

Cependant, la fiabilité de l'analyse de risque est limitée grâce à diverses sources d'incertitude. L'incertitude des paramètres, du modèle et d'incomplétude sont les principales sources d'incertitude. L'incertitude de paramètres découle de l'incapacité de définir des valeurs exactes à certains paramètres d'entrée utilisés pour l'analyse de la probabilité et de l'effet. L'incertitude de l'incomplétude provient de ne pas tenir compte de l'ensemble des contributions au risque dans le processus d'identification (certains événements initiateurs sont ignorés). L'incertitude du modèle n'est pas prise en compte dans ce travail.

L'INERIS (Institut national de l'environnement industriel et des risques) a développé une approche semi-quantitative d'intervalle pour l'évaluation de la probabilité des risques qui utilise des informations quantitatives si disponibles ou des informations qualitatives, sinon. Cependant, cette approche semi-quantitative d'intervalle présente certains inconvénients en raison de l'incertitude des paramètres.

L'information concernant les paramètres d'entrée des modèles d'effets est souvent incomplète, vague, imprécise ou subjective. En outre, certains paramètres peuvent être de nature aléatoire et ont des valeurs différentes. Cela conduit à deux différents types d'incertitude des paramètres. L'incertitude aléatoire dû à la variabilité naturelle. L'autre est l'incertitude épistémique, causée par le manque d'informations, par exemple, une imprecision de mesure.

---

De plus, dans les méthodes d'analyse de risque actuelles, l'étape d'identification est incomplète. Juste les scénarios liés à la sûreté causés par des événements accidentels sont pris en compte durant l'analyse. L'introduction de systèmes connectés et de technologies numériques dans l'Industrie crée de nouvelles menaces de cyber-sécurité qui peuvent entraîner des accidents de sûreté indésirables. Ces événements liés à la cyber-sécurité doivent être pris en compte lors de l'analyse des risques industriels.

Cette recherche vise à développer des méthodologies d'analyse d'incertitude pour traiter l'incertitude dans le processus d'analyse de risque de l'INERIS. En d'autres termes, analyser l'incertitude dans l'analyse de la probabilité, l'analyse des effets et l'étape d'identification. Dans ce travail, nous traitons les limites de l'approche semi-quantitative d'intervalle en introduisant la notion de nombres flous au lieu d'intervalles. Les nombres flous sont utilisés pour traiter l'incertitude dans les données d'entrée.

Une méthodologie hybride qui traite chaque cause de l'incertitude des paramètres dans l'analyse des effets avec la bonne théorie est développée. La théorie de la probabilité est utilisée pour représenter la variabilité, les nombres flous sont utilisés pour représenter l'imprécision et la théorie d'évidence est utilisée pour représenter l'ignorance, l'incomplétude ou le manque de consensus.

Une nouvelle méthodologie d'identification des risques qui considère la sûreté et la sécurité ensemble lors de l'analyse des risques industriels est développée. Cette approche combine Noeud-Papillon (BT), utilisé pour l'analyse de sûreté, avec une nouvelle version étendue de l'arbre d'attaque (AT), introduite pour l'analyse de cybersécurité des systèmes de contrôle industriel. L'utilisation combinée d'AT-BT fournit une représentation exhaustive des scénarios de risque en termes de sûreté et de sécurité.

---

# Abstract

The French ministerial order of 29/09/2005 imposes the assessment of risks of critical facilities with toxic, flammable, explosive or mixtures substances to prevent the occurrence of undesirable accidents and protect the surroundings people and the environment. The outcomes of risk calculations are used for permit granting and for land-use planning. A systematic risk analysis is made up of three steps: (i) identifying the undesirable risk scenarios that can lead to major accidents. A risk scenario is characterized by referencing to the potential event with its causes and consequences. (ii) Estimating the likelihood of occurrence of risk scenarios. (iii) Calculating the severity of consequences of the identified risk scenarios. Likelihood and severity analysis are carried out with the help of models that depend on several number of input parameters.

However, the trustworthiness of risk analysis is limited when inaccuracies in the results can occur, and are due to various sources of uncertainty. Parameter, model and completeness uncertainties are the main sources of uncertainty that affect an assessment. Parameter Uncertainty arises from the inability to set exact values for certain input parameters used for likelihood and severity analysis. Model uncertainty stems from the fact that risk models used for severity analysis are representations of reality and based on simple mathematical equations. Completeness uncertainty originates from not considering all contributions to risk in the identification process (some initiating events are ignored). In this study parameter and completeness uncertainties are addressed. Model uncertainty is difficult to quantify and it can be mitigated by validating the risk models against experiments.

Likelihood analysis can be qualitative or quantitative depending on the types of inputs data. Input data for likelihood analysis are either derived from databases (based on the data collected at the time of on-site investigations or similar facilities) or expert judgments if the former is not available. Qualitative analysis is subjective and not precise, while

---

quantitative analysis is often too expensive to perform. For these reasons, the INERIS has developed an interval semi-quantitative approach that uses both quantitative information if available or qualitative information if not. However, this interval semi-quantitative approach has some drawbacks due to parameter uncertainty and in some cases can lead to likelihood underestimation.

Models used for severity analysis are complex and depend on a large number of input parameters. However, Inability in determining precise values for models' input parameters may be faced due to time and financial constraints. Information regarding model parameters is often incomplete, vague, imprecise or subjective. Moreover, some of the parameters may be random in nature and have different values. This leads to two different types of parameter uncertainty that need to be accounted for an accurate risk analysis and effective decision-making. Aleatoric uncertainty arises from randomness due to natural variability resulting from the variation of a value in time. Or epistemic uncertainty caused by the lack of information resulting, for example, from measurement errors, subjectivity expert judgment or incompleteness.

Moreover, in today's risk analysis methodologies, the identification step is incomplete where only safety related scenarios caused by accidental events (component failures, human errors, etc.) are considered. The introduction of connected systems and digital technology in process industries creates new cyber-security threats that can lead to undesirable safety accidents. These cyber-security related events should be considered during industrial risk analysis. Safety and security are assessed separately when they should not be. This is because a security threat can lead to the same dangerous phenomenon as a safety incident. Thus, a new risk identification methodology that deal with completeness uncertainty by considering safety and security together during risk analysis is an important need.

This research aims to develop uncertainty analysis methodologies to evaluate industrial risks for critical facilities under parameter uncertainty. And proposing a modeling tool to complete the risk analysis process by introducing cyber-security related risks into the identification step. The research has the following specific objectives:

- To deal with parameter uncertainty during risk analysis, and provide a guidance for risk assessors and decision makers on how the best to handle parameter uncertainty and taking decisions in an uncertain universe.
  - ✓ To develop a methodology for likelihood analysis, update the interval semi-quantitative in order to remove the existing drawbacks, and to deal with parameter uncertainty that affect the input data;
  - ✓ To develop an uncertainty analysis methodology that separately handle aleatory and epistemic uncertainty in effect analysis using the best suitable represen-

---

tation and propagation theories regarding the causes of uncertainty and based on the only available information;

- To develop a risk identification methodology and risk modeling technique to introduce cyber-security related threats in industrial risk analysis for an exhaustive representation of risk scenarios;
- To demonstrate the applicability and effectiveness of the developed methodologies and the utility of the tool by applying them on real case studies.

In this work, we propose a fuzzy semi-quantitative approach to deal with parameter uncertainty in the likelihood analysis step. We handle the limits of the interval semi-quantitative approach by introducing the concept of fuzzy numbers instead of intervals. Fuzzy numbers are used to represent subjectivity in expert judgments and covers uncertainty in the quantitative data if this data exists. This proposed fuzzy-based methodology contributes to a simpler and effective alternative to the quantitative approach and more precise to the qualitative approach while keeping the virtue of being based on real accident frequency data if presented, and with the consideration of uncertainty.

A hybrid methodology to treat the two types of uncertainty separately in severity analysis is proposed. Probability theory is used to represent variability, fuzzy numbers are used to represent imprecision and evidence theory is used to represent vagueness, incompleteness and the lack of consensus. The represented parameters are propagated using MC simulation and/or fuzzy and evidence calculus depending on the type of the risk model. A comparison between the proposed methodology and the existing ones is performed to validate the effectiveness and preciseness of the method. Then guidelines for choosing the best approach to represent uncertain parameters based on the only available information is provided. And we proved that the use of an inappropriate approach in an inappropriate place may lead to under or overestimation of risk and subsequently to a bad decision.

A new risk identification methodology that considers safety and security together during industrial risk analysis is developed. This approach combines Bow-Tie Analysis (BTA), commonly used for safety analysis, with a new extended version of Attack Tree Analysis (ATA), introduced for security analysis of industrial control systems. The combined use of AT-BT provides an exhaustive representation of risk scenarios in terms of safety and security.

This research develops an approach for evaluating the likelihood level of safety/security risk scenarios based on two-term likelihood parts, one for safety and one for security. Two-term likelihood parts are used because safety and security events are different in nature. This differentiation helps in identifying the sequences of events (minimal cut sets) that are purely related to safety, security or to both. The resulting output of different types

---

of cut sets offers richer information for decision making.

# 1

## Introduction

**Summary:** This chapter presents the overall context of the thesis. In the first place, we present the motivations and objectives of conducting this work. Secondly, the main contributions of the thesis are highlighted. The chapter ends by outlining the structure of the thesis.

### Summary

---

<b>1.1</b>	<b>Background . . . . .</b>	<b>8</b>
<b>1.2</b>	<b>Motivations and objectives . . . . .</b>	<b>10</b>
<b>1.3</b>	<b>Contributions and Methodology . . . . .</b>	<b>12</b>
<b>1.4</b>	<b>Thesis outline . . . . .</b>	<b>14</b>

---



## 1.1 Background

Disaster and major industrial accidents (explosion, dispersion, etc.) in critical facilities classified SEVESO pose a significant threat to humans and the environment. Managing risks linked to these facilities is of crucial importance to minimize and prevent the associated hazards by implementing the right measures to ensure appropriate preparedness and that risks are managed according to defined acceptance criteria.

In 1990, the French Government established the INERIS institute as the National competence center for Industrial Safety and Environmental Protection. INERIS has developed expertise in the areas of chronic and hazardous risks. After the major accident of Toulouse and the related consequences as presented in Figure 1.1, the French ministerial laws paid more attention on industrial major risks and encourage the use of probabilistic analysis in all regulatory matters. INERIS and based on its expertise has developed a systematic probabilistic analysis methodology to analyze risks in order to protect people and the environment from major accidents. This methodology is made up of three steps:

1. risk identification: explore how an undesirable hazard can be developed starting from causes and ending with the consequences;
2. likelihood analysis: estimate the likelihood of identified risk scenarios;
3. severity analysis: calculate the impact of identified risk scenarios on surrounding environments in terms of people.



**Figure 1.1** – Toulouse chemical factory explosion consequences.

However, due to the uncertainty involved, the credibility of risk analysis results is still a major issue. Addressing uncertainty during risk analysis has become an important part

for a health risk analysis. International regulatory guides and risk management standards recognize the importance of the identification and treatment of uncertainty that are part of the risk analysis. The international standardization ISO-31000 defines risk as the effect of uncertainty on objectives. Regulatory commissions further note in their policy statements that the "treatment of uncertainty is an important issue for effective decision-makings". These references provide guidance on this subject to varying degrees. However, they do not provide explicit methodologies on the treatment of uncertainty regarding the sources and causes of this uncertainty. Most risk analysis studies struggle with how addressing uncertainty and yet uncertainty is not systematically treated. An important aspect in obtaining meaningful risk analysis is knowing what are the sources of uncertainty, the causes of these sources and the impact of these uncertainties on the analysis predictions. Capturing uncertainty regarding its sources and causes is vital in order to perform a sound uncertainty assessment.

Uncertainty in risk analysis has different sources, it can be either parameter, model or completeness. Parameter uncertainty is generated from the inability of giving precise values to some input parameters. Parameter uncertainty can be aleatoric caused by natural variability (randomness associated with the parameters of the model), or epistemic due to lack of information caused from imprecision, subjectivity, etc. The impact of parameter uncertainty is gained through quantification. Model uncertainty is due to assumptions and simplifications made during building risk models. Completeness uncertainty is caused from omitting some risk contributors intentionally or not. The causes of completeness uncertainty are characterized in terms of how they affect the assessment (e.g., introduction of a new type of initiating events, changes in the likelihood representation of events due to the introduction of initiating events, etc.).

In risk analysis, confidence in the information used for likelihood and severity analyses is an important issue. Uncertainty can occur in the input parameters of the likelihood or effect mathematical models (the frequency of occurrence value of an initiating event for example). This parameter uncertainty can result in risk under or overestimation and then to an inappropriate decision. Consequently, most of today's risk analysis only consider safety related risk causes generated from accidental component failures and human errors. However, introducing technology and connected systems into critical facilities has generated new type of risk causes that are related to cyber-security. Security threats that may lead to major accidents are not yet considered during industrial risk management. This result in completeness uncertainty that should be addressed (cyber-security related risk are omitted). Thus, initiation of work to modify the analysis to be up-to-date and meet the today's safety needs is necessary. For these reasons, this work aims to analyze:

- parameter uncertainty in quantifying the likelihood and severity of the identified

risk scenarios;

- the effect of neglecting or not considering exhaustively all kind of risk scenarios.

The methodology for treatment of uncertainties in this study is intended to provide a reasonable process to support the decision making. It should be noted that the treatment of uncertainty methodology contained in this report can be followed to treat uncertainty in any field.

The following chapter will give a background to the motivation behind the research, as well as the gap that presented the research opportunity. It then highlights the key contributions, and follows up with an overview of the thesis outline.

## 1.2 Motivations and objectives

The output of risk analysis used for decision making may be inaccurate due to parameter uncertainty that affect the inputs used for conducting the analysis. Assess parameter uncertainty is inevitable to conduct a valuable risk analysis and make the right decision. Addressing parameter uncertainty is achieved by: (i) representing the uncertain input data, and (ii) propagating these uncertain data after being represented through the risk model to obtain representations of uncertainty for the outputs.

However, existing risk analysis frameworks developed for decision making do not address parameter uncertainty correctly in a proper way. Existing approaches to address parameter uncertainty are either probabilistic, no probabilistic or a mix of both. As we mentioned before, parameter uncertainty can be aleatory or epistemic. The literature review on uncertainty modeling acknowledge a hybrid or mix (probabilistic non-probabilistic) approach to address these two types of parameter uncertainty separately in a single framework. However, several hybrid approaches have been proposed by several authors in the literature, but these mixed approaches suffer from some limitations regarding the mathematical theories used for representing of available data as well as the propagation of these data after being represented. In addition, there is no guideline to select a proper parameter uncertainty analysis techniques according to the types and causes of parameter uncertainty, and in respecting to models used in the analysis for propagation of represented uncertain data.

On the other hand, today critical facilities replace mechanical devices and closed systems by digital devices and interconnect systems. The introduction of technology and connected systems by embedding sensing, computing, and communication into critical facilities creates new and challenging threats to safety that are related to cyber-security. These threats may lead to major accidents that affect human lives and the environment.

Yet, cyber-security related threats are not considered during industrial risk analysis.

Safety and security are assessed separately when they should not be. Security experts interest only on CIA (confidentiality, integrity and availability) when analyzing security for critical facilities. Likewise, many risk assessors, do not automatically think about how a hacker might exploit plant designs to harm people or cause physical damage (so-called adversarial thinking). For these reasons, risk analysis for critical facilities requires expansive thinking about security that affect safety. Research that bridges this divide for a complete risk analysis is necessary as critical facility becomes increasingly dependent on cyber-control systems. Thus, finding ways how safety and security measures can be leveraged together to improve the overall security and safety of such systems and facilities would be valuable.

This thesis is motivated by the need to study the unreliable treatment of parameter uncertainty on the risk analysis results, and the impact of no-considering cyber-security. Diagram 1.2 shows how this research approached key questions regarding quantifying parameter uncertainty and not considering cyber-security during industrial risk analysis.

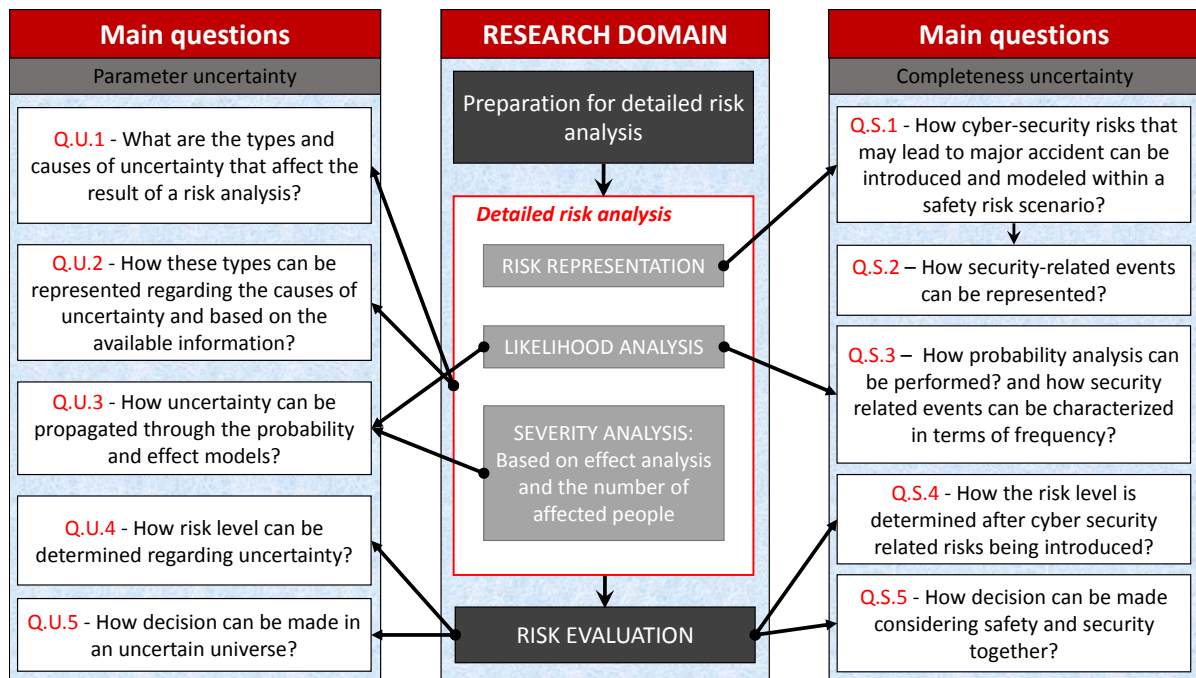


Figure 1.2 – Research overview and main questions.

The objective of this thesis is to provide technical guidance for establishing the level of confidence that can be placed in a decision based on the analysis of risk by: (i) providing a guideline on modeling and quantifying data uncertainty to help risk assessors choose the most accurate and suitable representation of the available data, (i) developing a risk evaluation method that can analyze and demonstrate causal relationships in high safety/security risk scenarios. Achieving these goals will provide an exhaustive and global industrial safety/security risk analysis framework that consider uncertainties for more

effective decision-making.

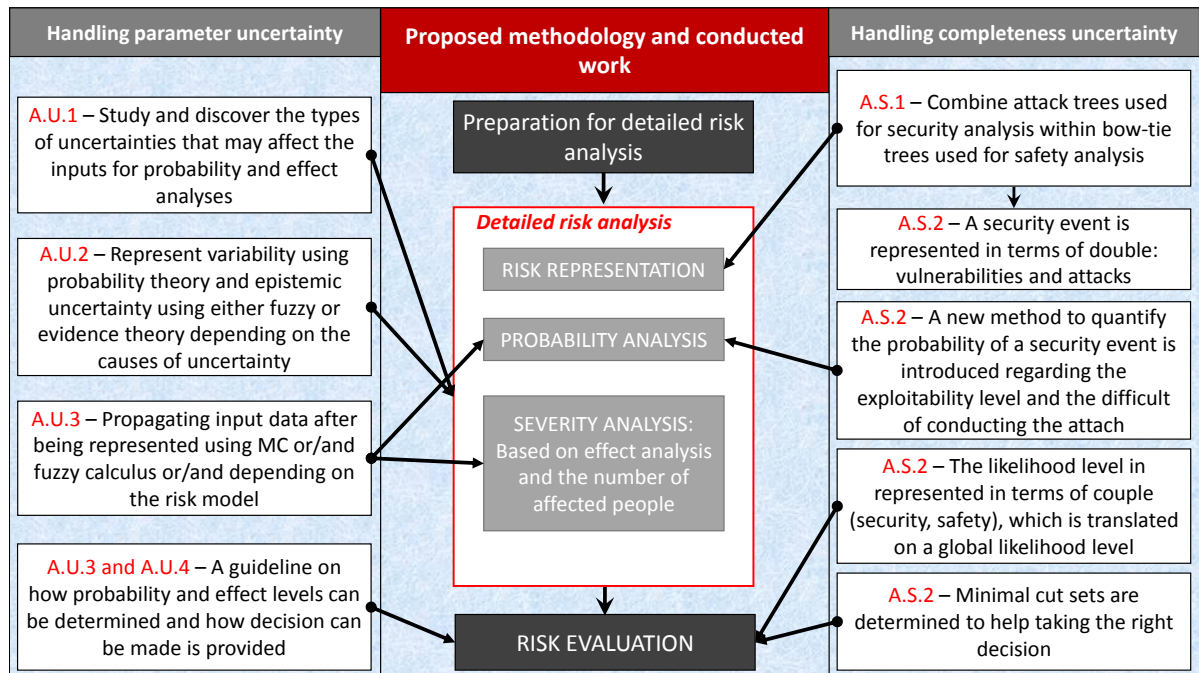
The specific aims of this research include:

- assessing data or parameter uncertainty during likelihood and effect analyses
  - ✓ examine the causes of parameter uncertainty in both likelihood and severity analyses;
  - ✓ review the mathematical representations of uncertain knowledge presented in the literature;
  - ✓ propose robust approaches to handle parameter uncertainty in likelihood and effect analyses by providing the most suitable representation to each cause of parameter uncertainty.
- Handling completeness uncertainty by considering cyber-security during risk analysis
  - ✓ propose a new global definition of risk that covers safety and security;
  - ✓ develop an identification and representation method of a safety/security risk scenario;
  - ✓ develop a likelihood analysis methodology to quantify the likelihood a safety/security risk scenario with the consideration of the difference in frequency of occurrence between safety and security related events;
  - ✓ provide a global safety/security likelihood levels to be used for decision making.
- provides guidance on how to treat uncertainties associated with industrial risk analysis used for decision making. Perform uncertainty analyses on real case studies to understand the impact of the uncertainties on the risk analysis results, and to provide examples that help risk assessors in analyzing risks under uncertain environment.

### 1.3 Contributions and Methodology

In this work, a global industrial risk analysis methodology that address parameter and completeness uncertainties is developed. Diagram 1.3 shows the proposed hierarchy to answer the key research questions and achieve the objectives. The diagram shows at a macro level how parameter uncertainty is treated, and how cyber-security related risks are introduced (completeness uncertainty). Taken together, the proposed methodology allow us to answer the oriented questions posed in the previous section as depicted in Diagram 1.3.

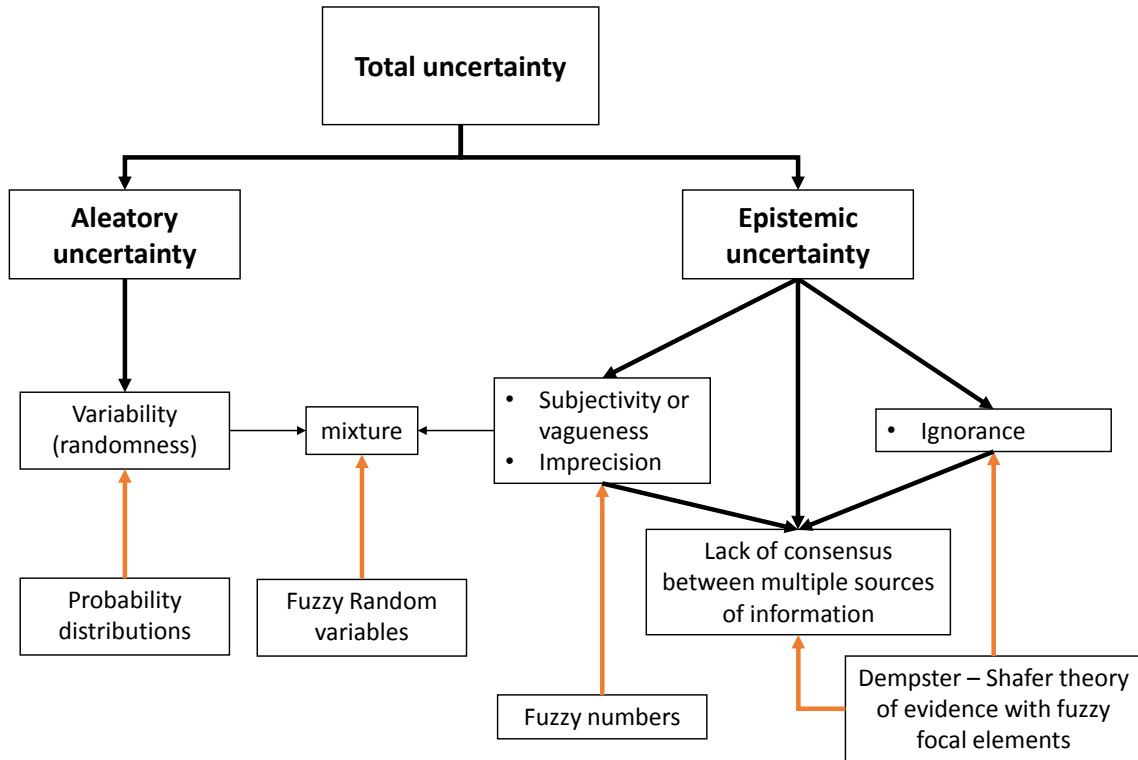
To provide a sound parameter uncertainty assessment, in this study, the most commonly used approaches to deal with parameter uncertainty: interval analysis, fuzzy theory,



**Figure 1.3** – The research methodology this work provides.

probability theory, evidence theory, and the mixed probabilistic-fuzzy approach are reviewed. The drawbacks of these approaches in dealing with certain causes of parameter uncertainty are identified. Then, an uncertainty analysis guidance is concluded to help risk assessors in determining the best representation method for the input parameters regarding the type of available information (cause of parameter uncertainty) as presented in Figure 1.4. Based on this guideline, new uncertainty analysis methodology is developed. In this methodology, randomness (probability theory), possibility (fuzzy numbers) and belief (evidence theory) are combined to treat the different causes of parameter uncertainty separately with the right mathematical theory. Coupled with completeness uncertainty, a new definition of risk is introduced where a risk scenario now covers safety and security sequences of events. The bow-tie analysis used to model safety related risks scenarios is combined with an extended version of the attack tree analysis that is introduced for security analysis to result in a cyber-bow-tie analysis. In sum, we are able to logically support answers to the following key questions:

1. How parameter uncertainty can be best addressed and based on the available information?
2. How safety and security related risk that can lead to major accidents can be analyzed together?
3. How should risk management efforts be allocated by implementing safety and security measures that work together so as to more effectively mitigate the probability and impacts of accidents?



**Figure 1.4** – The most suitable theories to represent uncertain input parameters regarding the causes of parameter uncertainty that affect each parameter.

## 1.4 Thesis outline

Chapter 1 is an introductory chapter on motivation behind the research, its approach, analysis contributions, and an outline for the following thesis.

Chapter 2 precises the context of the study. risk analysis for critical facilities is detailed. Then, the major problem (uncertainty) from what today’s industrial risk analysis methodologies suffer is highlighted. The main sources of uncertainty: (i) completeness, (ii) model and (iii) parameter uncertainties that might affect an assessment are exhibited. The causes of each source of uncertainty are discussed. Then, an overview on the steps behind dealing with the presented uncertainty sources is presented. At the end, we identify at which stages and where these uncertainties affect the risk analysis process.

Chapter 3 presents a state of the art review on two major fields: (i) literature review on parameter uncertainty analysis and representing, propagating of uncertain knowledge. These most common approaches used to characterize parameter uncertainty in model inputs obtained from different sources, such as statistical data and expert judgments are detailed. An overview on interval analysis, fuzzy theory, probability theory, evidence theory, and the mixed probabilistic-fuzzy approach is given. (ii) cyber-security for critical

facilities and industrial control system. This in order to treat the know completeness uncertainty. The definition of security related risks is provided, and an exhaustive review on the different threats and vulnerabilities on control systems are highlighted. This review will pave the way to handle the known completeness uncertainty by introducing cyber-security within industrial risk analysis.

In Chapter 4, we propose a new fuzzy semi-quantitative methodology to handle parameter uncertainty during likelihood analysis. A semi-quantitative approach was proposed by the INERIS for likelihood analysis as an alternative to the drawbacks existed in the pure quantitative or qualitative approach. This semi-quantitative approach uses interval analysis to characterize input parameters. However, interval analysis may lead to likelihood underestimation in some cases due to parameter uncertainty. Thus, fuzzy theory is introduced as an alternative to interval theory to deal with parameter uncertainty due to imprecision of recorded data and subjectivity of expert elicitation. The qualitative approach proposed in chapter 4 for safety/security likelihood analysis using ATBT is replaced by the fuzzy semi-quantitative to dispose the shortcomings of the qualitative approach. The application of the proposed approach is demonstrated using a case study that cover safety and security.

In Chapter 5, we take a deeper look at parameter uncertainty that affects the effect analysis of risk scenarios. A fuzzy-probabilistic approach is proposed to deal with two causes of parameter uncertainty: imprecision and randomness. Then a comparison between the proposed approach and the existing approaches for uncertainty analysis is done by the application of these approaches to a loss of containment scenario (LOC), representing one of the most likely situations to occur in industry. The overall aim is to compare the different approaches, and to identify which mathematical theories should be used to represent uncertainty regarding the available information. Indeed, the existing uncertainty quantification approaches can lead to different representations of uncertainty in the outputs and hence to different decisions. We prove that the use of an inappropriate approach in an inappropriate place may lead to under or overestimation of risk and subsequently to a bad decision. Based on this comparison, a global methodology that mixes probability theory, fuzzy numbers and Dempster-Shafer theory of evidence is proposed to treat each cause of parameter uncertainty with the best suitable theory.

Chapter 6 addresses one aspect of completeness uncertainty which is known. A new methodology to introduce cyber-security related risks within industrial risk analysis is proposed. This approach combines Bow-Tie Analysis, commonly used for safety analysis, with a new extended version of Attack Tree Analysis, introduced for security analysis of industrial control systems. The combined use of bow-tie and attack tree provides an exhaustive representation of risk scenarios in terms of safety and security. The difference



in nature between safety and security initiating event causes some obstacles for likelihood analysis. For these reasons, we propose an approach for evaluating the safety/security likelihood level based on two-term likelihood parts, one for safety and one for security. The likelihood analysis approach presented in this chapter is qualitative. The aim is to show how can separately characterize safety and security events in order to evaluate safety/security scenarios in terms of likelihood. The shortcomings of this approach are listed and discussed. The end of this chapter, the application of this approach is demonstrated using the case study of a risk scenario in a chemical facility.

The end of this document is a summary of the initial objectives, the proposed methodologies and main contributions, the key findings and lessons of the thesis with conclusions and proposes work for future research.

# 2

## Context & problematic

**Summary:** In this chapter, we present the context of the risk analysis process used by the INERIS. The different steps the INERIS follows to analyze risks for critical facilities are detailed (Section 2.2).

Moreover, the major problem that faces the risk analysis prediction is introduced. The different sources, types and causes of uncertainties are presented (Section 2.3). Why uncertainty should be considered during risk analysis (Section 2.4) and the different steps to deal with these uncertainties are discussed (Section 2.5). Then, where these uncertainties affect the INERIS's risk analysis process is discovered (Section 2.6).

### Summary

---

<b>2.1</b>	<b>Introduction</b>	<b>19</b>
<b>2.2</b>	<b>Understanding hazard analysis for critical facilities in France - The INERIS methodology</b>	<b>20</b>
2.2.1	Representation of risk scenarios	21
2.2.2	Probability analysis using bow-tie based on the interval semi-quantitative approach	23
2.2.3	Severity analysis	28
2.2.4	Decision-Risk matrix	29
<b>2.3</b>	<b>Introducing uncertainty in industrial risk analysis: the different sources of uncertainty</b>	<b>31</b>
2.3.1	Parameter uncertainty	31
2.3.2	Model uncertainty	36
2.3.3	Completeness uncertainty	36
<b>2.4</b>	<b>Why uncertainty represents a major problem and should be considered?</b>	<b>37</b>
2.4.1	Red-River flood - Why risk analysts should be explicit about uncertainty?	37

2.4.2	The ASSURANCE benchmark project: Assessment of Uncertainties in Risk Analysis of Chemical Establishments . . . . .	39
<b>2.5</b>	<b>Addressing uncertainty during risk assessment . . . . .</b>	<b>41</b>
2.5.1	Steps to address parameter uncertainty . . . . .	41
2.5.2	Addressing completeness uncertainty . . . . .	41
<b>2.6</b>	<b>Where uncertainties affect the INERIS risk assessment process . . .</b>	<b>42</b>
2.6.1	Uncertainty in likelihood analysis: drawbacks of the interval semi-quantitative approach . . . . .	43
2.6.2	Uncertainty in effect analysis . . . . .	44
2.6.3	Completeness uncertainty in the risk identification process: No-considering of cyber-security threats . . . . .	47
<b>2.7</b>	<b>Conclusion . . . . .</b>	<b>48</b>

---

## 2.1 Introduction

The French regulation introduces the obligation for SEVESO critical facilities to carry out a compulsory hazard analysis and to update it every five years. The hazard analysis has to identify all known and possible major accident scenarios, together with their prevention and mitigation barriers. This analysis must present the probability and the possible consequences of each identified major accident. The INERIS provides this service by a team of risk experts that follow a systematic methodology based on the INERIS expertise. In the current Chapter, the INERIS risk analysis methodology used for SEVESO critical facilities is described as presented in Section 2.2.

In Section 2.3, the major problem of the hazard analysis process (uncertainty) is presented. The different sources of uncertainty: parameter, model and completeness uncertainties are detailed.

- parameter uncertainty: relates to the uncertainty in the computation of the input parameter values for conducting the assessment such as initiating event frequencies, etc.,
- model uncertainty: relates to the uncertainty in the assumptions made in order to build the models used in the assessment.
- completeness uncertainty: relates to contributions to risk that have been excluded from the assessment intentionally or not.

To illustrate the effects of uncertainties and why uncertainties should be considered during risk assessment, two different illustrative examples are given in Section 2.4. The first example is a real major accident that happened from not being open about uncertainty. The second example is about conclusions extracted from the European benchmark study ASSURANCE on the effect of uncertainty on the risk analysis prediction.

Uncertainties affect the risk assessment process at different phases. Section 2.6 identifies the sources of uncertainty that affect each phase in the INERIS risk analysis process. Parameter uncertainty affects likelihood analysis due to some limits in the interval semi-quantitative approach. The effect analysis step suffers from parameter uncertainty due to the lack of information regarding values of some input parameters and the natural variability of others. Completeness uncertainty affects the identification phase as well as the likelihood analysis phase due to non-consideration of cyber-security related threats.

Section 2.7 draws some conclusions and summarizes the Chapter.

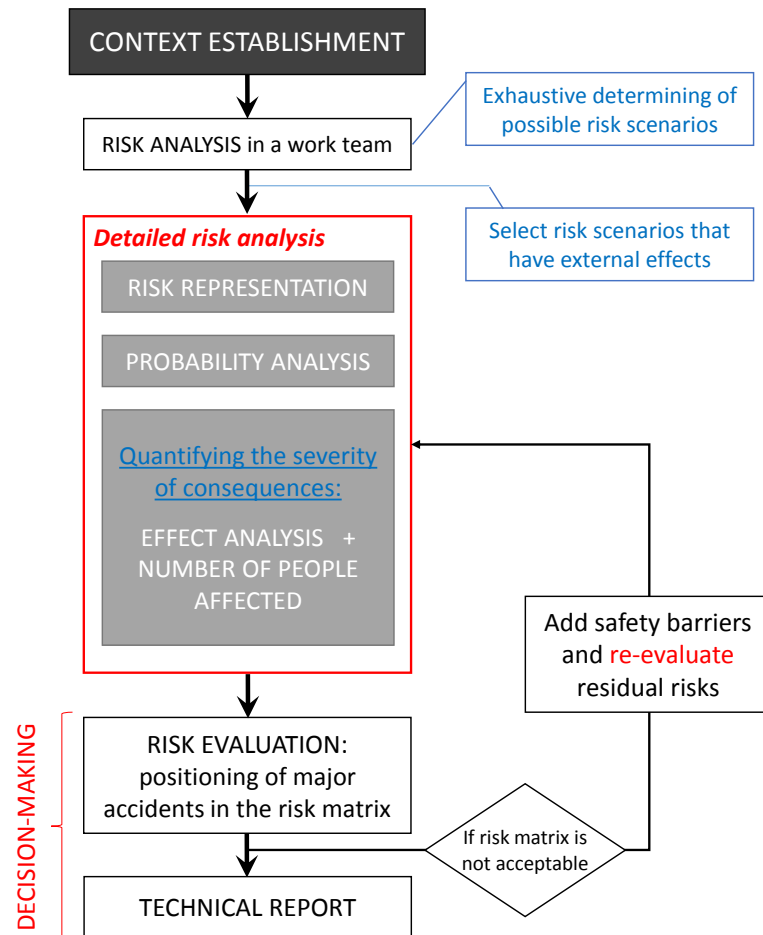
## 2.2 Understanding hazard analysis for critical facilities in France - The INERIS methodology

In the French regulatory framework industrial activities are classified according to their potential hazard. For high hazard facilities (also called ‘AS’ or ‘Upper tier Seveso establishments’), a hazard analysis is compulsory. Hazard analysis aims to demonstrate the public safety and it represents a tool for the inspection by the authorities. The results of the hazard analysis are used for the permit to operate and definition of the requirements for risk reduction on site.

In this section, we will present the different steps of conducting a systematic hazard analysis of high hazard facilities. Figure 2.1 presents the steps for assessing risks of critical industries in France as proposed by the INERIS [90].

- context establishment: aims to define the objectives of the risk management (what should be protected) and the decision criteria (when a risk is acceptable and when it is not);
- risk analysis: this step is the most substantial part of the hazard analysis process for decision making. It aims to analyze in details risk scenarios that may have effects outside the facility being studied. It involves the identification of mechanisms that lead to failure of equipment and loss of containment (central events), and the identification of consequence effects that may follow the loss of containment. In the INERIS method, the working group identifies the central events to be used in the risk analysis. INERIS then aims to find an exhaustive list of causes for the central events defined. The main sub-steps of this step are as follows:
  - ✓ risk identification: risk identification is the step of modeling risk scenarios (undesirable events) that could potentially harm people and the environment. A risk scenario is modeled as a sequence of events starting from the causes of the undesirable event ending by the related consequences (dangerous phenomena). This step is detailed in Section 2.2.1.
  - ✓ likelihood analysis: aims to calculate the probability of occurrence of the identified risk scenarios. More details are given in Section 2.2.2.
  - ✓ severity analysis: combination of dangerous phenomena effect intensity and the vulnerability of people potentially exposed at a given point. This step is more detailed in Section 2.2.3.
- decision making: the results of likelihood and severity analyses are mapped in a risk matrix to give a level of risk to each risk scenario. This matrix allows the authorities to assess the acceptability of risks generated by a facility in a given environment.

If the risk matrix is acceptable, the facility is considered to be compatible with its environment and a permit to operate will be given. If not, unacceptable risks should be treated and residual risks are then evaluated. Or the permission is withdrew or refused. More details on conducting this step are given in Section 2.2.4.



**Figure 2.1** – The hazard analysis process for high hazard critical facilities.

Risk analysis is the most critical step in the hazard study process. This step represents a challenging problem due to uncertainty. In order to highlight this problem, the rest of this section details the different steps of a detailed risk analysis. The representation of risk scenarios, likelihood analysis and severity analysis are presented.

### 2.2.1 Representation of risk scenarios

This step aims to model each risk scenario from its root causes via central event to related consequences. In the methodology used by INERIS, the representation of accident scenarios is usually realized through a modeling methodology such as bow-tie analysis. Bow-tie analysis is a very prominent method to identify and model risk scenarios

[55]. It presents a combination between fault tree analysis (FTA) and event tree analysis (ETA). FTA and ETA respectively describe the relationship between the undesirable event (the main event of a risk scenario), its causes and its consequences for a systematic representation of hazard ([136]; [53]).

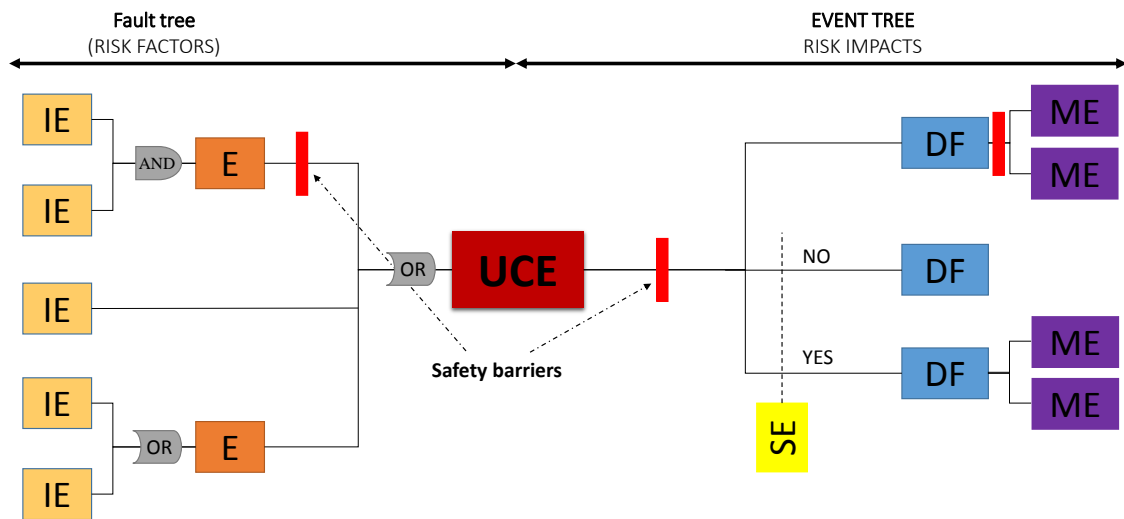
To identify and represent risk scenarios, a working group will be created. It could, for example, consist of the plant safety manager, several operators and risk experts. This working group will identify the following elements of each accident scenario:

- The undesirable central event (UCE) or main event to be considered;
- The initiating events (root causes) lying underneath the undesirable central event. Typical root causes are seal failure, operator errors, falling objects, etc.;
- The secondary events such as ignition after the UCE;
- The consequence events (dangerous phenomena) and the associated major accidents of the central events;
- The safety barriers which may prevent the occurrence of the accident. Prevention (before the central event) and protection (after the central event) barriers are considered if they meet the following requirements:
  - ✓ independence regarding the occurrence of the event they prevent;
  - ✓ effectiveness;
  - ✓ response time adapted to the kinetic of the accident they prevent;
  - ✓ maintainable;
  - ✓ testable.

When a mitigation barrier is identified in an accident scenario, both the scenario describing the consequences following the function and the malfunction of the safety barrier need to be taken into account.

These risk contributors are then connected together to construct the bow-tie. Figure 2.2 presents a schematic diagram of the bow-tie analysis and how the risk contributors listed below are connected using AND or OR gates. If any of the root causes can cause the intermediate event, an OR operator is used. If multiple root causes are required for the occurrence of the intermediate event, an AND operator is used. The definition of each term is detailed in Table 2.1.

Once the identification process is realized, the scenarios' likelihoods can be calculated. Bow-tie analysis and in addition of being a tool for representing a risk scenario, it implicitly provides a likelihood analysis methodology. This methodology aims to propagate the frequency or probability of initiating events after being characterized through the bow-tie in order to calculate the probability of dangerous phenomena.




**Figure 2.2** – Elements of a “Bow-Tie” diagram.

In France, probability analysis can be qualitative, quantitative or semi-quantitative. Quantitative information to perform a precise likelihood analysis is often missing due to time and financial constraints. On the other hand, quantitative information, if available, is lost by using a qualitative probability analysis. For these reasons, the INERIS has developed an interval semi-quantitative approach as a solution to handle these issues. In the next section, a brief details on likelihood analysis using bow-tie based on the interval semi-quantitative approach is provided.

### 2.2.2 Probability analysis using bow-tie based on the interval semi-quantitative approach

Evaluating probability of an accident using bow-tie analysis is performed by: (i) characterizing input data, (ii) propagating these characterized inputs through the bow-tie analysis. Characterizing input data aims to qualitatively, quantitatively or semi-quantitatively represent the information provided either by experts or derived from historical data, with the consideration of uncertainty. Propagating the characterized inputs through the bow-tie aims to calculate the probability of ERC, PhDs (dangerous phenomena) and AMs. This is done by solving the AND and OR gates, considering the occurrence of secondary events and the existence of risk barriers. The interval semi-quantitative approach uses frequency classes instead of probability to characterize inputs and outputs for probability analysis, section 2.2.2.3 shows the translation of frequency classes obtained after conducting the interval semi-quantitative approach into probability levels outlined by the French ministerial laws.



Abbreviation	Signification	Definition
IE	Initiative event	Direct cause of a loss of containment or physical integrity
E	Event	Physical integrity caused by the occurrence basic events
UCE	Undesirable central event	The unwanted event such as a loss of containment, etc.
SE	Secondary event	Characterize the source term of an accident, event such as ignition occur after the UCE
DF	Dangerous phenomenon	Physical phenomena that can cause major accidents, explosion, dispersion, fire
ME	Major Effect	Damages caused by the effects of an PhD on people, environment or goods
	Safety barrier	Measures taken place to reduce the probability of undesirable event and the effects of accidents
AND, OR gates		Describe the relationships between events

**Tableau 2.1** – Abbreviations, significations and definitions of elements listed in the Bow-Tie diagram.

### 2.2.2.1 characterizing input data

The interval semi-quantitative approach characterizes the relevant input data on a scale of different classes characterized by probability or frequency ranges. For probability analysis using bow-tie based on the semi-quantitative approach, three types of nodes should be characterized:

- basic events - in terms of annual frequency of occurrence (Section 2.2.2.1.1);
- secondary events - in terms of conditional probability of occurrence after the occurrence of ERC (Section 2.2.2.1.2);
- risk barriers - in terms of confidence levels (CL). A CL depends on the probability of failure for each risk barrier (Section 2.2.2.1.3).

Before starting, it should be noted that there is a difference between probability and frequency. But in a simplified manner, these two notions of probability and frequency coincide when frequencies are low (less than 1 time every 10 years) [40].

#### 2.2.2.1.1 Characterizing frequencies of basic events

The interval semi-quantitative scale used by INERIS to characterize the input data is presented in Table 2.2. A class of frequency in terms of an interval is given to each basic event as input. Each class is considered to cover a broad range of occurrence frequencies. Giving a frequency class to an input event is based on the process of asking experts, or by translating the quantitative data into a class. A class is linguistically elicited from experts since they prefer linguistic judgments rather than precise value. The translation

of a quantitative value into a class is performed based on Equation 2.1 below:

$$Class(f) = -Ent(\log(f)) - 1. \quad (2.1)$$

where,  $Ent$  is the integer part,  $\log$  is the logarithm base 10 and  $f$  is the frequency value of the event under translation (e.g. an event with a frequency equals  $4 \times 10^{-4}$  is of class F3).

F-2	$10^{+1}/year \leq frequency \leq 10^{+2}/year$	10 to 100 times/year
F-1	$10^0/year \leq frequency \leq 10^{+1}/year$	1 to 10 times/year
F0	$10^{-1}/year \leq frequency \leq 10^0/year$	1 time every 1 to 10 years
F1	$10^{-2}/year \leq frequency \leq 10^{-1}/year$	1 time every 10 to 100 years
FX	$10^{-(X+1)}/year \leq frequency \leq 10^{-X}/year$	

**Tableau 2.2** – Determining the frequency classes based on the semi-quantitative approach.

### 2.2.2.1.2 Characterizing the occurrence probabilities of secondary events (ESs)

INERIS has proposed a semi-quantitative method to evaluate the occurrence probability of secondary events [65]. The probability of failure of an SE based on this semi-quantitative approach can be equal to 1 for a conservative analysis, or a factor of 10 inspired from (i) databases after being rounded-up to sub power of 10 or (ii) from expert opinions. For example, it can be  $10^{-1}$  for a leakage of a product which the expert judges to be low flammable,  $10^{-2}$  for less flammable material.

The different probability values used to characterize the probability of ESs are presented in Table 2.3.

Probability values	$10^{-3}$	$10^{-2}$	$10^{-1}$	1
Probability Class	P3	P2	P1	P0

**Tableau 2.3** – Representative values of the probability of failure of ESs.

### 2.2.2.1.3 Characterizing probability of failures for risk barriers in terms of Confidence Level - CL

Because little feedback from industrial sites is available, there is insufficient data to calculate the failure probabilities of risk barriers. To compensate for this weakness, INERIS has proposed a semi-quantitative approach to give a confidence level for a risk barrier based on norms ISO IEC 61508 and 61511 ([85]; [36]). These norms provide principles that allow attributing CLs for the safety instrumented systems. INERIS has expanded

these principles to cover the whole type of industrial risk barriers. The CL shall be derived by taking into account the required risk reduction that is provided by the risk barrier (see [87] for more guidance) or probability of failure data if existed. Table 2.4 presents how CLs can be determined based on either the average probability of failure on demand ( $PFD_{avg}$ ) or the risk reduction (RR). This CL will be used in calculating the output frequency classes of ERC, PhDs and EMs.

Confidence level (CL)	Average probability of failure on demand ( $PFD_{avg}$ )	Risk reduction (RR)
4	$10^{-5} \leq PFD_{avg} < 10^{-4}$	$10\ 000 < RR \leq 100\ 000$
3	$10^{-4} \leq PFD_{avg} < 10^{-3}$	$1\ 000 < RR \leq 10\ 000$
2	$10^{-3} \leq PFD_{avg} < 10^{-2}$	$100 < RR \leq 1\ 000$
1	$10^{-2} \leq PFD_{avg} < 10^{-1}$	$10 < RR \leq 100$
0	$10^{-1} \leq PFD_{avg} < 1$	$1 < RR \leq 10$

**Tableau 2.4** – Confidence Level: probability of failure on demand.

### 2.2.2.2 propagating input data

After characterized the inputs for probability analysis, these characterizations are then propagated through the bow-tie in order to give occurrence frequency classes for ERC, PhDs and AMs [86]. Only the propagation rules for the AND and OR gates (Eqs 2.2 et 2.3, respectively) based on the interval semi-quantitative approach are described in section to be used in the next section. The other roles will be described in details in Chapter 5.

- In the case of OR gate, the frequency class of the intermediate event is equal to the minimum frequency class of the root causes as presented in Eq 2.2. The minimum operator is used here conversely to the usual because the used scale is exponential;

$$Class(OR_{output}) = \min[Class(EI1), \dots, Class(EIn)] \quad (2.2)$$

- In the case of AND gate, the frequency class of the intermediate event is equal to the maximum of the frequency classes of the input root causes. The output frequency class of an AND gate with  $n$  input is as calculated below:

$$Class(AND_{output}) = \max[Class(EI1), \dots, Class(EIn)] \quad (2.3)$$

Finally, after the end of the propagation process, a frequency class is given to each of the ERC, PhDs and AMs.

### 2.2.2.3 Final probability

In France, the ministerial laws precise that the probability of dangerous phenomena should be determined regarding the scale presented in Table 2.5

Qualitative scale	Level	Designation	Quantitative meaning
<b>Likelihood</b>	E	<b>Very unlikely event:</b> event that is practically impossible, very low chance of occurrence	$10^{-5}$
	D	<b>Unlikely event:</b> Low chance of occurrence, already happened but the presence of risk measures reduce the probability of occurrence	$10^{-4}$
	C	<b>Moderate event:</b> similar events have already been seen in the activity sector	$10^{-3}$
	B	<b>Likely event:</b> already occurred or may occur during operational life of the installation	$10^{-2}$
	A	<b>Very likely event:</b> can frequently occur (several times) during operational life	

**Tableau 2.5** – Probability of occurrence scale extracted from the French ministerial order of 29/09/2005 related to the evaluation of risk.

In the interval semi-quantitative approach, annual frequency classes are given to the outputs (ERC, PhDs and EMs). These frequency classes are translated into probability level as follows:

- if the frequency class is greater than zero (frequency is less 1 time/year), then the approximations of frequency classes in terms of probability levels is as presented in Table 2.6. Level E is given to any frequency classes less than 5.

Quantitative scale	$10^{-5}$	$10^{-4}$	$10^{-3}$	$10^{-2}$	
Frequency classes	F5	F4	F3	F2	F1
Probability levels	<b>E</b>	<b>D</b>	<b>C</b>	<b>B</b>	<b>A</b>


**Tableau 2.6** – Transforming of frequency classes into probability levels.

- if the frequency class is less than zero (for example, an event may occur two times a month and it is class of frequency is F-2), then the probability level can not be assimilated to the frequency class. In this case, the level is approximated as:  $\text{level(probability)} = [10^{-1}, 1]$  which is of level A.

### 2.2.3 Severity analysis

In the French regulatory context, the severity of risks are determined regarding the intensity of effects of dangerous phenomena and the number of persons exposed. Dangerous phenomena may have three types of effects: thermal, toxic and over-pressure. Different levels of intensity are attached to each type of effect as presented in Table 2.7. Three levels are distinguished to classify the intensity of dangerous phenomena:

- significant lethal effect (about 5% probability of fatality);
- lethal effect (about 1% probability of fatality);
- irreversible effect (irreversible health effects).



Threshold effects on human	TYPES of effects		
	Thermic	Toxic	Over-pressure
<b>LETHAL SIGNIFICANT (SELS)</b>	<b>8kW/m<sup>2</sup></b> ou (1 800 kW/m <sup>2</sup> ) <sup>4/3,s</sup>	<b>CL 5%</b>	<b>200 mbar</b>
<b>LETHAL (SEL)</b>	<b>5kW/m<sup>2</sup></b> ou (1 000 kW/m <sup>2</sup> ) <sup>4/3,s</sup>	<b>CL 1%</b>	<b>140 mbar</b>
<b>IRREVERSIBLE (SEI)</b>	<b>3kW/m<sup>2</sup></b> ou (600 kW/m <sup>2</sup> ) <sup>4/3,s</sup>	<b>SEI</b>	<b>50 mbar</b>
<b>INDIRECT (broken windows)</b>			<b>20 mbar</b>

**Tableau 2.7** – French end-point values for the intensity of thermal radiation, toxic and over-pressure effects.

The INERIS uses complex mathematical models to perform effect analyses for the identified risk scenarios. These models are representations of phenomena being studied. The aim of effect analysis is to determine the distances of affected zones (distance modeling) by the defined types of effect. The used mathematical models depend on large number of input parameters, such of these parameters are:

- occurrence conditions of the phenomenon under study: the term sources by defining for example the size of the crack, temperature, pressure, etc.
- conditions related to the environment: meteorological data such as the stability classes;
- etc.

Values of input parameters to perform effect analyses are determined based on experimental or statistical data. However, if experimental data is not available, the INERIS is

based on the state of the art and the experience of its experts to set the values of input parameters used for effect analysis.

Finally, the severity of a given potential major accident is deduced from the number of persons exposed to the pre-defined levels of intensity as presented in Table 2.8.

Resulting severity class	Affected area (Intensity)		
	LETHAL SIGNIFICANT (SELS)	LETHAL (SEL)	IRREVERSIBLES (SEI)
DESASTROUS	> 10	> 100	> 1000
CATASTROPHIC	1 à 10	10 à 100	100 à 1000
IMPORTANT	1	1 à 10	10 à 100
SERIOUS	0	1	1 à 10
MODERATE	0	0	< 1

**Tableau 2.8** – French scale for the classification of the severity of a potential accident.

### 2.2.4 Decision-Risk matrix

Probabilities of major accidents and their severity of consequences are introduced in a decision tool called risk matrix. This matrix allows the authorities to assess the societal acceptability of the risk generated by a facility in a given environment. If the risk matrix is acceptable, the facility is considered to be compatible with its environment. Table 2.9 is the risk matrix used by the French authorities to assess this compatibility. Each major accident identified during the risk analysis and based on its probability range (Table 2.5) and its severity class (Table 2.8) is mapped on one of the cells in Table 4. Once all identified accidents are characterized with regard to this matrix, the overall acceptability of the matrix, and thereby the overall acceptability of the permit of the facility, is assessed.

This risk matrix consists of three areas:

- an acceptable area (in white): if all identified scenarios are in the acceptable area, the permit to operate is granted;
- an unacceptable area (in red): if one or more scenarios are in the unacceptable area, the permit to operate is not granted (risk measures need to be set) or withdrawn (if risks are really high);
- an 'ALARP' (As Low As Reasonably Practicable) area (in yellow): for each accident scenario in this area, continuous improvement of the safety is asked to operators. There are two specific cases among the ALARP areas:

probability gravity	VERY LOW E	LOW D	AVERAGE C	HIGH B	VERY HIGH A
DISASTROUS					
CATASTROPHIC	ALARP class 2	ALARP class 2			
IMPORTANT	ALARP	ALARP	ALARP class 2		
SERIOUS			ALARP	ALARP class 2	
MODERATE					ALARP class 2

Acceptable

Risk to be reduced

Unacceptable

**Tableau 2.9** – French risk matrix.

- ✓ the area disastrous/E (upper left box): If an accident scenario is in this area, a distinction is made between new and existing facilities. For a new facility the situation is acceptable if and only if this scenario has at least one barrier, and if this barrier was not considered, the remaining frequency would still be E. For an existing facility, ALARP class 2 conditions is applied for the scenarios in this box;
- ✓ ALARP class 2: The total number of accidents in the ALARP 2 boxes of the diagram must be five or lower. If there are more than five accident scenarios, additional technical barriers must be installed in such a way that the amount of ALARP class 2 accident scenarios reduces to five (or less). More than five ALARP class 2 accidents scenarios are acceptable if and only if:
  - ◇ all these ALARP class 2 accident scenarios have probability class E;
  - ◇ all these ALARP class 2 accident scenarios have at least one barrier;
  - ◇ if this barrier was not considered, the remaining frequency would still be E (for each of these ALARP class 2 accident scenarios). All other situations with more than five ALARP class 2 accident scenarios are unacceptable.

If a facility generates a risk that is considered as unacceptable, the operator has the responsibility, on its own funds, to improve the safety in the establishment and to install additional safety measures. The safety must be improved until the situation becomes acceptable or ALARP. If the situation is ALARP, the operator must prove in the risk analysis that all risk reducing measures at an acceptable cost have been implemented.

## 2.3 Introducing uncertainty in industrial risk analysis: the different sources of uncertainty

Uncertainty is a measure of the “goodness” of an estimate. Without such a measure, it is impossible to judge how closely the estimated value relates to or represents reality. However, to help understand the concept of uncertainty, and to be able to treat uncertainty in a structured manner, this Section presents a detailed summary on sources of uncertainty and their related causes. Uncertainty in risk analysis has three main sources presented as follows [47]:

- parameter uncertainty: is the uncertainty in the values of the parameters of a model given that the mathematical form of that model has been agreed to be appropriate. Detailed definition of parameter uncertainty, its types and causes are given in Section 2.3.1;
- model uncertainty: is related to an issue for which no consensus approach or model exists and where the choice of approach or model is known to have an effect on the risk analysis prediction. See Section 2.3.2 for further details;
- completeness uncertainty: relates to risk contributors that are not considered in the analysis. Section 2.3.3 highlights the different causes of completeness uncertainty.

### 2.3.1 Parameter uncertainty

Parameter uncertainty relates to the uncertainty in the computation of the input parameter values used to quantify the likelihood and severity of a risk scenario[47]. Examples of such parameters are initiating event frequencies, failure probabilities, and wind speed (used in effect analysis of risk). This uncertainty arises from the inability to set exact values for these input parameters ([33]; [170]; [53]). Where some of these parameters vary randomly while others are imprecisely known due to measurement errors, lack of knowledge or other causes. This results in two different types of parameter uncertainty: aleatoric and epistemic.

However, misunderstandings and conflicts are presented by confusing the “causes of parameter uncertainty” with the “types of parameter uncertainty”. Types of parameter uncertainty are an upper level in the hierarchy of defining parameter uncertainty comparing to the causes of parameter uncertainty. In other words, type is determined regarding the causes of parameter uncertainty, where each type has different causes. In what follows, the main types and causes of parameter uncertainty are separately discussed as presented in Sections 2.3.1.1 and 2.3.1.2, respectively.



### 2.3.1.1 Types of parameter uncertainty

In risk studies, we differentiate between two types of parameter uncertainty affecting the analysis as follows:

- stochastic or aleatory uncertainty is due to randomness or natural variability (e.g. wind speed, time of failure of a mechanical component). The wind speed can be  $3m/s$  at time  $t$  and  $5m/s$  at  $t + \delta t$ . Thus, if the parameter sometimes has one value and sometimes has another value, then it has aleatory uncertainty ([69]; [23]; [70]; [78]);
- epistemic uncertainty is due to imprecision, vagueness or ignorance (e.g. a release breach diameter). If the parameter has one value but we don't know what it is due to lack of information, then the parameter has epistemic uncertainty ([130]; [49]; [173]; [156]; [61]). If a person "X" is guessing the height of his/her friend "Y," the answer would be uncertain (around  $a$  cm or between  $a$  cm and  $b$  cm). In this case, epistemic uncertainty is attached. This uncertainty can be reduced by adding information (ask the person about his/her height).

It should be noted that epistemic uncertainty can be reduced by further studies while aleatory uncertainty cannot ([14]; [46]). Additional effort may yield a better estimate of the magnitude of variability, but it will not tend to reduce it. However, both types of parameter uncertainty can influence the same parameter. While some parameter values are random and known statistically in terms of probability distributions, the mean or standard deviation of their distribution is uncertain, i.e. it is not precisely known. This means that this parameter is affected by aleatory and epistemic uncertainty. In what follows, we shall consider this last mixture as a new type of parameter uncertainty.

For risk evaluation to be effective, these different types of uncertainty relating to the analysis's input parameters must be represented and quantified. Figure 2.3 summarizes the different types of parameter uncertainty with practical examples of aleatory, epistemic and the mixture uncertainties. In the example of epistemic uncertainty in Figure 2.3, we suppose that the expert has no empirical data about the depth of the vessel at first.

In the next section, the causes of these types of parameter uncertainty will be detailed.

### 2.3.1.2 Causes of parameter uncertainty

In industrial risk analysis, different causes of parameter uncertainty can be found ([172]; [167]):

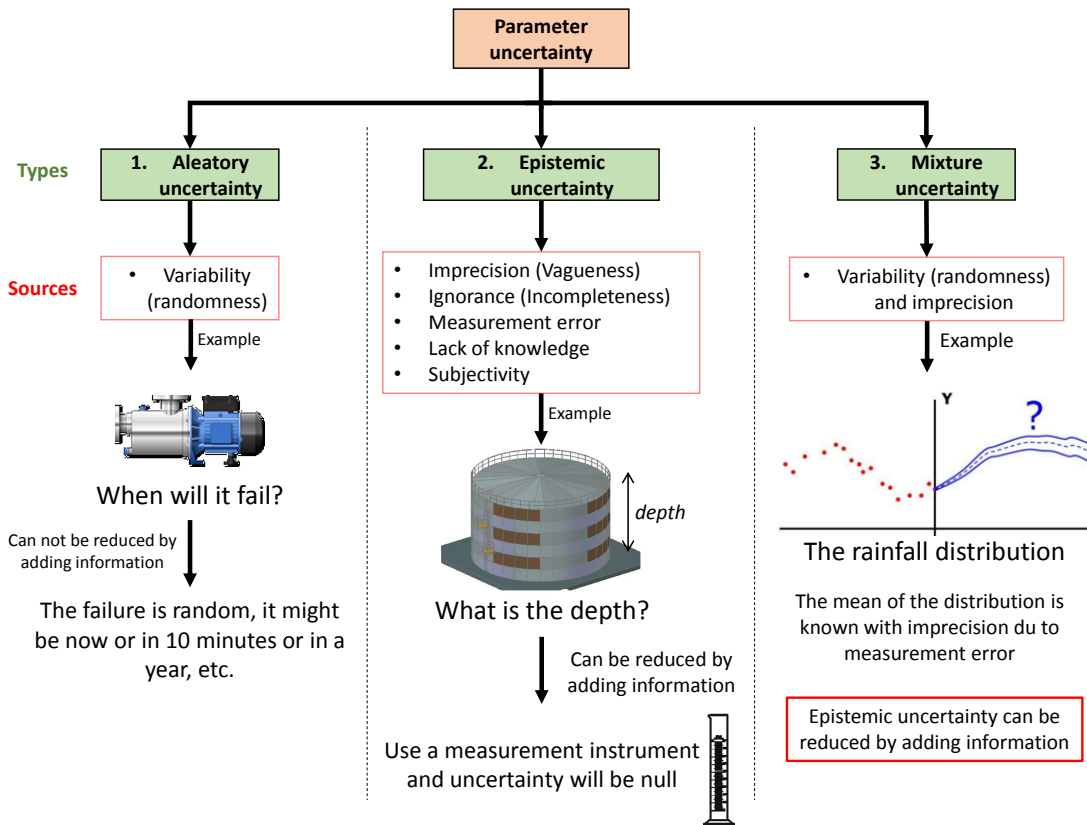


Figure 2.3 – Types of parameter uncertainty.

**Natural variability**

Natural variability refers to changes in the value of a parameter with time. Variability may result in a range of possible outcomes of a given situation. This cause of uncertainty represents the only cause of aleatory uncertainty.

**Lack of information or knowledge**

Lack of information represent the principle cause for uncertainty. It represents the situation when analysts approximately characterize the parameter of interest instead of giving an exact characterization. Providing approximation for input parameters takes place if there is no enough information to characterize these parameters, or when analysts believe that error factors are attached to the measurements. Lack of information has different classes presented as follows:

**Measurement errors or imprecision** Measurements refer to the engineering, i.e. measurement of physical quantity such as weight, temperature, length, etc. that are often affected by uncertainty. The quality of measurements depends on the experience of the analysts who perform the measurements and the used instrumentation. Mistakes

can be done that lead to errors in the measurement. This measurement errors result in imprecision about the true value of the parameter being measured. Another source of imprecision is the process of asking experts. Experts often provide a range (interval) about the value of a parameter instead of a precise value (for example, the true value of X is belong to the interval [a,b]).

**Subjectivity or vagueness** Subjectivity represents uncertainty in interpreting the linguistic terms (words such as high, very high, etc.) provided by experts. However, these words mean different things to different analysts. A simple example of this subjectivity from a study done by [141] is presented in Figure 2.4. In that study risk analysts have been asked about what they understood by the words low, moderate and high. The histograms in Figure 2.4 show the range of variation in the individual interpretation of those words. We can clearly see the wide difference between the interpretations of the words.

This cause of uncertainty can be reduced by translating the qualitative words into a quantitative scale (e.g. ask the expert what does he/she mean by high quantitatively), and gathering multiple opinions from different experts. But, resorting to the elicitation of multiple expert opinions will generate an new cause of uncertainty which is the lack of consensus(see Section 2.3.1.2 for more details).

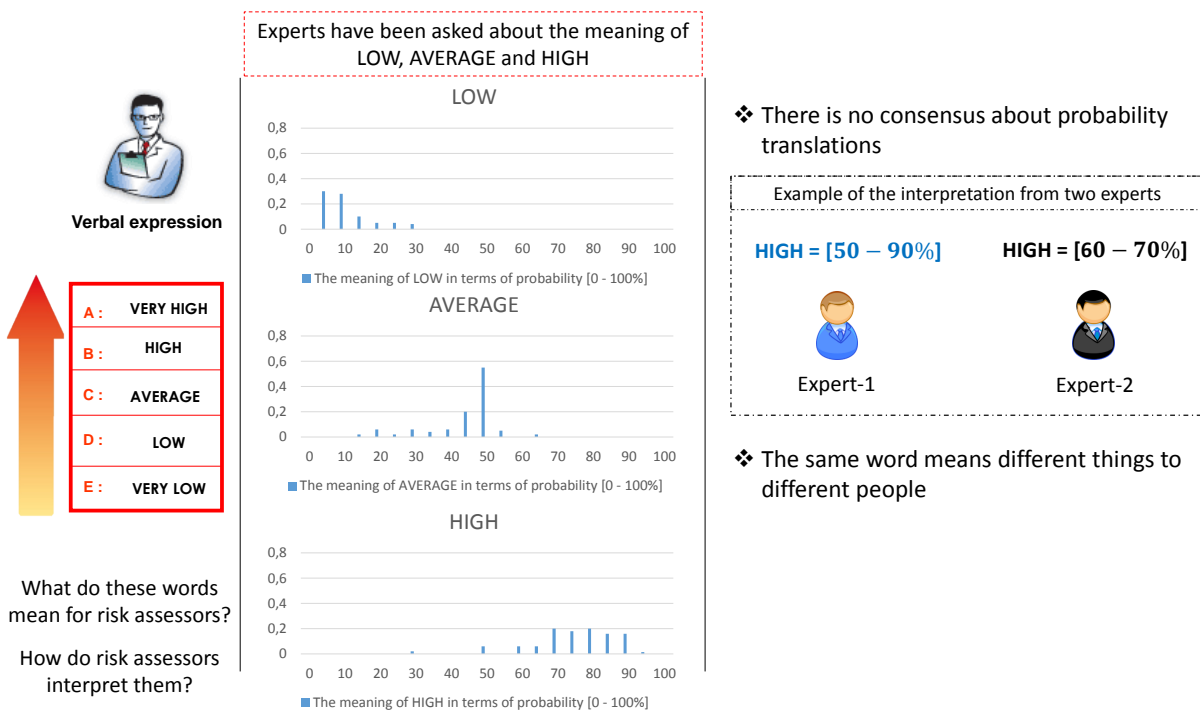


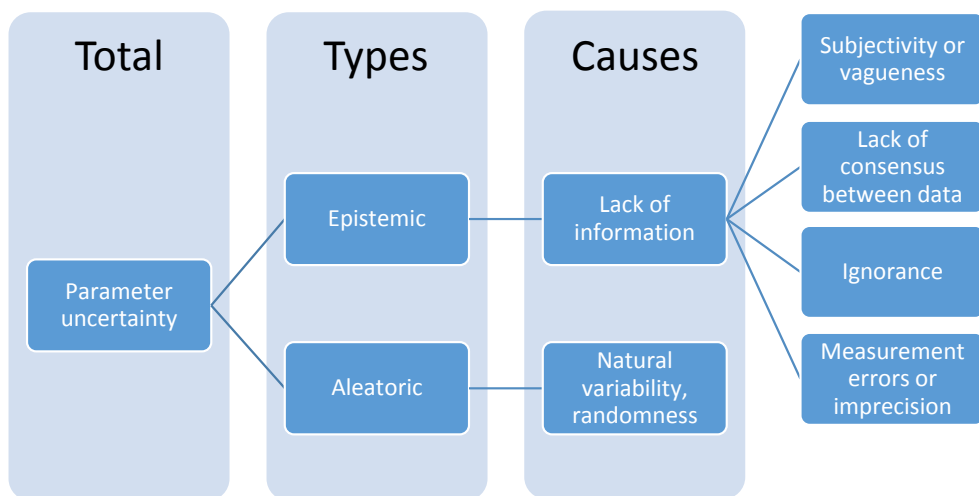
Figure 2.4 – Uncertainty in using expert elicitation to characterizing likelihood of inputs.

**Lack of consensus between sources of data** This cause is generated when a database or an expert suggests a description or a value for an input parameter, while others suggest a different one. This conflict can be due to different sources:

- different expert might have different opinions regarding a specific input depending on their own experiences and knowledge;
- values in databases: (i) can be wrong affected by errors (but not identified as wrong), (ii) are bias if the databases are generic and not relevant to the problem at hand, or (iii) calculated using different not accurate methods and models.

**Ignorance** The term “ignorance” means a partial incertitude that arises because of limits on empirical study or prediction. By means of the following example, suppose that we want to model some data about the seasonal weather. Let us assume that the possible outcome is either wet or dry. Now the meteorological institute provided evidence that it will be a dry season with 0.6 probability and another set of evidence suggests a wet season with a probability of 0.2. Since the probability of a dry and a wet season adds up to  $(0.6 + 0.2 = 0.8)$ , the remaining 0.2 actually implies ignorance with regards to the weather. In other words, 20% of the season might be wet or dry, with uncertainty (ignorance) about the true percentages.

It should be noted that, moving from an uncertain parameter to less uncertain or certain parameter caused by lack of information can be achieved by gathering more data about the parameter being characterized. To conclude this section, Figure 2.5 presents the relations between types and causes of parameter uncertainty.



**Figure 2.5** – Classification of parameter uncertainty.

### 2.3.2 Model uncertainty

Model uncertainty stems from the fact that risk models are based on simple mathematical equations which represent reality but cannot completely characterize the complex physical processes of any given phenomenon ([76]; [138]). Different approaches may exist to represent certain aspects of a phenomenon and none is clearly more correct than another. Model uncertainty is generated from two main causes:

1. limitations in the analyst's phenomenon knowledge;
2. deliberate simplifications introduced by the analyst.

In real life when analyzing complex phenomena, compliance between the model assumption and the properties of the system being analyzed never exists in an absolute sense. Then a very important question may be posed in this situation which is, the model can be accepted in spite of infringing one or more of the conditions supporting the model.

As an example, consider application of the model given by Eq 2.4 in order to predict the velocity of an object dropped from a height  $h$ . Where  $g$  is the acceleration due to gravity. In this model, it can be argued that the air resistance against any dropped object is not considered. The air resistance can modify the speed of a dropped object, and it is influenced by the mass and the shape of the object (quantities that are not included).

$$v = \sqrt{2gh} \tag{2.4}$$

Model uncertainty can be mitigated by validating the models against experiments [159]. In [169], the authors suggest the “adjustment factor approach” to quantify model uncertainty. The principle of this approach is to employ the best model available, denoted  $Y^*$ , and compensate for the error associated with  $Y^*$  by introducing a factor  $E$ . This adjustment factor might be additive ( $E_a$ ) or multiplicative ( $E_m$ ), resulting in  $Y = Y^* + E_a$  or  $Y = Y^* \times E_m$ .

However, this source of uncertainty is not addressed in this work. All the risk models used are already validated and appropriate.

### 2.3.3 Completeness uncertainty

This section defines completeness uncertainty in the scope of industrial risk analysis. Completeness uncertainty relates to risk contributors that are not addressed during the analysis process. The causes of this source of uncertainty either are known but not included in the analysis or not known and therefore not addressed in the analysis.

The known completeness uncertainty (not included in the analysis) could have significant impact on the analysis outputs. Causes of this known incompleteness are the

following:

- The scope of the analysis does not include some classes of initiating events because they are related to new technologies introduced in the facility under study and are complex to be introduced in the analysis (see Figure 2.6(a) - left side);
- Some of the risk contributors are omitted from the analysis because their relative contribution is believed to be negligible (see Figure 2.6(a) - right side).

However, when a risk analysis is used for critical decision making that concerns human lives, its scope needs to be exhaustive to match the changes and cover the nowadays risk causes. If the analysis is incomplete, then either the analysis process should be upgraded to include the missing piece(s) or it should be demonstrated that the missing elements are not significant risk contributors. Section 2.6.3 details the known cause of completeness uncertainty that is not considered in most nowadays risk analyses. Then, in Chapter 3 we proposed an approach to treat this known cause of incompleteness to provide a more exhaustive and precise risk analysis.

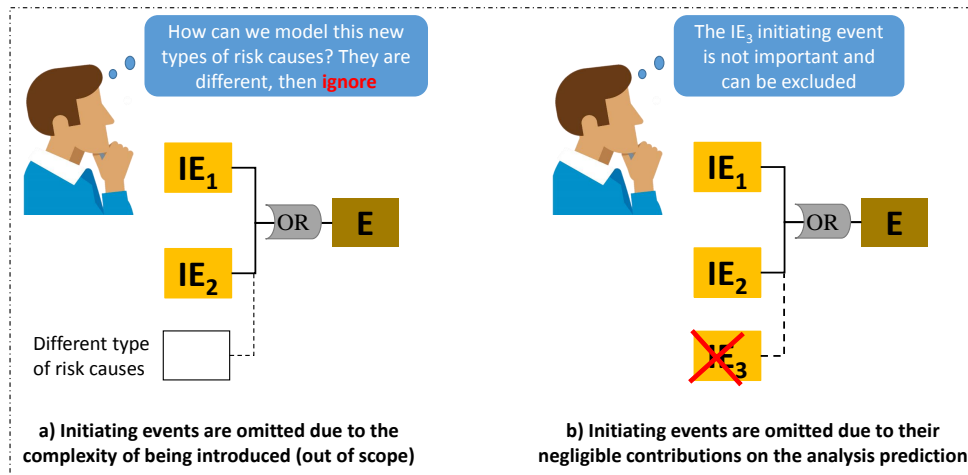
The second type of completeness uncertainty is the unknown. The causes of this type are due to the fact that some initiating events are unknown and the risk assessors are not sure about if the analysis they made is complete or not, see Figure 2.6(b). These causes of incompleteness are complex to be addressed during risk analysis.

## **2.4 Why uncertainty represents a major problem and should be considered?**

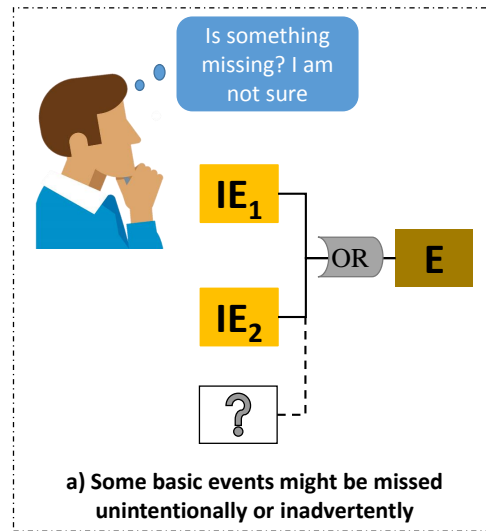
In this section, we will present the problem of not considering uncertainty during risk analysis by the help of real examples as presented in Sections 2.4.1 and 2.4.2. These example can illustrate the impact of not considering uncertainty extremely well. Section 2.4.1 presents the Red River flood accident of 1997 which devastated Grand Forks and other cities along the Minnesota-North Dakota border, USA [141]. In Section 2.4.2, the results of an European benchmark exercise on the effect of uncertainty in risk analysis are discussed.

### **2.4.1 Red-River flood - Why risk analysts should be explicit about uncertainty?**

At the forefront of the story, the river was flooding each year, so community leaders in cities located along the Red River had to manage the risk of flooding by levees. The national risk analysis service provided predictions on how high the river would reach in



(a) Causes of the known completeness uncertainty.



(b) Cause of the unknown completeness uncertainty.

**Figure 2.6** – Examples of known and unknown causes of the completeness uncertainty.

Grand Forks. In 1997, the river height prediction was 49 feet while the levee height was 51 feet. However, uncertainty surrounding the estimate was calculated and it was  $\pm 9$  feet. But, this uncertainty was not considered or communicated to decision makers because of concerns of alarming the public. Even so it was clear that there is a possibility that the flood surpasses the levee.

In actual fact the flood in that particular year went to 54 feet. The consequences of that was too high when damages were estimated about 3.5 billion \$. The other consequence was the lost of trust and confidence in the authority for managing the flood risk, and also the authority for who provided the predictions. Ironically, some residents and Grand Forks

officials insisted that the city would not have been lost if the analysis would have been more accurate. In other words, uncertainty has been considered. The Grand Forks mayor said, “I don’t like to be critical, but we were told 49 feet by the weather service”, had the city known how high the waters would rise, the devastation “would have been preventable”. Figure 2.7 shows the serious consequences from not having considered uncertainty.

Thus, two important reasons for considering uncertainty :

1. it is necessary for decision making, you need to know about uncertainty to make a good decision;
2. if you are not open about uncertainty, then you risk losing credibility and trust.



**Figure 2.7** – Grand Forks after levee over-topped.

### **2.4.2 The ASSURANCE benchmark project: Assessment of Uncertainties in Risk Analysis of Chemical Establishments**

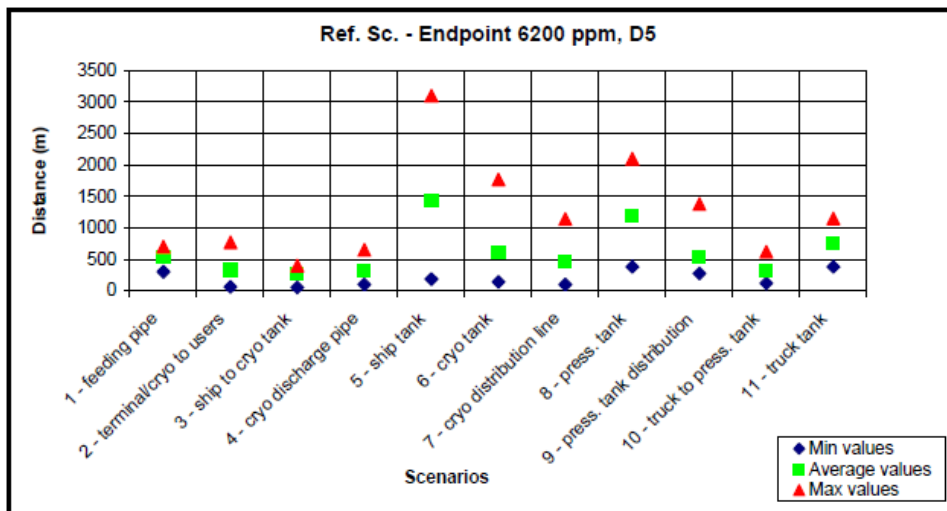
ASSURANCE is a benchmark exercise, which was completed in 2001. Seven teams from different European countries have joined this project. Each one of the joined teams and based on its own input data and expertise has performed a risk analysis on the same risk scenarios in an ammonia storage facility. The results for probability and effect analysis of each team for the studied risk scenarios are presented in Table 2.10 and Figure 2.8, respectively.



#	Top Event <sup>a</sup>	Partner number							Range of deviation
		3	4	1	5	7	2	6	
1	Major ammonia leak from 8'' feeding pipe	2.1 10 <sup>-4</sup>	5.0 10 <sup>-6</sup>	9.5 10 <sup>-5</sup>	1.6 10 <sup>-5</sup>	2.0 10 <sup>-5</sup>	7.7 10 <sup>-6</sup>	3	5.0 10 <sup>-6</sup> - 2.1 10 <sup>-4</sup>
2	Breakeage of 4'' pipe 241P-067-P1349	3.9 10 <sup>-4</sup>	1.0 10 <sup>-4</sup>	2.0 10 <sup>-4</sup>	5.9 10 <sup>-5</sup>	7.3 10 <sup>-4</sup>	4.5 10 <sup>-4</sup>	2	5.9 10 <sup>-5</sup> - 7.3 10 <sup>-4</sup>
4	Rupture or disconnection between ammonia ship and unloading arm 241-ME1	5.8 10 <sup>-3</sup>	5.0 10 <sup>-3</sup>	4.8 10 <sup>-4</sup>	4.1 10 <sup>-6</sup>	1.0 10 <sup>-5</sup>	4.8 10 <sup>-4</sup>	4	4.1 10 <sup>-6</sup> - 5.8 10 <sup>-3</sup>
7	Rupture of 10'' pipe 241P-089-P1283	4.0 10 <sup>-4</sup>	2.0 10 <sup>-8</sup>	3.9 10 <sup>-8</sup>	7.0 10 <sup>-5</sup>	1.7 10 <sup>-4</sup>	-----	2	2.0 10 <sup>-8</sup> - 4.0 10 <sup>-4</sup>
7*	Rupture of a ship tank	5.7 10 <sup>-5</sup>	-----	2.3 10 <sup>-7</sup>	2.3 10 <sup>-6</sup>	4.9 10 <sup>-6</sup>	2.3 10 <sup>-7</sup>	-----	2.3 10 <sup>-7</sup> - 5.7 10 <sup>-5</sup>
9	Rupture of cryogenic tank 241-S1	Contained leak: 1.0 10 <sup>-6</sup> Uncontain leak: 4.0 10 <sup>-8</sup>	-----	5.0 10 <sup>-7</sup>	5.0 10 <sup>-8</sup>	5.0 10 <sup>-7</sup>	1.0 10 <sup>-8</sup>	4	1.0 10 <sup>-8</sup> - 1.0 10 <sup>-6</sup>
10	Rupture of 20'' pipe 241P-015-P1284	9.0 10 <sup>-5</sup>	1.0 10 <sup>-6</sup>	7.6 10 <sup>-6</sup>	8.8 10 <sup>-7</sup>	9.7 10 <sup>-7</sup>	1.0 10 <sup>-6</sup>	2	8.7 10 <sup>-7</sup> - 9.0 10 <sup>-5</sup>
14	Rupture of one of the ten pressurised tanks	2.5 10 <sup>-6</sup>	5.0 10 <sup>-7</sup>	1.6 10 <sup>-6</sup>	1.3 10 <sup>-5</sup>	2.0 10 <sup>-6</sup>	5.0 10 <sup>-7</sup>	3	5.0 10 <sup>-7</sup> - 1.3 10 <sup>-5</sup>
15	Rupture of 4'' pipe on the distribution line of tank 241-V1	2.3 10 <sup>-4</sup>	2.0 10 <sup>-5</sup>	6.0 10 <sup>-5</sup>	1.1 10 <sup>-5</sup>	4.9 10 <sup>-7</sup>	3.4 10 <sup>-8</sup>	2	3.4 10 <sup>-8</sup> - 2.3 10 <sup>-4</sup>
17	Rupture or disconnection between ammonia truck and unloading arm	3.7 10 <sup>-3</sup>	6.0 10 <sup>-5</sup>	4.7 10 <sup>-6</sup>	6.8 10 <sup>-5</sup>	1.0 10 <sup>-6</sup>	1.5 10 <sup>-7</sup>	1	1.5 10 <sup>-7</sup> - 3.7 10 <sup>-3</sup>
18	Catastrophic rupture of a truck tank	2.3 10 <sup>-7</sup>	1.2 10 <sup>-7</sup>	1.1 10 <sup>-8</sup>	7.4 10 <sup>-9</sup>	2.7 10 <sup>-8</sup>	1.5 10 <sup>-9</sup>	1-2	1.5 10 <sup>-9</sup> - 2.3 10 <sup>-7</sup>

• Grey tanned cells contain the lower assessments. Black tanned cells contain the upper assessments

**Tableau 2.10** – Deviation between frequencies of the top events of the common scenarios analyzed by the partners (events per year).



**Figure 2.8** – Variation of results for the consequence assessment of the reference scenarios.

From Table 2.10 and Figure 2.8, the results of the analysis show a great deviation between the assessed frequencies and consequences of risk scenarios. The spread in the results is due to uncertainty. However, this uncertainty will obviously be transferred to the final risk evaluation and affect the decision making, mainly land use planning, emergency planning and acceptability of risk. A risk scenario can be acceptable for a team and not acceptable for another.

Therefore, the need to analyze these uncertainties, take measures to reduce them, and inform the other stakeholders of risk-informed decisions about their existence and

implications is a principal obligation for the risk analysis community. The outcome of the ASSURANCE project should be seen in this context, i.e. as a contribution to the understanding of the uncertainties.

In the next section, the steps behind addressing parameter and completeness uncertainties will be presented.

## 2.5 Addressing uncertainty during risk assessment

The following provides an overview on how to address parameter and completeness uncertainties. Detailed guidance on how we treated these uncertainties in the INERIS risk analysis process appears in Chapters 3, 4, and 5.

### 2.5.1 Steps to address parameter uncertainty

Addressing this source of uncertainty is based on two main steps as presented in Figure 2.9 and detailed below:

1. uncertainty characterization: meaningfully representing the uncertain parameters using mathematical theories and based on the available data;
2. uncertainty propagation: propagating the mathematical representations generated from step-1 through the mathematical models in order to be able to represent parameter uncertainty in the outputs.

### 2.5.2 Addressing completeness uncertainty

Although the analysis of parameter uncertainty is fairly mature and is addressed adequately through the use of mathematical theories on the values of the parameters, the analysis of the completeness uncertainty cannot be handled in such a formal manner. As presented in Section 2.3.3, completeness uncertainty has two types: known and unknowns. The typical response to the known completeness uncertainty is to understand the importance of being up-to-date to introduce all risk contributors that can lead to major accidents. Chapter 3 addresses the known completeness uncertainty that faces today's risk assessment methodologies (omission of cyber-security related risks). If not, risk assessors should prove that omission of some risk contributors from the analysis would not affect the output prediction neither the decision making.

The true unknowns (i.e., those related to issues whose existence is not recognized) can be addressed analytically regarding the believe of who perform the analysis. However, in the interests of making defensible decisions, these unknowns can be addressed during the decision making by safety margins.

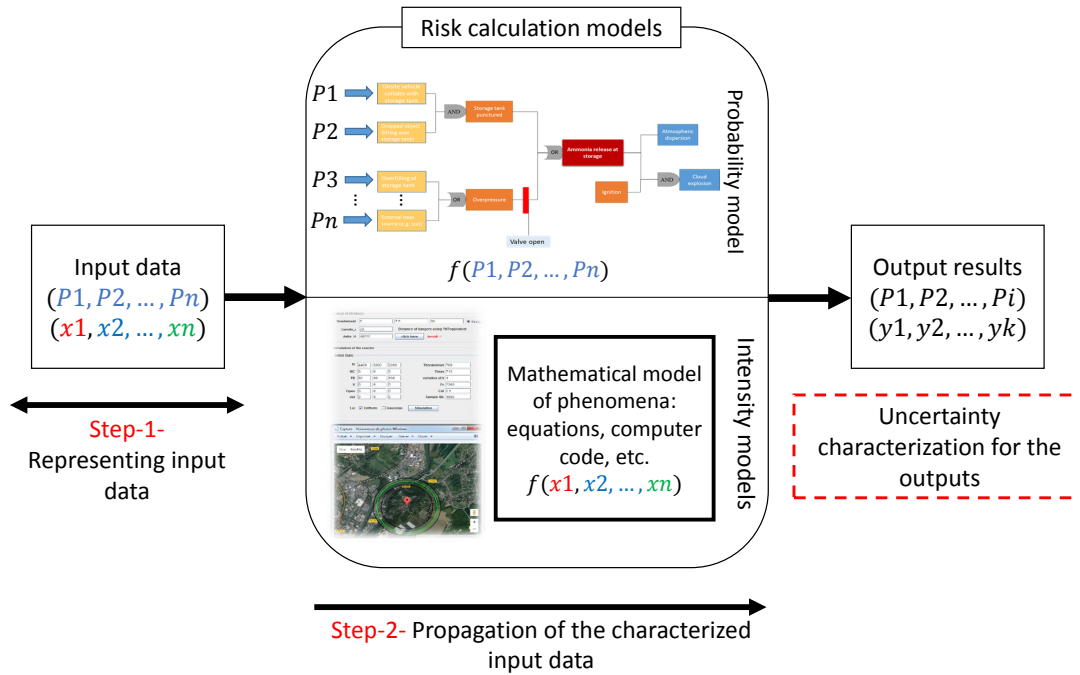


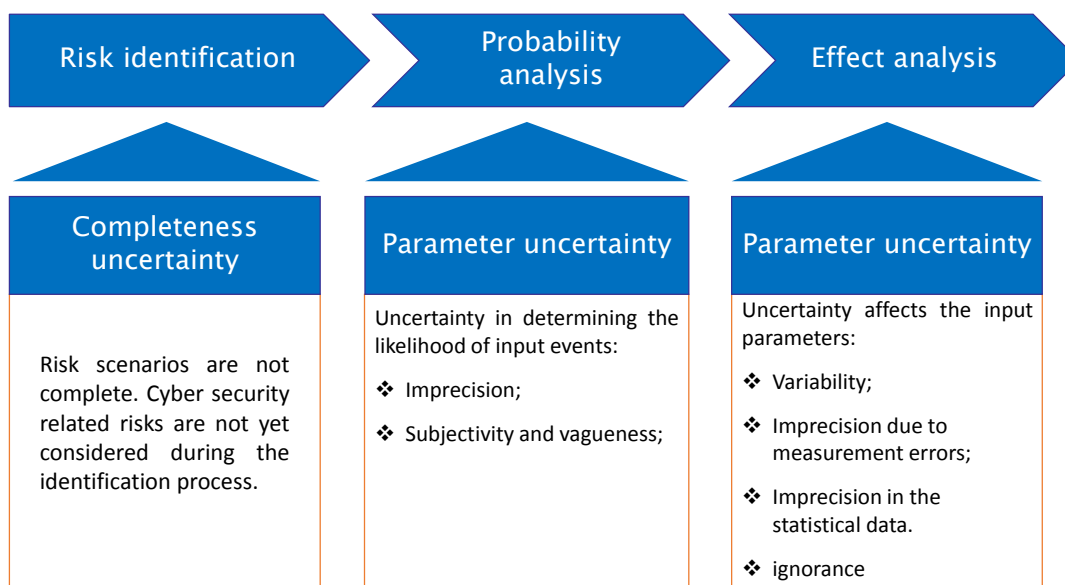
Figure 2.9 – Steps to address parameter uncertainty during risk assessment.

## 2.6 Where uncertainties affect the INERIS risk assessment process

In this section, we present the sources of uncertainties that affect each step of the INERIS risk analysis process as shown in Figure 2.10. Then we discuss the causes of each source of uncertainty at each step as follows:

- The interval semi-quantitative approach used by the INERIS for likelihood analysis suffers from some limitations due to parameter uncertainty. Section 2.6.1 details these limitations and highlights their impacts on the predictions of the likelihood analysis.
- Inability in determining precise values for risk models’ input parameters used for effect analysis may be faced due to parameter uncertainty. Some of these input parameters are random in nature and have different values, while information regarding other parameters is often incomplete, imprecise or vague. These different causes of parameter uncertainty should be treated separately with different representation theories. However, inappropriate representation of the available information may lead to under or overestimation of risk and subsequently to a bad decision as detailed in Section 2.6.2.
- The identification step is incomplete where only safety related scenarios caused by accidental events (component failures, human errors, etc.) are considered. This

no-consideration of cyber-security threats results in completeness uncertainty that should be addressed. Section 2.6.3 details this problem as well as the obstacles in dealing with it in order to introduce cyber-security related risks within industrial risk analysis.



**Figure 2.10** – Where uncertainties affect the risk assessment process.

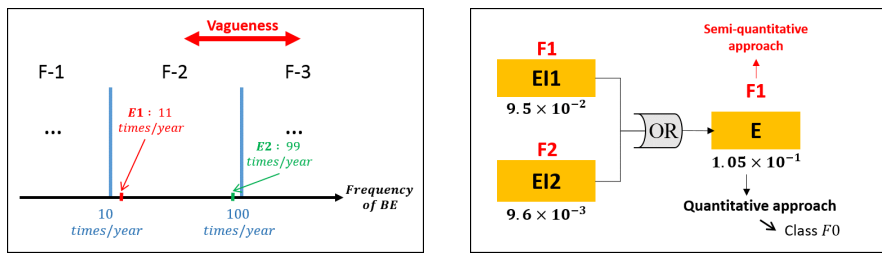
### 2.6.1 Uncertainty in likelihood analysis: drawbacks of the interval semi-quantitative approach

This section highlights the causes of parameter uncertainty in the likelihood analysis process. However, the interval semi-quantitative used by the INERIS has some limits that can lead to probability underestimation in some cases and need to be addressed. These limitations are presented by converting the statistical data into interval classes:

- *Limit-1*: The discreteness of the frequency classes makes the conditions on the border between two intervals not well defined. Vagueness on the extent of half the range of the interval to which category it belongs is presented. The same class is given to different frequencies even if the difference between them is too remarkable (see Figure 2.11(a), the same class  $F - 1$  is given to events E1 and E2 where their frequencies are 11 and 99 per year respectively);
- *Limit-2*: The interval representation of frequency classes can lead to probability underestimation. Figure 2.11(b) presents an example of an OR gate of two inputs  $EI1$  and  $EI2$  with quantitative frequencies equal  $9.5 \times 10^{-2}$  and  $9.6 \times 10^{-2}$  respectively. Suppose that these quantitative information are certain, the quantitative

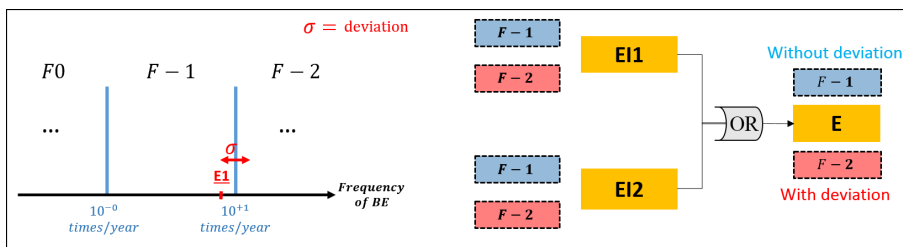
output frequency of the OR gate is equal to  $1.05 \times 10^{-1}$ . The translation of this quantitative output frequency using the semi-quantitative scale results in a class  $F0$  (see Equation 2.2 for how to solve the OR gate). While a frequency class equals  $F1$  is obtained based on the interval semi-quantitative approach;

- *Limit-3*: If we were wrong in determining the statistical data, i.e. an error factor may affect the statistical value. This error or deviation can lead to a different class and thus to a different result. Figure 2.11(c) shows an example of an OR gate with two input events EI1 and EI2. The occurrence frequencies of the both input events are equal to 9 times/year. Suppose we have an error factor due to the lack of information or a measurement error, and the frequencies can be higher by a factor of 2 times/year (e.g. 11 times/year). This deviation changes the class of events E1 and E2 from  $F - 1$  to  $F - 2$ . The output of the OR gate is affected by this error where its class is changed. Thus, a small deviation can lead to a great change in the output.



(a) The same class is given to two different values.

(b) Deviation due to imprecision.



(c) Probability underestimation.

**Figure 2.11** – Limitations presented in the traditional semi-quantitative approach.

### 2.6.2 Uncertainty in effect analysis

We shall now introduce the problem that risk analysts face when analyzing the effect of a risk scenario. Complex mathematical models that depend on an important number of input parameters are used in order to calculate the effect of an accident. However, input

parameters for these models are fraught with uncertainties. These uncertainties might be of different types (aleatoric, epistemic or mix of both types) depending on the types of these input parameters and the sources of available data. The natural variability of some of these parameters and the lack of knowledge and empirical data in the determination of others make it difficult to assess the degree of exposure of certain effects. The particular context of uncertainties in effect analysis is shown in Figure. 2.12.

This figure shows a representation of mathematical models utilized for effect analysis. An effect model is made up of three: inputs, the model equation (the mathematical model used to represent the dangerous phenomenon being studied), and outputs. The challenge is to address the issues relating to the representation of parameter-related information (see figure. 2.12). However, this information may not necessarily be accurate. It may be based on experimental data and measurement, and therefore be known statistically, or based on expert opinion and therefore be known with a certain degree of inaccuracy.

One method commonly used to address these uncertainties is the probabilistic method, which applies probability theory and relies on a statistical representation of available information. This approach aims to represent uncertain parameters using probability distributions. This means that information on model's parameters are generally supposed to be random in nature (variability). Then these distributions are propagated through the model by applying the Monte Carlo technique in order to get probability distributions for the outputs. But, in order to build probability distributions, statistical data are required and these are often missing. This means that assumptions must be made in order to build the distributions. Indeed, in a risk analysis context for industries classified SEVESO, the information available concerning certain parameters is often imprecise or incomplete due to lack of data. The calibration of probability distribution by making assumptions on this type of knowledge becomes subjective and arbitrary.

However, as we are going to prove in this work that the use of probabilistic method to address uncertainty is not always the right choice, and choosing of the appropriate method depends on the context and the type of uncertainty that affect the parameter being studied. The use of an inappropriate approach in an inappropriate place may lead to under or overestimation of risk and subsequently to a bad decision.

Thus, Our challenge is how to represent uncertainty relating to parameters when there is insufficient information available for statistical identification (when there is epistemic uncertainty). Incorporating uncertainty into effect analysis is an important issue of widespread interest.

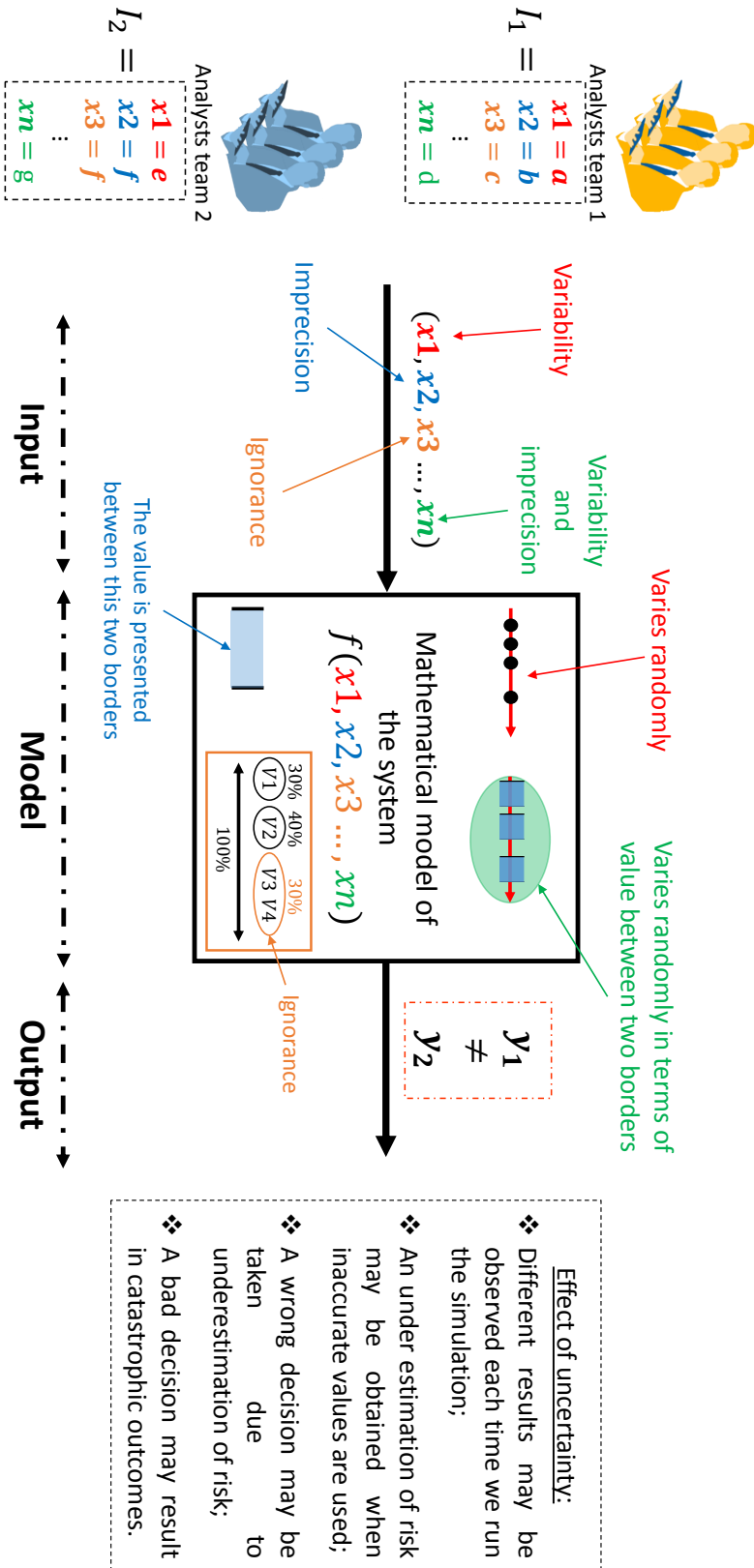


Figure 2.12 – Parameter uncertainty in effect analysis.

### 2.6.3 Completeness uncertainty in the risk identification process: No-considering of cyber-security threats

In this section we present the known cause of completeness uncertainty that makes the identification step of the risk analysis process incomplete. New type of initiating events related to cyber-security are generated due to the development of industries. Thus, the analysis process should be updated to meet the development of critical industries and to assure an appropriate risk analysis and effective decision-making.

Traditional industries were based on mechanical devices and closed systems [101]. Only safety related risks generated from accidental component failures and human errors need to be addressed during risk analysis of these industries. However, today, industries are influenced by the development of digital technology related to instrumentation and industrial automation (IA). Supervisory Control And Data Acquisition (SCADA) systems are introduced to monitor and control equipment that deals with critical and time-sensitive materials or events. The shift from analog equipment towards technologies has a number of benefits concerning production, but it also presents challenges [145]. This introduction of automation technology increases the degree of complexity and communication among systems. The use of Internet for connecting, remote controlling and supervising systems and facilities has generated a new type of risk causes that related to cyber-security. These systems and facilities have become more vulnerable to external cyber attacks. These new security threats can affect the safety of systems and their surrounding environments in terms of people, property, etc. ([89]; [99]). Then, introducing cyber-security related risks that may lead to major accidents into industrial risk assessment is an important need.

The differences and similarities between safety and security are studied by many authors ([101]; [62]). In general, safety is associated with accidental risks caused by component failures, human errors or any non-deliberate source of hazard, while security is related to deliberate risks originating from malicious attacks which can be accomplished physically (which are excluded in this study) or by cyber means. In addition, causes of accidents related to safety are internal and considered to be rare events with low frequency. Causes of security accidents can be internal or external (attacks via insider agents or outsiders) and are classified as common events. Table 2.11 shows the difference between safety and security regarding different criteria.

Until today, industrial risk analysis does not take into consideration the cyber-security related risks that can affect the safety of the system and lead to major accidents. Systems are designed to be reliable and safe, rather than cyber-secure. These various types of critical infrastructure have unique requirements for both functionality and up-time, but they were not built from the ground up to be security focused. At the time that many of



these critical systems were being built, cyber terrorism, corporate cyber espionage, and global interconnectivity simply were not issues of the day. Today is a far different story. In recent years, there has been an increasing number of cyber attacks that target critical facilities (e.g., Stuxnet in 2010 and Flame in 2012). According to Dell's annual threat report [42], cyber attacks against SCADA systems doubled in 2014. Dell SonicWALL saw global SCADA attacks increase against its customer base from 91,676 in January 2012 to 163,228 in January 2013, and 675,186 in January 2014. Many authors have studied the potential impact of security related threats on the safety of critical facilities and highlight the importance of analyzing safety and security risks together [99]. Theft, corruption, or outright destruction of these systems can have a crippling effect on human safety, financial stability, and standard of living. Thus, concerns about approaches for industrial risk analysis that consider safety and security together are a primary need.

	Safety	Security
The nature of risk	Technical and controllable problem, caused by any accidental sources of hazard	Human strategic aggressor, caused from malicious cyber-attacks
Type of intent	Internal, non-deliberate	Intentional, malicious, internal or external
Likelihood of occurrence	Rare events with low frequency of occurrence	Classified as common events with high likelihood of occurrence
Definition of risk (Section 6.2.1 and 6.2.2)	$R_{security} = (tv, Ptv, Xtv)$	$R_{safety} = (Se, Pe, XS)$

**Tableau 2.11** – Difference between safety and security.

## 2.7 Conclusion

This Chapter established the context of the study by presenting the different steps to conduct a systematic industrial risk analysis. The definition, sources, types and causes of uncertainty in risk analysis are presented. The main outcome of this Chapter is the

definition of the research problem, and highlighting the obstacles that face risk assessors when performing risk analysis for today's critical facilities.

First we present the steps of performing a risk analysis for critical facilities based on the INERIS approach. The identification and modeling of risk scenarios using the bow-tie analysis is detailed. The INERIS semi-quantitative approach to conduct likelihood analysis is presented. The characterization of input data and the propagation through the bow-tie model in order to calculate the likelihood of outcomes are detailed. Analyzing the effects of risk scenarios is also detailed. The decision making process to see if the identified risk scenarios are acceptable or not after being analyzed in terms of likelihood and effects is presented.

Uncertainty represents a major problem when analyzing major risks. Uncertainty analysis is the measure of the "goodness" of an estimate. The different sources of uncertainty with their causes are defined:

- Parameter uncertainty include not only imprecision due to small samples of recorded data, but also uncertainties in experts' judgments of parameter values when there are not recorded data;
- Model uncertainty is the indefiniteness in the model's comprehensiveness (i.e., does the model considers all variables and the relations between these variables which can significantly affect the results). Model uncertainty is similar in nature to the completeness uncertainty but occur at the modeling level of consequences and not at the initial stage of the analysis.
- Completeness uncertainty is the uncertainty as to whether all the significant risk contributors have been considered in the analysis. There are two types of completeness uncertainty: (1) known uncertainty (risk contributors are omitted intentionally due to the complexity of introducing them, or these contributors do not have any significance on the prediction), and (2) unknown completeness uncertainty (risk contributors are not known or unintentionally missed);

The different sources of uncertainty that affect the INERIS risk analysis process are identified. Parameter and completeness uncertainties affect the analysis process at its different steps. Probability and effect analyses suffer from parameter uncertainty. Completeness uncertainty affects the risk identification step. As we are going to prove in this work, inappropriate treatment of parameter uncertainty may have important effects on the results and lead to under or overestimation of risk and subsequently to a bad decision. Completeness uncertainty is faced in analyzing risks for critical facilities that use technologies to control and monitor their systems.



# 3

## Literature review

**Summary:** This chapter is divided in two parts. The first part of this Chapter presents a state of art on the different theories and approaches to treat parameter uncertainty as presented in Section 3.1. The most commonly used approaches to quantify uncertainty in risk analysis: interval analysis, fuzzy theory, probability theory, evidence theory, and the mixed probabilistic-fuzzy approach are presented. Characterizing parameter uncertainty in model inputs and propagating these inputs through the model based on these approaches are detailed.

The second part presents an introduction to industrial control systems, cyber-security and cyber-security analysis in critical facilities (Section 3.3). This introduction is followed by an overview on existing approaches to analyze security risks alone and within safety for critical facilities as presented in Sections 3.5 and 3.6, respectively. Advantages and limits of these approaches are also discussed.

### Summary

---

<b>3.1</b>	<b>State of art on approaches for parameter uncertainty analysis . . . . .</b>	<b>53</b>
3.1.1	Interval analysis . . . . .	53
3.1.2	Probabilistic approach . . . . .	54
3.1.3	Fuzzy approach . . . . .	55
3.1.4	Evidence approach . . . . .	58
<b>3.2</b>	<b>Industrial Automation and Control System - IACS . . . . .</b>	<b>63</b>
<b>3.3</b>	<b>Cyber-security for industrial control systems . . . . .</b>	<b>64</b>
<b>3.4</b>	<b>The security risk analysis process for industrial control systems . . .</b>	<b>66</b>

<b>3.5</b>	<b>A review of cyber-security risk analysis approaches for Industrial Control Systems . . . . .</b>	<b>66</b>
3.5.1	Attack-tree-based approaches . . . . .	67
3.5.2	Security Modeling with BDMP: From Theory to Implementation [126] . .	70
3.5.3	Evaluating the risk of cyber attacks on SCADA systems via Petri net analysis with application to hazardous liquid loading operations [80] . . .	71
3.5.4	Network Vulnerability Assessment using Bayesian Networks [108] . . . . .	72
3.5.5	Discussion . . . . .	72
<b>3.6</b>	<b>Existing Approaches that combine safety and security for industrial control systems . . . . .</b>	<b>74</b>
3.6.1	Integrating cyber attacks within fault trees [68] . . . . .	74
3.6.2	Modeling safety and security inter-dependencies with BDMP (Boolean logic driven Markov processes) [125] . . . . .	75
3.6.3	Studying interrelationships of safety and security for software assurance in cyber-physical systems: Approach based on Bayesian belief networks [98]	75
3.6.4	Modeling and analysis of safety-critical cyber physical systems using state/event fault trees [134] . . . . .	76
3.6.5	Discussion . . . . .	76
<b>3.7</b>	<b>Conclusion . . . . .</b>	<b>78</b>

---

## 3.1 State of art on approaches for parameter uncertainty analysis

Parameter uncertainty analysis is a process of representing, propagating of uncertain knowledge to provide uncertainty representation for the output results. In risk analysis, uncertainty in input parameters is described using mathematical theories in term of intervals or distribution functions. Uncertainty in parameters is then propagated through the model using mathematical rules or simulation such as interval roles or MC simulation.

Uncertainty analysis has been performed as part of risk analysis in various studies in different fields and applications: characterization of uncertainty in a chemical plant based on an analytical network process [103]; use of an integrated approach to modeling uncertainty in risk assessment [14]; study of the uncertainty related to injury exposure risks ([113]; [154]); ecological risks ([61]; [81]); chemical consequence risks [130]; software development risks [109]; transportation risks [97]; dam safety risks [30]; human health risks ([95]; [103]); microbial risks [39]; bridge risks [156]; risks relating to construction projects ([165]; [155]; [21]; [128]); and supply chain risk management ([91]; [137]; [75]).

The rest of this section provides an overview of how parameter uncertainty can be analyzed using interval analysis, the probabilistic approach, fuzzy theory and evidence theory. However, these theories are not compared in this chapter. In Chapter 5, we apply them on a real case study and compare their results.

### 3.1.1 Interval analysis

Interval analysis is the simplest way to represent uncertainty as it requires the smallest amount of information: the lower and upper or minimum and maximum values are the only information needed to build the interval and represent the uncertain parameter of interest [116].

Propagating uncertain parameters represented by intervals is done by answering the following question: what are the smallest and the largest possible values that can be obtained when applying mathematical operations to the interval? In answer to this question, a number of authors ([151]; [10]; [11]) have developed what is called interval arithmetic to perform the basic mathematical operations on interval analysis. For instance, let  $[a] = [\underline{a}, \bar{a}]$ ,  $b = [\underline{b}, \bar{b}]$  be real intervals while  $o$  is one of the basic mathematical operations (addition, subtraction, multiplication or division), in other words  $o \in \{+, -, \times, /\}$ . This makes it possible to define the corresponding operations for intervals  $[a]$  and  $[b]$  by

$[a]o[b] = \{aob \mid a \in [a], b \in [b]\}$ , where  $0 \notin b$  can be shown as follows:

$$\text{Addition: } [a] + [b] = [\underline{a} + \underline{b}, \bar{a} + \bar{b}].$$

$$\text{Subtraction: } [a] - [b] = [\underline{a} - \bar{b}, \bar{a} - \underline{b}].$$

$$\text{Multiplication: } [a] \times [b] = [\min\{\underline{a}\underline{b}, \underline{a}\bar{b}, \bar{a}\underline{b}, \bar{a}\bar{b}\}, \max\{\underline{a}\underline{b}, \underline{a}\bar{b}, \bar{a}\underline{b}, \bar{a}\bar{b}\}].$$

$$\text{Division: } [a]/[b] = [\min\{\underline{a}/\underline{b}, \underline{a}/\bar{b}, \bar{a}/\underline{b}, \bar{a}/\bar{b}\}, \max\{\underline{a}/\underline{b}, \underline{a}/\bar{b}, \bar{a}/\underline{b}, \bar{a}/\bar{b}\}].$$

For example, let us suppose that the uncertainty associated with the two parameters X and Y is represented by the intervals  $[0.1, 0.2]$  and  $[0.15, 0.35]$ , respectively. Using the equations given above, the output of  $Z = f(X + Y)$  is the interval  $Z = [0.1 + 0.15, 0.2 + 0.35] = [0.25, 0.55]$ .

As explained above, we can see that interval analysis is very simple to understand offering a straightforward means of representing and propagating uncertainty. It should be noted that for interval analysis, there is no structured shape of likelihood. Thus, a decision is based on either the minimum or maximum value since no distribution shape is provided. This means that the decision may be pessimistic and overestimate the risk.

### 3.1.2 Probabilistic approach

The probabilistic approach is the most common approach used to represent parameter uncertainty ([13]; [52]). It enables uncertainty to be quantified, mainly by using distributions of random variables instead of fixed values. A distribution describes the range of possible values (e.g. wind speed varies between 1 and 3 mph), and shows which values within the range are most likely ([15]; [26]).

Distribution probability covers discrete and continuous cases. In discrete cases, a discrete probability distribution function  $d(x) : \omega \rightarrow [0, 1]$  exists where  $\sum d(x) = 1$  ( $\omega$  is the sample space containing all the possible values for the random variable). In continuous cases, on the other hand, a Probability Density Function (PDF)  $p(x)$  exists such that  $\int_{\omega} p(x)dx = 1$  [171].

The probability of any subset  $S$  of  $\omega$  is:

$$P(S) = \sum_{x \in S} P(x); \text{ where } P(x) = d(x) \text{ in the discrete case.} \quad (3.1)$$

$$p(S) = \int_S p(x)dx; \text{ in the continuous case.} \quad (3.2)$$

In continuous cases, calculating the probability that a specific element  $x$  will occur is based on the Cumulative Distribution Function (CDF). This represents the probability that a value smaller than or equal to  $x$  will occur.

$$P(x) = P([-\infty, x]) = P(X \leq x) = \int_{-\infty}^x p(t)dt, \forall x \in \omega. \quad (3.3)$$

Additional information on probability theory can be found in the work of various authors ([149]; [121]; [41]; [129]).

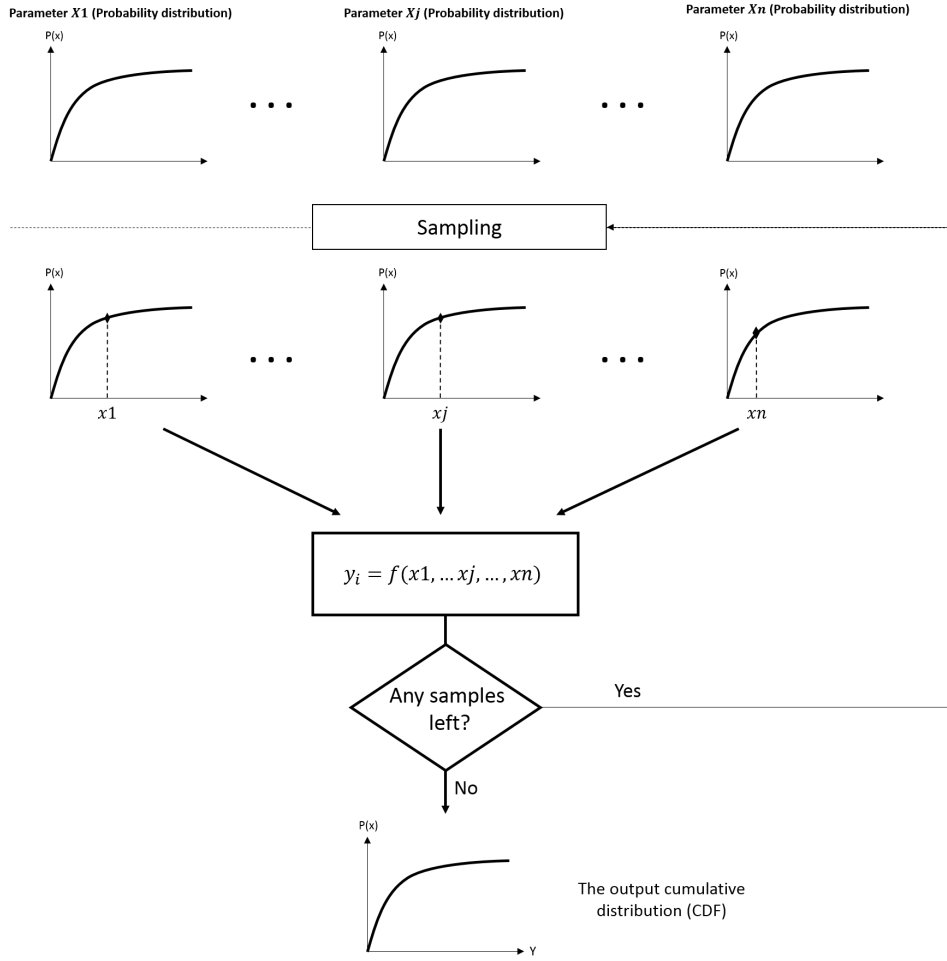
For propagation, Monte Carlo (MC) is the most widely used method to propagate uncertainty when probability distributions are attached to uncertain variables ([37]; [20]). For instance, let us consider a model whose output is a function  $Z = f(Y) = f(X_1, X_2, \dots, X_n)$  of  $n$  uncertain variables which are represented by probability distributions. The propagation of uncertainty calculated using the Monte Carlo method is based on the following main steps (see figure 3.1):

1. set  $i = 1$  and  $N =$  number of samples;
2. take sample  $i$ ,  $y_i = (x_1, x_2, \dots, x_n)$  where  $x_i, i = 0, \dots, n$  are randomly generated from each distribution  $X_i$ ;
3. calculate  $z_i = f(y_i) = f(x_1, x_2, \dots, x_n)$ ;
4. go back to step 2 if any samples are left, otherwise go to 5;
5. generate the *pdf* of results( $Z$ ) by butting all the  $z_i$  on a histogram and the *CDF* of  $Z$  as  $F(z) = \frac{1}{N} \sum_{i=1}^N H_i$  where  $H_i = 1$  if  $z_i < z$  and 0 elsewhere.

### 3.1.3 Fuzzy approach

By definition, fuzzy numbers represent a family of nested sets or so-called  $\alpha - cuts$  [164]. A fuzzy variable is associated with a possibility distribution or membership function ( $\mu$ ) in the same manner as a random variable is associated with a probability distribution ([131]; [96]). [4] and [1] provide a detailed explanation of the difference between probability and possibility. However, considering a fuzzy set  $F$  on the range  $X$  and a given  $x \in X$ , the membership function value  $\mu_{F,X}(x)$  represents the degree of compatibility of the value  $x$  with the concept expressed by  $F$  [119]. Furthermore, fuzzy numbers represent an important alternative for information representation, especially when this representation is given by experts and is hence by nature qualitative ([132]; [88]; [111]; [56]; [55]; [43]; [136]).



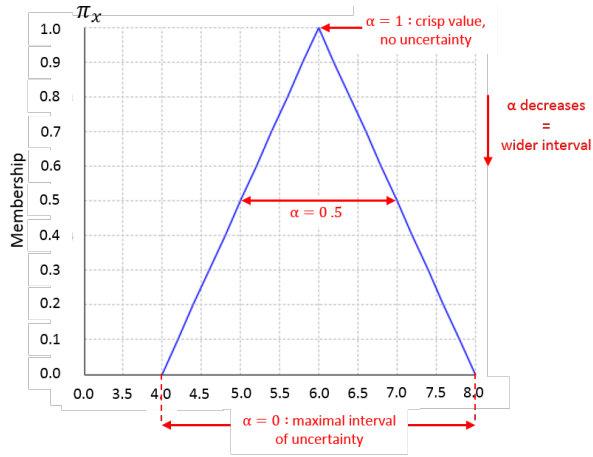


**Figure 3.1** – Monte Carlo propagation of probability distributions.

To explore this further, let us take an example where the only information about an uncertain parameter  $x$  is that it can take a value in the interval  $[4, 8]$ , where 6 represents the most likely value to be the true value of  $x$ . This information can be represented using a triangular fuzzy number (TFN) on the interval  $[4, 8]$  with a possibility degree equal to 1 at 6 (see figure 3.2). The membership function for a TFN is presented as follows:

$$\mu_{(A)}(x) = \begin{cases} \frac{x - a_1}{a_2 - a_1} & a_1 \leq x \leq a_2 \\ \frac{a_3 - x}{a_3 - a_2} & a_2 \leq x \leq a_3 \end{cases} \quad (3.4)$$

where  $(a_1, a_2, a_3)$  are the lower, most likely and upper values respectively. Based on this equation (eq. 3.4), we can calculate any  $\alpha$  – cut interval  $(A_\alpha)$  at any degree  $(\alpha)$  between 0 and 1.  $A_\alpha^X = \{x : \pi_X(x) \geq \alpha\}$  is the set of  $x$  values for which the possibility function is greater than or equal to  $\alpha$ . For example,  $A_{0.5}^X = [5, 7]$ , which means that there is a 50% possibility that this interval contains the true value.



**Figure 3.2** – Triangular possibility distribution on the interval  $[4, 8]$ , where the most possible value equals 6.

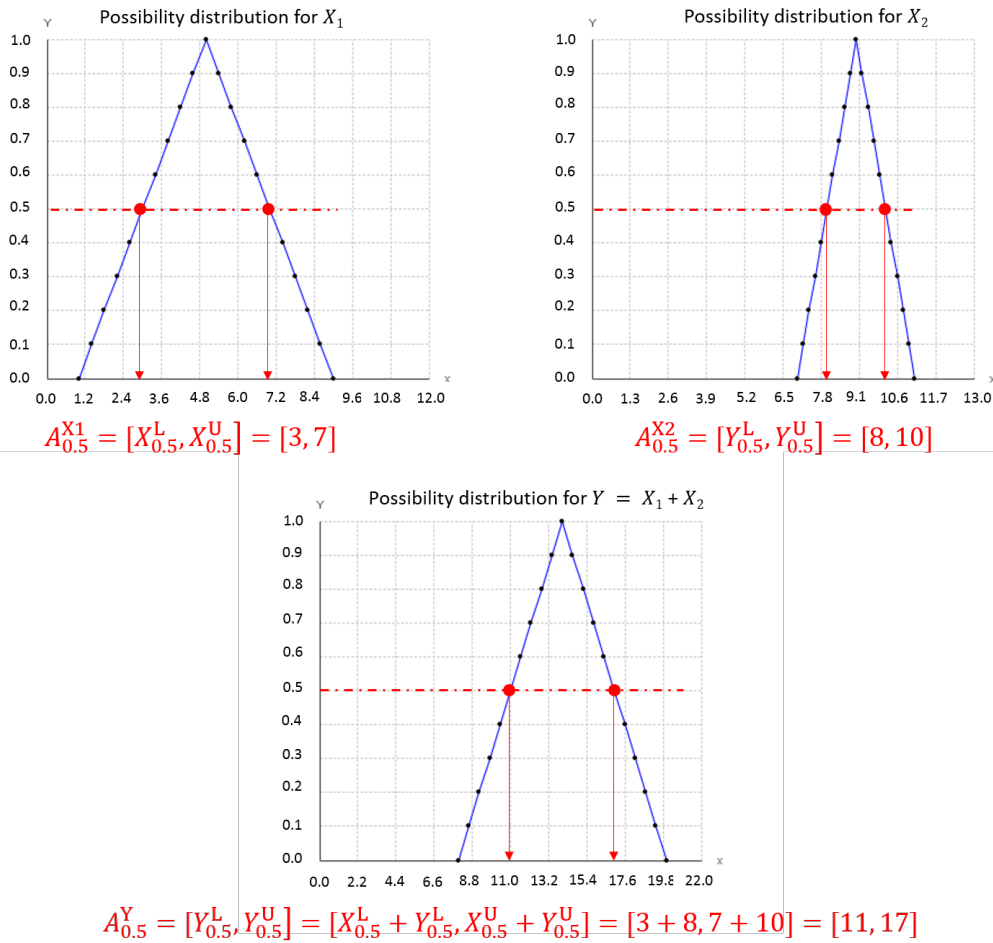
With respect to uncertainty propagation, let a model  $Y = f(X_1, X_2, \dots, X_n)$ , where all the inputs are affected by uncertainty and represented by fuzzy sets with their possibility distributions  $\pi_{X_1}, \pi_{X_2}, \dots, \pi_{X_n}$ . The propagation of uncertainty can be performed using fuzzy calculus based on interval analysis with different  $\alpha$  – cuts ([161]; [58]; [48]).

The different steps for propagating uncertainty using alpha-cuts or fuzzy calculus are presented as follows:

1. Set  $\alpha$  – cut = 0;
2. Select the  $\alpha$  – cut intervals  $(A_{X_1}^\alpha, A_{X_2}^\alpha, \dots, A_{X_n}^\alpha)$  from the fuzzy inputs. An interval is obtained from each input parameter.
3. Calculate the smallest (lower) and largest (upper) values of the result  $Y$  at the set  $\alpha$  – cut by performing interval analysis to obtain  $A_Y^\alpha = f(A_{X_1}^\alpha, A_{X_2}^\alpha, \dots, A_{X_n}^\alpha)$ .
4. If  $\alpha = 1$ , then the propagation is finished and the possibility distribution is calculated (a set of nested  $\alpha$  – cuts representing the fuzzy output is obtained). If  $\alpha < 1$ , then set  $\alpha = \alpha + D\alpha$  ( $D\alpha$  is determined step-wise as  $[0 : \frac{1}{q} : 1]$ , where  $q \in \mathbb{N}$ ) and return to step 2.

To explore this further, let us take a simple example of  $Y = X_1 + X_2$ , where  $X$  and  $Y$  are uncertain and represented by triangular possibility distributions  $\pi_{X_1}(x_1) = (1, 5, 9)$  and  $\pi_{X_2}(x_2) = (7, 9, 11)$  (see figure 3.3). The bottom of figure 3.3 shows the triangular distribution  $\pi_Y$  of the output obtained by fuzzy-interval analysis and with  $D\alpha = 0.1, q = 10$ . For instance, the output interval at  $\alpha = 0.5$  is:  $A_Y^{0.5} = A_{X_1}^{0.5} + A_{X_2}^{0.5} = [A_{X_1L}^{0.5} + A_{X_2L}^{0.5}, A_{X_1U}^{0.5} + A_{X_2U}^{0.5}] = [3 + 8, 7 + 10] = [11, 17]$ . See figure 3.3.

It should be noted that the fuzzy approach assumes that the input parameters are totally dependent [124].



**Figure 3.3** – The triangular distributions of  $X_1$ ,  $X_2$  and  $Y = X_1 + X_2$  are shown in the top right, top left and bottom of the figure respectively.

### 3.1.4 Evidence approach

Evidence or Dempster-Shafer theory ([143]; [147]; [44]; [73]; [160]) provides an alternative way of representing uncertainty but with less restrictive statements about likelihood than probability theory [78]. Evidence theory is an extension of probability theory which allows us to quantify the probability of intervals instead of precise values when input variables are uncertain. These intervals are called focal elements and their probabilities are called basic probability assignments (bpa). In summary, an application of evidence theory involves the specification of a triple  $(\rho, S, m)$ , where

- $\rho$  is the set of everything that could occur;
- $S$  is a collection of subsets of  $\rho$ ;

—  $m$  is the bpa associated with each subset  $A_i \in S$ , where  $\sum_{A_i \in S} m = 1$ .

Evidence theory is characterized by two measurement functions called belief and plausibility, which are defined from the mass distribution  $m$  as follows:

$$Bel(A) = \sum_{D \subseteq A} m(D)$$

$$Pl(A) = \sum_{D \cap A \neq \emptyset} m(D)$$

Moreover, as in probability theory, a CDF can be used to provide summaries of the information contained in a probability space. Similarly, cumulative belief and plausibility functions can be constructed and used to summarize the information provided by an evidence space. Cumulative belief and plausibility (CBF and CPF) are defined by sets of points as follows:

$$CBF = \{[v, Bel(\rho_v)], v \in \rho\} \quad (3.5)$$

$$CPF = \{[v, Pl(\rho_v)], v \in \rho\} \quad (3.6)$$

where  $\rho_v$  is defined as  $\{x : x \in \rho \text{ and } x < v\}$ . Therefore  $Bel(\rho_v)$  and  $Pl(\rho_v)$  are respectively the belief and plausibility that a value smaller than or equal to  $v$  will occur.

To explain the propagation of uncertainty using this approach, let us take the same model  $Y = f(X) = f(X_1, X_2, \dots, X_n)$  but with  $n$  variables described here by evidence spaces. Each one of the input variables is characterized by its discrete set of focal elements and corresponding probability masses ( $X_{ij} = \{[x_{ij}^L, x_{ij}^U], m_i\}$   $i = 1 \dots n$  and  $j = 1 \dots q_i$ , where  $n$  is the number of input parameters and  $q_i$  is the number of focal elements of the parameter  $X_i$ ). To explore this approach further, two different propagation methods for independent and dependent input parameters are presented in the next two sections (3.1.4.1 and 3.1.4.2). However, Section 3.1.4.3 discusses why evidence theory is the right method to treat ignorance.

#### 3.1.4.1 Independent input parameters

This method assumes that there is total independence between the input parameters. The body of evidence of the output is computed by constructing the Cartesian product, where the probability masses (*bpas*) are obtained by the traditional multiplication of the input masses of the uncertain inputs.

For example, let  $X$  be represented by the body of evidence as follows:

$Z = X + Y$	$[1, 3] : 0.2$	$[3, 5] : 0.5$	$[4, 6] : 0.3$
$[1, 4] : 0.4$	$[2, 7] : 0.08$	$[4, 9] : 0.2$	$[5, 10] : 0.12$
$[2, 6] : 0.1$	$[3, 9] : 0.02$	$[5, 11] : 0.05$	$[6, 12] : 0.03$
$[4, 8] : 0.5$	$[5, 11] : 0.1$	$[7, 13] : 0.25$	$[8, 14] : 0.15$

**Tableau 3.1** – Body of evidence for the output  $Z = X + Y$  obtained by means of the Cartesian product.

$$X = \{[1, 3] : 0.2; [3, 5] : 0.5; [4, 6] : 0.3\}$$

and  $Y$  by

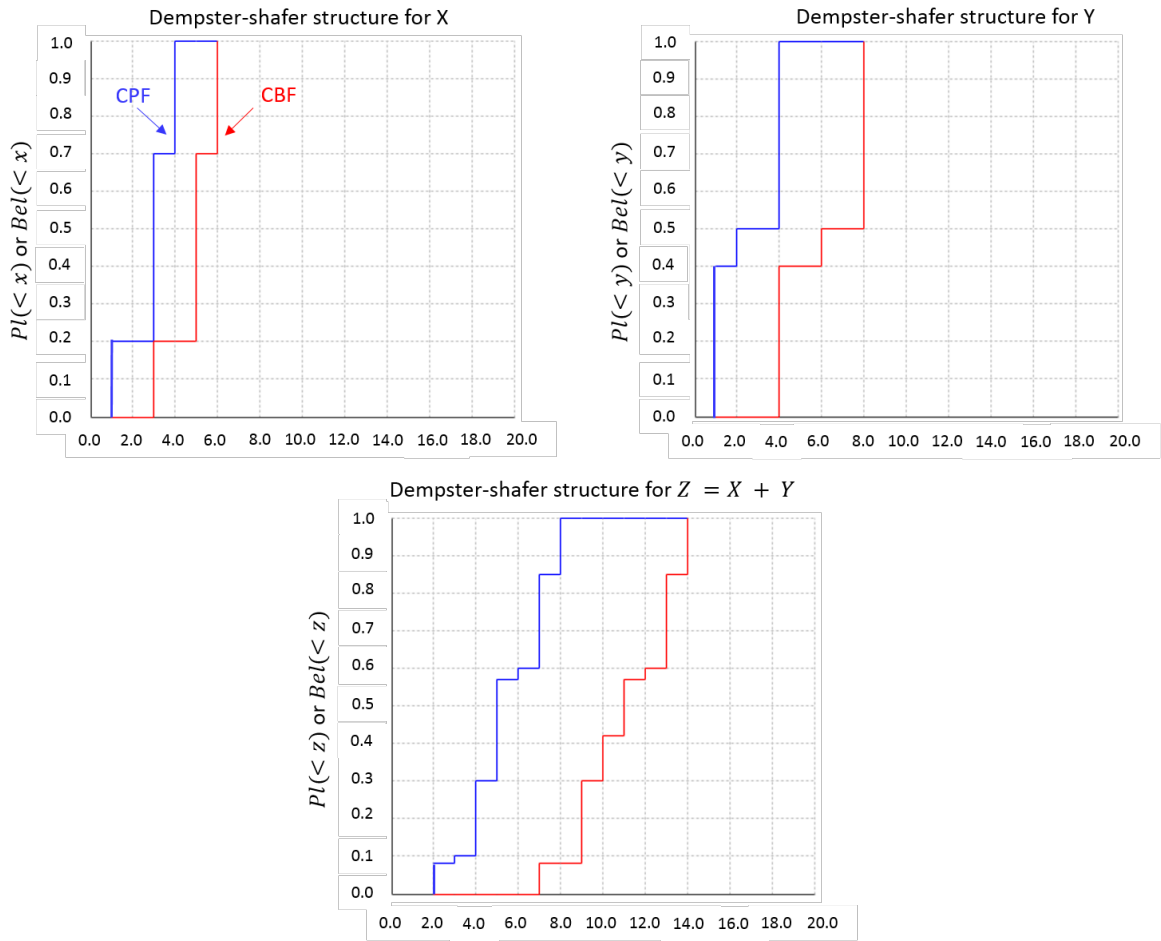
$$Y = \{[1, 4] : 0.4; [2, 6] : 0.1; [4, 8] : 0.5\}$$

The result of  $Z = X + Y$  is calculated and depicted in table 3.1 and figure 3.4 respectively.

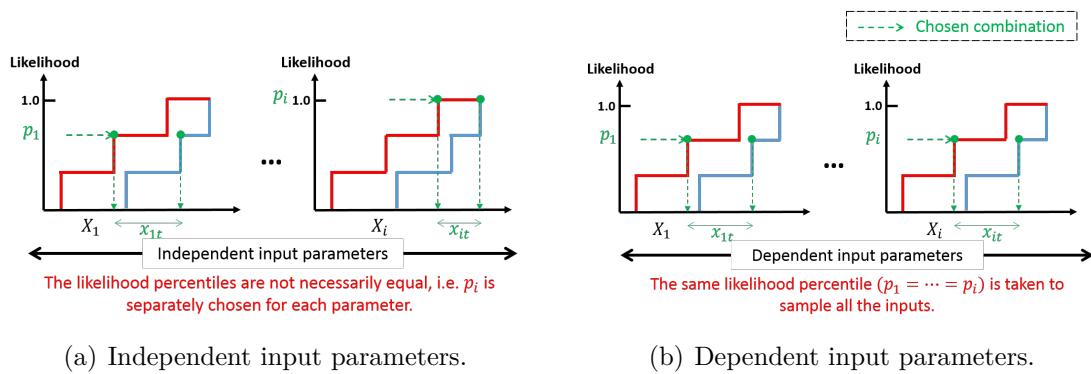
### 3.1.4.2 Dependent input parameters

In this case, the combinations are taken differently than in the case of independence. Combinations at the same level of likelihood are chosen (see figure 3.5, right) instead of all the possible combinations (figure 3.5 left). The CBFs and CPFs of the input parameters are used to generate these combinations in order to calculate the CBF and CPF of the output. The main calculation steps are listed below:

1. based on the CBFs and CPFs of the inputs, determine the vector of all the different percentiles. The different percentiles are the likelihood percentiles each time the CBF or CPF is changed. For the example taken in section 3.1.4.1, and based on the inputs  $X$  and  $Y$ , the vector of different percentiles is  $v = \{0.2, 0.4, 0.5, 0.7, 1.0\}$ .
2. for each percentile from the vector derived in step 1, an interval ( $x_{it}$ ,  $t$  is the chosen percentile) is derived from the cumulative distributions of each input parameter  $X_i$  (see figure 3.5, right);
3. calculate the output interval  $y_t = f(x_{it})$ . This interval represents the cumulative output for the percentile  $t$  (the minimum and maximum of  $y_t$  are the values of  $y$  where CBF and CPF equal  $t$ );
4. return to step 2 if any percentile is left, otherwise go to 5;
5. generate the CBF and CPF of the output.

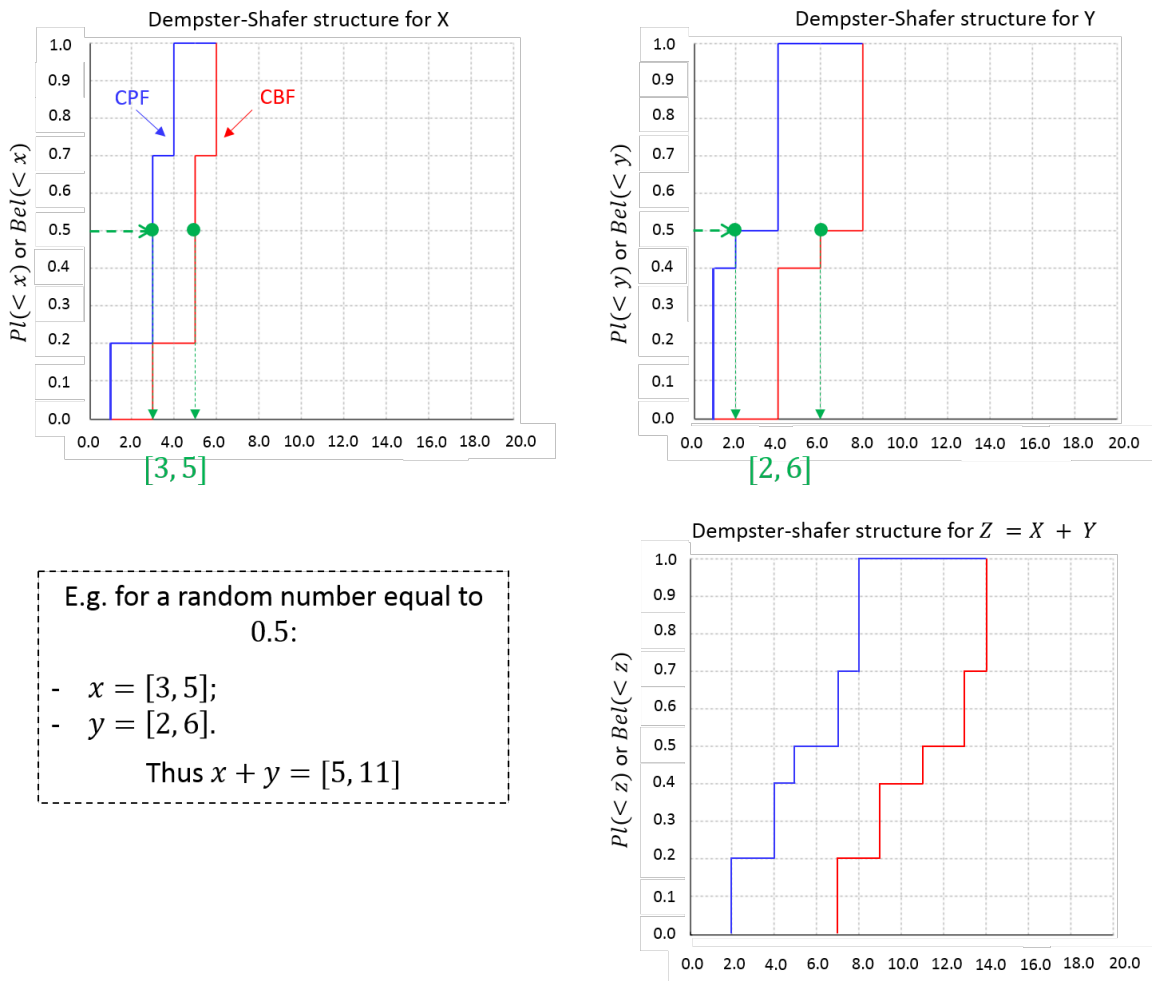


**Figure 3.4** – Upper and lower CDFs of  $X$ , i.e.  $Pl_X(-\infty, x)$  and  $Bel_X(-\infty, x)$  in the top left; upper and lower CDFs of  $Y$ , i.e.  $Pl_Y(-\infty, y)$  and  $Bel_Y(-\infty, y)$  in the top right; upper and lower CDFs of  $Z = X + Y$ , i.e.  $Pl_Z(-\infty, z)$  and  $Bel_Z(-\infty, z)$  in the bottom middle.



**Figure 3.5** – Propagation of uncertainty represented in terms of bodies of evidences when input parameters are considered to be independent (left) and dependent (right).

The result for the same example in section 3.1.4.1 when  $X$  and  $Y$  are considered to be totally dependent is depicted in figure 3.6.



**Figure 3.6** – Body of evidence for the output  $Z = X + Y$  when total uncertainty between X and Y is considered.

### 3.1.4.3 Why evidence theory is the best to treat ignorance

Evidence theory describes uncertainty by defining a subsets  $S$  of focal elements where the real value lies in. These focal elements are assigned with degrees of likelihood  $m$  (*bpa*) between 0 and 1. The unassigned *pba* (i.e.,  $1 - m(S)$ , where  $m(S) = \sum m(s_i)$  and  $s_i$  is the *pba* of the focal element  $i$ ) represents the ignorance or incompleteness in the expert knowledge [55]. In terms of example, let us suppose that the subjective probability of an event to occur is either low or high. And evidences have been collected about this event where it can be low with a likelihood of 0.6 and high with a likelihood of 0.2. Since the sum of the probabilities is not equal to 1, the remaining 0.2 implies ignorance with regard to the POC. Perhaps we might think that this problem can be solved by probability theory by choosing a suitable outcome space  $\{[low], [high], [low, high]\}$ . Then we define now that  $p([low]) = 0.6$ ,  $p([high]) = 0.2$ , and  $p([low, high]) = 0.2$ . But, if we compute the probability of the union of low and high, then  $p([low \cup high]) = p([low]) + p([dry]) = 0.8$ ,

which is in contradiction with  $p([low, high]) = 0.2$ .

## 3.2 Industrial Automation and Control System - IACS

Industrial automation is the use of Industrial Control System (ICS), such as computers and information technologies for handling different processes in an industry. The use of ICS helps in increasing productivity, quality and flexibility in the manufacturing process [12].

The SCADA system is one of the most important parts of IACS, which refers to an industrial computer system that monitors and controls processes and systems distributed over limited or large geographical areas ([120]; [31]). The principal function of SCADA is acquiring the data from devices such as valves, pumps, etc. and providing control of all of these devices using a host software platform ([106]; [139]). The monitoring of the process is provided using a remote method of capturing data and alarm events, where instruments can be regulated and turned on and off at the right time. The SCADA system also provides more functions such as displaying graphics, alarming facilities and storing data. Malfunctions of SCADA may cause undesirable consequences ranging from financial loss to environmental damages [123].

SCADA systems throughout the world supervise and control electric grids, power plants, water systems, chemical plants, pipelines, manufacturing, transportation, and other physical processes [157]. Figure 3.7 shows the basic hierarchy and architecture of an IACS, which is classified into five distinct levels. SCADA operates on levels 1 and 2. The different levels of IACS are presented as follows:

- level 0 - field instruments: the lowest level of the control hierarchy which includes to sensors, pumps, actuators, etc. that are directly connected to the plant or equipment. They generate the data that will be used by the other levels to supervise and control the process;
- level 1 - control level using Programmable Logic Controller (PLC): PLC is an adapted industrial digital computer that controls the manufacturing processes. It is linked to the field instruments, and to the SCADA host software using a communication network;
- level 2 - SCADA: monitor, maintain and engineer processes and instruments;
- level 3 - MES: this level is responsible for process scheduling, material handling, maintenance, inventory, etc;
- level 4 - ERP: the top level of the industrial automation which manages the whole control or automation system. This level deals with commercial activities including production planning, customer and market analysis, orders and sales, etc.



Industrial communication networks are most prominent in IAS which represents the link that relays data from one level to the other in order to provide continuous flow of information. This communication network can be different from one level to another.

The SCADA system represents the most sensitive and targeted part of the industrial automation in terms of cyber-security. Cyber attacks on the SCADA system are classified into three different categories: (i) hardware, (ii) software, (iii) communication network.

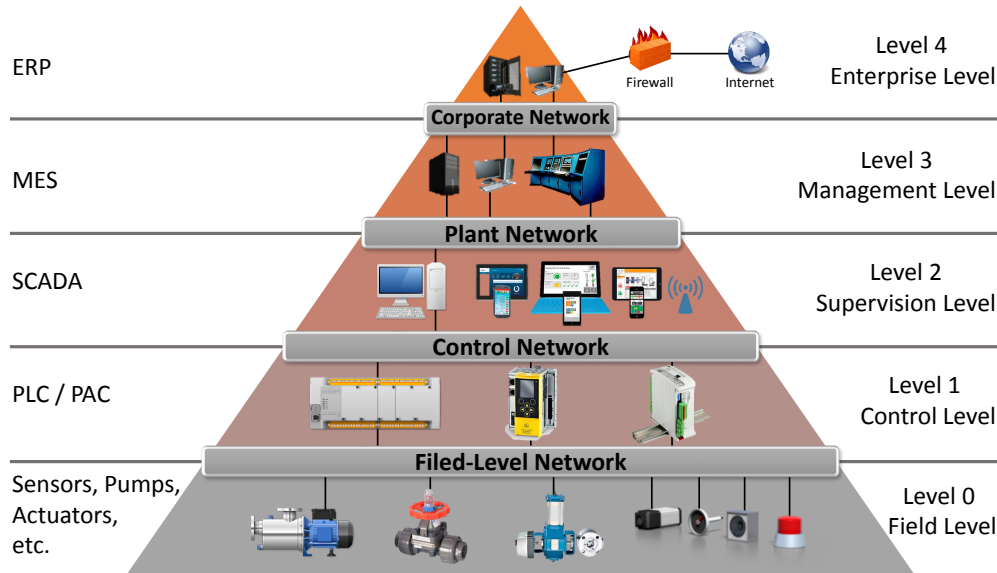


Figure 3.7 – Components and architecture of IAS.

### 3.3 Cyber-security for industrial control systems

Within modern communication environments, such as the corporate infrastructure for managing the business that drives operations in plant control systems, there are people, processes and technology related vulnerabilities that need to be addressed ([82]; [142]). Historically, these issues have been the responsibility of the IT security by providing security policies to protect vital information assets. As control systems become part of these large architectures, the main concern is providing relevant security procedures that cover the control system and assure the safety for the facility and the surrounding environments.

Introducing technology in critical facilities provided different benefits in terms of control and production, but it also generated challenges that are related to securing these facilities:

- increased connectivity: today’s ICS are increasingly connected to the enterprise business level and are accessible through the Internet. Even though these connections

improve operability, they also create security vulnerabilities because improvements in the security measures of control systems are not concurrent.

- inter-dependencies: due to the high degree of interdependency among infrastructure sectors, failures within one sector can spread into others. A successful cyber attack might be able to take advantage of these interdependencies to produce cascading impacts and amplify the overall economic damage.
- complexity: the demand for real-time control has increased system complexity in several ways: 1) access to ICS is being granted to more users business; 2) ICS are interconnected, and, 3) the degree of interdependency among infrastructures has increased. Dramatic differences in the training and concerns of those in charge of IT systems and those responsible for control system operations have led to challenges in coordinating network security between these two key groups.
- system accessibility: even limited connection to the Internet exposes ICS to all of the inherent vulnerabilities of interconnected computer networks, including viruses, worms, hackers, and terrorists. Control channels that use wireless or leased lines that pass through commercial telecommunications facilities may also provide minimal protection against forgery of data or control messages. These issues are of particular concern in industries that rely on interconnected enterprise and control networks with remote access from within or outside the company.
- information availability: manuals and training videos on ICS are publicly available, and many hacker tools can now be downloaded from the Internet and applied with limited system knowledge. Attackers do not have to be experts in control operations to create an impact.

ICS-CERT (Industrial control systems - cyber emergency response team ) listed over 250 attacks on ICS in 2013:

- 59% of attacks targeted the energy sector
- 79 attacks successfully compromised the target
- 57 attacks did not succeed in compromising the target
- 120 attacks were not identified/investigated

In addition, cyber-security should be put in place together with safety and not separated. Cyber-security protects control systems to keep the critical facility processes working safely. It ensures data communication confidentiality and integrity. It also keeps computers and the control systems protected from viruses, worms, Trojans and other sophisticated threats that can sabotage these systems and cause damages. Cyber-security assure access restrictions by allowing the right people the have access on controls and important information and keeps the wrong people out of the controls.

## 3.4 The security risk analysis process for industrial control systems

The methodology of security risk analysis comprises a number of basic steps. These differ between authors, but in general include:

1. risk identification:
  - asset identification: this step aims to identify what to protect: hardware, software, information, infrastructure, people, etc. Here we care about protecting people and the environment from major accidents;
  - threat identification: identify potential causes of an unwanted incident, which may result in harm to an asset or set of assets. A thorough understanding of the threats to the system is required in order to protect the assets;
  - vulnerability identification: identify weaknesses of the system that can be exploited by a threat or set of threats.
2. likelihood analysis: the aim of likelihood analysis is to estimate the probability of exposure of the system to a threat (or set of threats), and how likely these threats are able to exploit the system's vulnerabilities in order to harm its assets. Determining likelihood is related to the current security measures and the environment in which they are applied;
3. consequence analysis: calculating the impacts of threats exploiting vulnerabilities on the assets.

## 3.5 A review of cyber-security risk analysis approaches for Industrial Control Systems

This section reviews the state of art in cyber-security risk analysis of industrial control systems. We select the most common approaches that provide graphical representation of security risks. We choose to review only graph-based methods because INERIS uses a graph-based approach (bow-tie analysis) to analyze safety risks. Graph-based approaches are widespread and more practical for modeling systems' components and functionalities. Moreover, they are able to analyze the security of complex systems with a high level of precision and details. We describe the methods in terms of aim, concept and the application domain. These approaches are then analyzed and discussed to highlight their advantages and limits. Based on this review, we point out the most suitable approach to be used to deal with our problem.

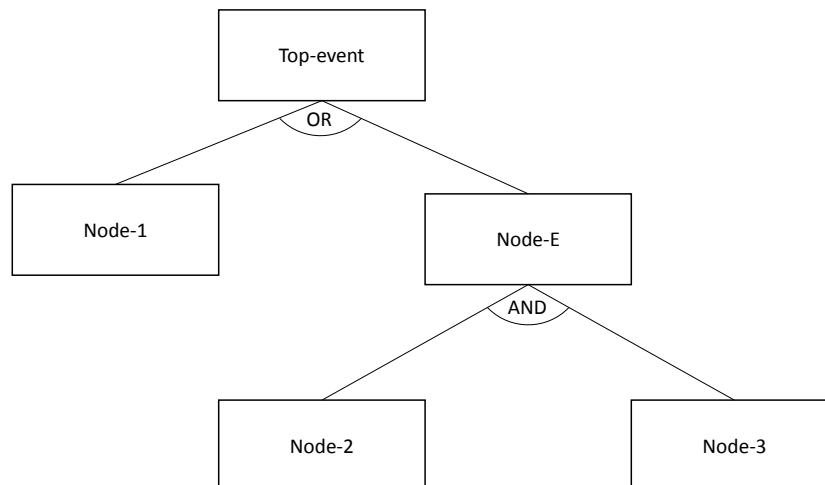
Before proceeding, graphical-based methods are risk analysis methods that are based on a graphical model such as fault tree, attack tree, etc. These methods, in the majority of cases, are supported by mathematical models to enable qualitative or quantitative likelihood analysis [32].

However, since the use of probabilistic risk analysis approaches is an obligation in France, only probabilistic approaches is reviewed in this section.

### 3.5.1 Attack-tree-based approaches

The « Attack Tree » technique as initially presented by [140] is a graph that describes the sequence of steps in order to perform an attack. It represents an attack against a system in a tree structure [67]. The root (main event) of the tree is the goal of an attack. This root is connected to intermediate and starting (leaf nodes) events in order to represent the different ways to achieve the attack.

The concept of the attack tree is inspired from the fault tree. As in fault tree, AND and OR gates are used to describe the combinations between nodes. Figure 3.8 presents a demonstrative example of attack tree as in its creation form by [140]. The goal of the attack is the top event. To achieve this goal, attackers can follow two paths, Node-1 or (Node-2 and Node-3). The Node-E is a sub-goal in the tree, and its children are ways to achieve this sub-goal. However, this initial version of attack tree shows only the steps that attackers should follow in order to compromise a system. In other word, system's vulnerabilities are not modeled.



**Figure 3.8** – Attack tree structure using an example as developed by [140].

In the rest of this section, we present some papers that are based on the attack tree methodology. In these papers, the attack tree is either extended (new modeling components are added) or combined with other models.

### 3.5.1.1 Attack trees for assessing vulnerabilities in SCADA [28]

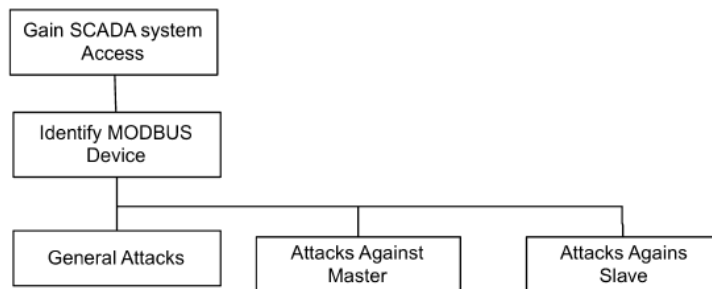
In this paper, [28] describe the application of the attack tree methodology to assess vulnerabilities in the common MODBUS/TCP-based SCADA system. The same modeling structure proposed by [140] is used in [28]. By assessing the possible attacks and security vulnerabilities on SCADA systems using attack trees, the authors suggest possible best practices for SCADA operators and improvement to the MODBUS communication protocol.

According to [28], several factors are taken in order to evaluate the security risks:

- Technical Difficulty of attack (it is believed to be the most critical indicator);
- Probability of Apprehension;
- Cost of Attack;
- Site Conditions;
- Installed Countermeasures.

The purpose of the assessment is to determine all the attacker final goals that intruders might attempt to achieve against a MODBUS-based SCADA system, and to identify all possible ways to achieve these final goals. To do this, a team of experts identifies all possible goals of attackers against the SCADA system and the causes of each goal to design the attack trees. Then, each leaf node of an attack tree is assigned a level of technical difficulty using a qualitative scale “Trivial-Moderate-Difficult-Unlikely”. Based on the AND/OR logical functions as the maximum/minimum of the children nodes difficulty values, the difficulties of leaf nodes are propagated through the tree and the difficulty of each attack goal is calculated. Then, each attack goal is ranked in terms of attack difficulty, likelihood of detection and the potential severity of impact.

The study presents a sample of attack tree (Figure 3.9) with the estimated attack difficulty of its nodes. The authors concluded that the attack trees can be a very useful tool for modeling threats and vulnerabilities in a wide variety of systems.

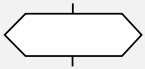
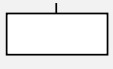




**Figure 3.9** – Attack tree example for the MODBUS-based SCADA system given in [28].

### 3.5.1.2 Through the Description of Attacks: a Multidimensional View [67]

In [67], an extended version of attack tree is proposed. This extended version allows the consideration of more information than just the attack steps, such as attacker resources, motivations, etc. The term, shape and definition of each event are presented in Table 6.1. Figure 3.10 shows an example of the attack tree as presented in [67].

**Tableau 3.2** – Description of events used for representing an attack scenario.

	shape	Signification	Definition
Input events		Operation	Any step representing an operation made by the attacker in order to perform the attack
		Vulnerability	Any step describing a vulnerability required in order to realize the attack
		Assertion	Any step representing assumptions, results, or requirements characterizing the attack process
		Intermediate/ top event	A security breach caused by the occurrence of input events

### 3.5.1.3 Attack countermeasure trees (ACT): towards unifying the constructs of attack and defense trees [135]

In [135], a novel attack tree model called attack countermeasure tree (ACT) is proposed. The aim of the ACT is to place attacks as well as defense mechanisms on the same model by combining the defense trees within the attack trees. There are three types of events in an ACT: attack event, detection event and mitigation event. Figure 3.11 presents an example of the ACT.

ACT provides qualitative and quantitative security analyses. Qualitative analysis allows the identification of the minimal combinations of attack events that lead to the occurrence of the top event and the determination of the most critical event in the ACT. Quantitative analysis using ACT allows the computation of success attacks probabilities based on the probabilities of single attack events. Several quantities can be calculated: the cost of attacks, the impact of attacks, birnbaum or reliability importance measure, risk to the attacker and the system, the benefits to attackers from attacks and the benefits of defenders from implementing countermeasures. Quantitative analysis using ACT can be viewed from two distinct viewpoints: attackers' viewpoint and defender's (or security analyst's) viewpoint. The measures such as attack cost and benefits of attackers reflect the attacker's perspective whereas the security investment cost, risk, impact and benefits of defenders represent the defender's perspective.

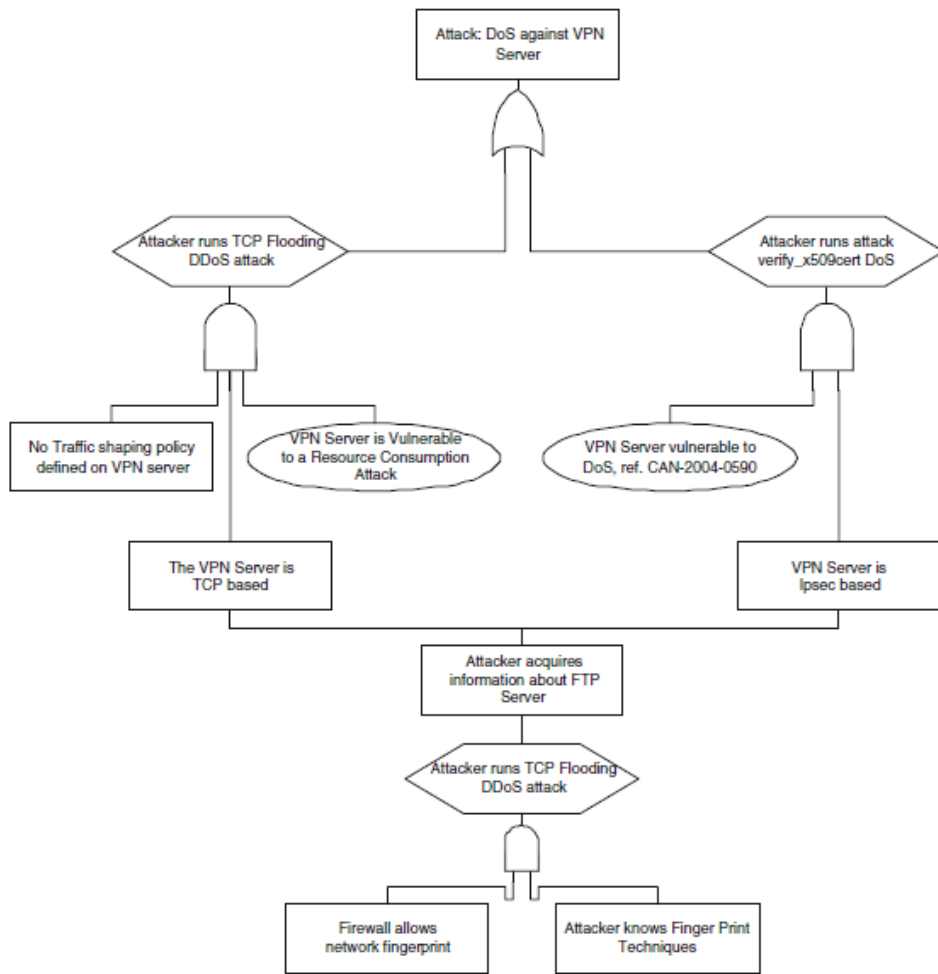


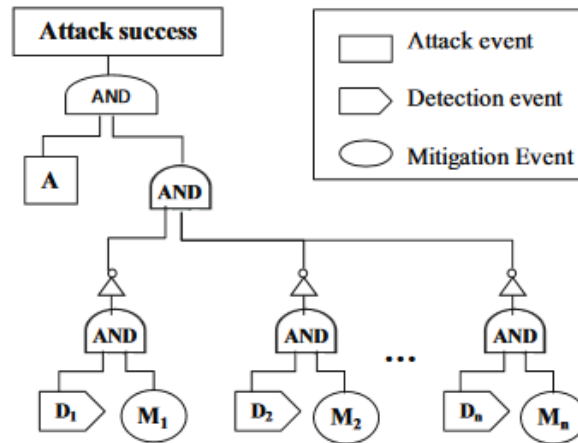
Figure 3.10 – Example of the attack tree proposed in [67].

The ACT in [135] has been implemented in the software tool SHARPE (Symbolic Hierarchical Automated Reliability and Performance Evaluator). The use of the ACT is demonstrated using three case studies (ACT for BGP attack, ACT for SCADA attack and ACT for malicious insider attack).

### 3.5.2 Security Modeling with BDMP: From Theory to Implementation [126]

Boolean logic Driven Markov Process (BDMP) is a good solution to model and analyze complex systems. The concept of BDMP is inherited from combining fault trees and Markov models. In [126], the BDMP is adapted to the security domain in order to graphically model cyber attacks.

Different types of leaf events are used for security analysis using BDMP: (1) the Attacker Action (AA), (2) the Timed Security Event (TSE), (3) the Instantaneous Security



**Figure 3.11** – ATC with one attack and multiple pairs of detection and mitigation events [135].

Event (ISE).

The BDMP provides a useful qualitative and quantitative security analysis. The application of the BDMP in the security domain is demonstrated using complex case studies: modeling STUXNET attack [100].

### 3.5.3 Evaluating the risk of cyber attacks on SCADA systems via Petri net analysis with application to hazardous liquid loading operations [80]

A analytic technique to analyze security risks for SCADA systems is developed in [80]. The technique constitutes a novel application of Petri net state coverability analysis coupled with process simulation. It consists of three main components:

- a Petri net-based model to represent the attack scenarios;
- a technique to assess both the depth and breadth of an attack;
- a technique for mapping Petri net coverability to operational consequences.

The objective of the method is to identify and mitigate the system’s vulnerabilities in order to prevent malicious induction of catastrophic process failure modes.

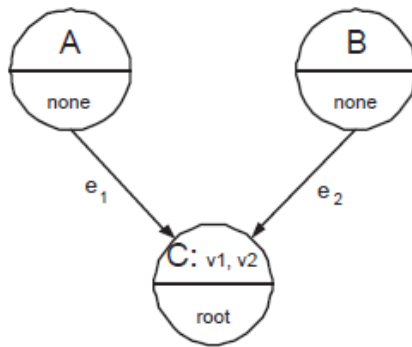
This approach does not take likelihood as a notion to evaluate the risk. Instead, risk is measured in terms of the extent to which an attacker can manipulate process control elements, the consequences due to disruption of the controlled physical process, and the vulnerability of the SCADA system to malicious intrusion. The method is demonstrated by the application on a hazardous liquid loading process.



### 3.5.4 Network Vulnerability Assessment using Bayesian Networks [108]

In [108], the Bayesian network is used to model attack paths. The proposed approach is called "Bayesian attack graph". The graph is made up of nodes and edges to relate the nodes. Each node in the graph represents a single security state. An edge corresponds to an exploitation of one or several vulnerabilities. A series nodes and edges represent a path of a potential attack.

For probability analysis, the probability of a successful exploit is assigned as the weight to each edge. Each node in the graph can be in two different states  $\{false : 0, true : 1\}$ . A true state indicates a compromised system state that accomplished by an attacker. Figure 3.12 presents a simple example of Bayesian attack graph.



**Figure 3.12** – A Simple Example of Bayesian Attack Graph [108].

Figure 3.12 shows three nodes A, B and C where each node represents a potential security violation state. Two vulnerabilities  $v_1$  and  $v_2$  exhibit on node C.  $v_1$  is exploitable from host A with the probability of success  $p(e_1)$ .  $v_2$  is exploitable from host B with the probability of success  $p(e_2)$ . Both exploitations result in the same compromised state on node C. The probability of C is calculated using the conditional probability specifications. The approach is tested on an experimental network.

### 3.5.5 Discussion

In this section, we summarized the graphical-based approaches that help integrating cyber-security within safety risk analysis. We reviewed the most common methods to analyze security for ICS that are either based on: attack tree, defense tree as well as other security analysis methods based on directed graphs (e.g. Petri net and Bayesian network) that fall under the category of probabilistic risk analysis (Section 3.5). In particular, we discuss the used models, type of the analysis and the application of the method. Comparison between these approaches is coming in the rest of this section.

Only graph-based approaches are reviewed here because we believe that they are the most promising and suitable approaches to meaningfully model and analyze risks. Large efforts are currently conducted to develop these approaches ([115], [102]).

Traditionally, risk analysis methods are classified into qualitative and quantitative ([148]; [93]; [110]). However, we classified the reviewed approaches regarding different categories as presented in Table 3.3. From Table 3.3, the results of comparison between the examined approaches is as follows:

NO.	Ref.	Key security risk concept in the attack tree				Type of likelihood analysis		Applicability	
		Asset	Vulnerability	Threat/Attack	countermeasures	Qualitative	Quantitative	Simple	Complex
1	[28]			✓		✓		✓	
2	[67]		✓	✓			✓	✓	
3	[134]			✓			✓	✓	
4	[125]			✓			✓		✓
5	[80]		✓	✓			✓		✓
6	[108]			✓			✓		✓

**Tableau 3.3** – List of approaches combining safety and security for ICS.

- key risk concepts: show which key concepts (Asset, Vulnerability, threat or/and countermeasures) are considered by the method examined for likelihood analysis and scenario representation. We conclude that none of the examined methods takes into account the existence of countermeasures, and only two have taken how vulnerable the system is to a specific attack step;
- type of likelihood analysis: Only the method proposed by [28] is qualitative while the others use quantitative likelihood analysis to evaluate the security risks. Qualitative analysis is easy to conduct but not always precise. Statistical data is needed to perform a quantitative likelihood analysis which is often unavailable. For security analysis, qualitative analysis is preferable since no much statistical data has been recorded to conduct a quantitative likelihood analysis;
- applicability of the methods for security analysis: applying methods that use attack trees is simple and rapid. While applying BDMP, Bayesian network or Petri net based approaches is time consuming and complex due to their dynamic concept. However, we can not ignore the effectiveness of these latter in reliability analysis of complex and dynamic system.

Based on this overview and discussion, we believe that attack tree is the most suitable method for us to use. Bow-tie analysis used by the INERIS for risk analysis is of static nature. Thus, there is no need for using a complex dynamic modeling for security analysis. A static simple modeling is more preferable.

However, existing attack tree based approaches do not provide a thorough modeling of a risk scenario. More information should be plotted on the attack tree beside the attack steps in order to detail the system' weaknesses. A system might be more or less vulnerable to a certain attack steps. This information will be helpful to provide more precise likelihood analysis and detailed risk representation. For this latter, a new extended attack tree will be proposed in order to provide a complete modeling of cyber attacks. This method is then combined within bow-tie analysis to integrates cyber-security risks within safety industrial risk analysis as presented in the next chapter.

## **3.6 Existing Approaches that combine safety and security for industrial control systems**

With the growing awareness that safety and security analyses should be coordinated in risk analysis for complex systems, working groups and researchers have initiated work to bridge the gap between safety and security [102]. Approaches to analyze safety and security together have been developed in some specific domains such as the nuclear domain [127], the aerospace industry [25] and in air traffic control management [133].

In this section and for the same reasons discussed in the previous section, the most reputed graph-based approaches are reviewed and discussed. For other methods, [102] provided a detailed survey on initiatives standards and approaches that integrate cyber-security and safety. We outline the challenges facing the domain and the limits of the existing approaches. We also give a hint of the solution.

### **3.6.1 Integrating cyber attacks within fault trees [68]**

In [134], a new graph-based method to analyze security within safety for complex systems is proposed. The method combines the attack tree introduced by [67] (reviewed in Section 3.5.1.2) within fault tree analysis. Combining attack trees with the fault tree allows capturing accidental (safety) and malicious (security) risks that can lead to safety accidents.

The approach is made up of the following steps:

1. constructing the fault tree: identify the top event and all its related sequences of causes;

### 3.6 Existing Approaches that combine safety and security for industrial control systems

---

2. security analysis of the fault tree: identify which events in the fault tree developed step 1 can be triggered using malicious attacks;
3. constructing the attack trees: for each event (e) identified in step 2, construct the attack tree and attach it to the event (e) in the fault tree using an OR gate with its children.

The approach also provides a quantitative probability analysis in order to calculate the probability of top event in the fault tree. Quantitative probability values are assigned to the leaf and basic events in the attack and fault trees, respectively. The mathematical equations to propagate the probability of input events through the proposed attack/fault tree model are provided. The approach is tested on a simple example.

#### **3.6.2 Modeling safety and security inter-dependencies with BDMP (Boolean logic driven Markov processes) [125]**

In [125], the BDMP discussed in Section 3.5.2 is applied to analyze safety and security inter-dependencies. BDMP is used to graphically model the different causes of undesirable events in a system. These causes can related either to accidental (safety causes) or malicious events (security causes) [102].

For probability analysis, each basic safety and security leaf in the BDMP is assigned a mean time to success and a mean time to failure, respectively. These input data is on assumptions made by security and safety experts. Then this input data is propagated through the BDMP to provide a quantitative likelihood estimation of the modeled undesirable event.

#### **3.6.3 Studying interrelationships of safety and security for software assurance in cyber-physical systems: Approach based on Bayesian belief networks [98]**

In this paper, [98] analyzed safety and security jointly using Bayesian belief network. The Bayesian technique was applied to determine the impact of equipment failures or security vulnerabilities on the overall security and safety of the entire system. Nodes and arcs in the Bayesian graph are used to model system components failures and the transition from a state to another, respectively.

The Bayesian approach in [98] provides a quantitative safety/security likelihood evaluation. The approach uses failure rates of system components and connections or likelihood of incidents impacting safety and security to quantitatively evaluate the achievement or

denial of safety and security. The approach is demonstrated using a case study of a SCADA system in an oil pipeline control.

### 3.6.4 Modeling and analysis of safety-critical cyber physical systems using state/event fault trees [134]

In [134], a common model to analyze safety and security aspects for critical physical systems is proposed. This method extends the state/event fault trees with an attacker model to develop a common model to deal jointly with safety and security. This extension provides a modeling and analysis approach that integrates security aspects into a safety model and enables quantitative analysis of safety and security.

The proposed modeling tool enables the trees to be modeled and converted into extended deterministic stochastic Petri nets for quantitative analysis; which are then analyzed using the Time Net tool and steady state analysis or Monte Carlo simulation.

This approach is evaluated case study of a tire pressure monitoring system.

### 3.6.5 Discussion

In this section, we reviewed the approaches that analyze safety and security jointly during risk analysis (Section 3.6). The approaches presented in the literature are based on either adapting an approach from the safety or security domains, or combining different methods from the two domains. Combining different techniques developed in each domain (safety and security) presents some advantages:

- relationships between safety and security become more visible;
- differentiation between risk causes related to each domain is provided;
- the use of existing models is easier than developing a new one to combine the both aspects of safety and security. Where in this case, guides about the methods and real case applications are available;
- more information for decision makers is presented. Separation of properties would permit recognizing if a high level risk is returned to safety or security causes.

However, in this section, we only reviewed papers that use: attack tree, BDMP, Bayesian Network or Petri net alone or in combination with other safety methods. In other words, we reviewed approaches that use the approaches presented in Section 3.5.

Fault trees and attack trees are the most used approaches for safety and security analysis respectively. Beside their static nature, they are simple to apply and provide a clear risk modeling of the system being studied. In addition, likelihood analysis using fault or attack trees is easy to conduct and demands low-level of expertise. Likelihood

### 3.6 Existing Approaches that combine safety and security for industrial control systems

analysis can be qualitative or quantitative. It can be manually conducted without the need of a computer software or simulation.

Bayesian belief networks is a good quantitative probabilistic framework to analyze risk but it also has many limits. This technique is time consuming, computationally expensive and hard to interpret. Petri nets provide qualitative and quantitative modeling of safety and security risks. The limitations of Petri net model are: not trivial to be built and can be difficult to analyze. The use Markov chain increase exponentially the complexity of computation and the number of states when the number of events increase.

Furthermore, in all the examined approaches, safety and security are treated within the same scale in probability analysis when they should not be. By providing single likelihood index for safety and security, decision makers would not be able to know if the unacceptable likelihood is generated from safety or security related causes. These limits will be handled by proposing an approach for evaluating the risk level based on two-terms likelihood parts, one for safety and one for security.

As conclusion, combining attack tree and fault tree is the most appropriate approach. However, due to the limits of existing attack tree models as discussed in Section 3.6.5, none of these approaches will be used. In the next chapter, an extended attack tree will be proposed and combined within the bow-tie analysis to provide an exhaustive risk analysis that jointly consider safety and security.

The theories presented in this section are rich in content and powerful to quantify parameter uncertainty. But, different types and causes of parameter uncertainty require different methods of analysis. Each theory suits a specific type of parameter uncertainty regarding the causes of this type. Thus, classifying parameter uncertainty into corresponding types according to their causes simplifies their analysis and modeling by the most suitable theories. Appropriate analysis of uncertainty will lead to more accurate risk predictions and consequently to a better decision making.

In Chapter 5, fuzzy theory will be introduced instead of interval theory to handle the limits of the interval semi-quantitative approach for likelihood analysis.

In Chapter 6, new approaches for parameter uncertainty analysis are developed to treat uncertainty regarding the type and causes. These proposed approaches and the approaches presented above shall be applied to an LOC scenario in order to calculate the toxic concentration at a specific end-point after an undesired event has occurred taking into account parametric uncertainty. The aim is to see the effect of this uncertainty on the output result, compare the different results of each approach and determine where a specific approach should be used and where it should not. Java software was developed to implement the theories of uncertainty representation, the propagation algorithms and the risk models.

## 3.7 Conclusion

The first part of this chapter reviews the most common approaches to handle parameter uncertainty. The purpose of quantifying uncertainty is to meaningfully represent uncertain parameters, then to propagate these representations through mathematical models in order to be able to represent uncertainty in the outputs. In this chapter, representing and propagating parameter uncertainty using: interval analysis, fuzzy theory, probability theory, evidence theory are presented. These approaches are used to characterize uncertainty in model inputs obtained from different sources, such as statistical data and expert judgments, and to which different types of parameter uncertainty can be attached. But, these theories are of different nature. In other words, each approach represent and propagate parameter uncertainty in a specific way and different than the others. Simple examples on how addressing parameter uncertainty based on each theory are also provided.

However, these approaches were not compared in this chapter. Chapter 5 shows the application of these approaches on a real case study and discuss the difference between them. The purpose of this comparison is checking how these approaches differ in terms of analyzing parameter uncertainty and sizing up the advantages and disadvantages of each one. This comparison aims to identify which theory or approach is the most suitable to deal with parameter uncertainty regarding the types and causes of this uncertainty.

The part focuses on cyber-security of critical industries and how it might affect the safety these industries. Cyber attacks on ICS and SCADA systems are increasing everyday. Today's security analysis approaches focus on protecting information and assuring the availability of the system where they ignore safety. Luckily, until now major disasters due to cyber-attacks have not been occurred. But, without taking precautions we may not hope for this to happen in future as attackers get more sophisticated, experienced and malicious.

However, due to the incompleteness of safety risk analysis methods, the actual value of the probability of undesirable events occurrence is higher than estimated. For example, for incidents in power industry it was noted that "While these may not be frequent in an absolute sense, there are good reasons to believe that they will be far more frequent than quantitative tools such as probabilistic risk assessments predict". Therefore, introducing cyber-security is important to provide a precise risk evaluation.

This second part of this chapter contains a structured comprehensive overview of graph-based security and security/safety risk analysis methods. Overall, the findings of the first part of this chapter are:

- a review on the state of art in security analysis for ICS and safety/security combined

for critical facilities;

- highlighting the advantages and disadvantages of the reviewed methods and pointing out the most suitable to be used;
- outline the research challenges in the domain of considering cyber-security related risks with safety analysis that existing approaches did consider.

This review indicates that despite the fact that existing approaches can jointly consider security within safety but they still suffer from challenging limits. Risk analysis methodology that jointly consider safety and security can be improved in terms of: (1) Modeling attack as well as system characteristics on the same view, and (2) considering the difference in likelihood of occurrence between safety and security related causes. Addressing these limits is very important to provide a more detailed and effective risk analysis for decision making.





# 4

## Treatment of uncertainty in probability analysis: a fuzzy semi-quantitative approach

**Summary:** In this chapter, we develop a fuzzy semi-quantitative bow-tie analysis to address data uncertainty and as an alternative for losing quantitative data or adding unjustified information. A fuzzy-based approach is used for handling subjectivity and measurement errors in the input parameters. The application of the proposed approach is demonstrated using the case study of a loss of containment scenario (LOC) in a chemical facility.

### Summary

---

<b>4.1</b>	<b>Introduction</b>	<b>83</b>
<b>4.2</b>	<b>Proposed methodology: fuzzy semi-quantitative approach</b>	<b>84</b>
4.2.1	Characterizing inputs for probability analysis using the fuzzy semi-quantitative approach	84
4.2.2	Representing (fuzzifying) input data based on the proposed fuzzy scales	87
4.2.3	Propagating fuzzy frequencies through the Bow-Tie	88
4.2.4	Decision making under fuzzy environment	92
4.2.5	Handling the existing limitations of the interval approach	93
<b>4.3</b>	<b>Case study</b>	<b>95</b>
<b>4.4</b>	<b>Conclusion</b>	<b>97</b>

---

This work was carried out in collaboration with the INERIS. The results developed in this chapter have been presented in the following articles:



[5] H. ABDO, J-M. FLAUS, F.MASSE. *Fuzzy semi-quantitative approach for probability evaluation using Bow-Tie analysis*. Safety and Reliability Theory and Applications (2017).



[2] H. ABDO, AND J-M. FLAUS. *Uncertainty quantification in bow-tie analysis: A mixed approach of fuzzy theory with Dempster-Shafer theory of evidence*. In *Risk, Reliability and Safety: Innovating Theory and Practice: Proceedings of ESREL*, Glasgow, Scotland (2016).

## 4.1 Introduction

Probability analysis may be qualitative or quantitative depending on the circumstances [104]. Data for performing a probability analysis is either from historical incident data or expert elicitations [3].

Qualitative probability analysis uses a scale of qualitative expressions (low, medium, high, etc.) to describe an event's probability or frequency. The advantage of the qualitative methodology is its simplicity of applying and understanding by the relevant personnel. Expert judgments represent an important source of data to apply this methodology which are subjective in nature. This subjectivity represents a disadvantage when quantitative or more precise information is available.

The quantitative approach uses a numerical scale with real values to describe the event frequencies based on statistical data. A problem of this approach is the imprecision and lack of such data can affect the quality of the analysis [2]. The quantitative methodology is often too expensive and complex to be performed in terms of time and cost since statistical and empirical data are needed. It can lead to probability underestimation if uncertainty is not taken into consideration [4]. For these disadvantages, a semi-quantitative approach represents a better alternative based on the available information.

INERIS has developed an interval semi-quantitative bow-tie analysis to model and quantify the probability of risk scenarios based on the available information as presented in Chapter 2, Section 2.2.2. This approach is mainly based on the INERIS expertise and the results of the European project ARAMIS ([83]; [84]). It uses historical accident data (quantitative) or expert elicitations (qualitative) if the former is not available. It is easy to use, effective and implicitly takes uncertainty into consideration. Based on an international benchmark exercise, this semi-quantitative approach proved to be effective and precise for estimating the probability of risks in comparison with the approaches used in the United Kingdom, the Netherlands and the Walloon Region of Belgium which are either qualitative or quantitative ([71]; [105]).

However, as highlighted in Chapter 2, Section 2.6.1, the interval semi-quantitative presents some limits. In some cases, this interval semi-quantitative approach can lead to probability underestimation. We handle these limits by introducing the concept of fuzzy numbers instead of intervals. Fuzzy numbers are used to represent subjectivity in expert judgments and covers uncertainty in the quantitative data if this data exists. This proposed fuzzy semi-quantitative bow-tie contributes to a simpler and effective alternative to the quantitative approach and more precise to the qualitative approach while keeping the virtue of being based on real accident frequency data if presented, and with the consideration of uncertainty.

In order to present the advantages offered by this fuzzy semi-quantitative approach for probability analysis, this Chapter is structured as follows: Section 4.2 examines the proposed methodology to deal with the limitations of the interval semi-quantitative approach, and how the methodology handles the limits of the interval semi-quantitative approach. Section 4.3 presents a case study and compares the fuzzy-based approach with the interval-based approach and the quantitative approach. Finally, a number of conclusions are drawn in Section 4.4.

## 4.2 Proposed methodology: fuzzy semi-quantitative approach

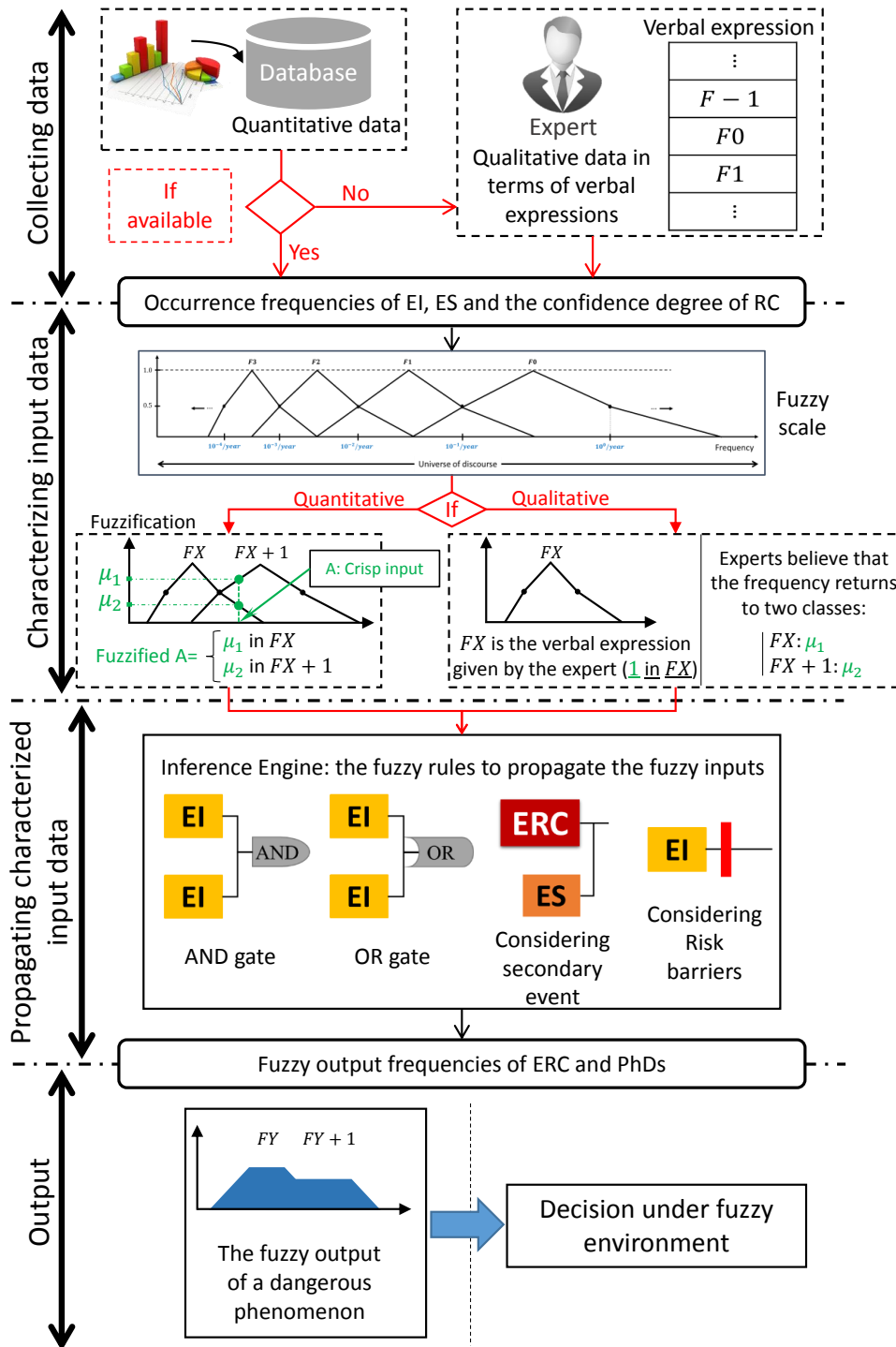
In this section, we present the added improvements to the interval semi-quantitative approach. The characterization of input data in terms of fuzzy numbers and the propagation rules to calculate the occurrence probabilities of ERC and outcomes are discussed in Sections 4.2.2 and 4.2.3, respectively. The framework developed in Figure 4.1 presents the steps of the proposed methodology for probability estimation under uncertainty using bow-tie analysis. The details of each step in this framework is given in the rest of this section.

### 4.2.1 Characterizing inputs for probability analysis using the fuzzy semi-quantitative approach

In this section, we will detail how intervals are replaced by fuzzy numbers to characterize input data in order to handle the limits of the interval semi-quantitative approach. Fuzzy scales are proposed to characterize the frequency of basic events, probability of occurrence of SEs and the CLs of risk barriers as respectively presented in Sections 4.2.1.1, 4.2.1.2 and 4.2.1.3. Section 4.2.2 describes how input data are translated into fuzzy inputs based on the proposed fuzzy scales.

#### 4.2.1.1 Define basic event frequencies using a fuzzy scale

Fuzzy numbers are used to express the linguistic frequencies as shown in Figure 4.2. In this fuzzy scale, a fuzzy number covers three intervals comparing to the interval scale. As shown in Figure 4.3, a class  $FX$  in the fuzzy scale covers the class  $FX$  and half of  $FX - 1$  and  $FX + 1$  in the interval scale. A fuzzy class covers three interval class in order to handle the vagueness on the borders between two classes in the interval scale. The middle point in the interval class  $FX$  is considered to be with the highest membership degree equals 1 for the fuzzy class  $FX$ . This membership degree decreases in the both



**Figure 4.1** – Framework for estimating the probability of accidents in Bow-Tie analysis based on the fuzzy semi-quantitative approach.

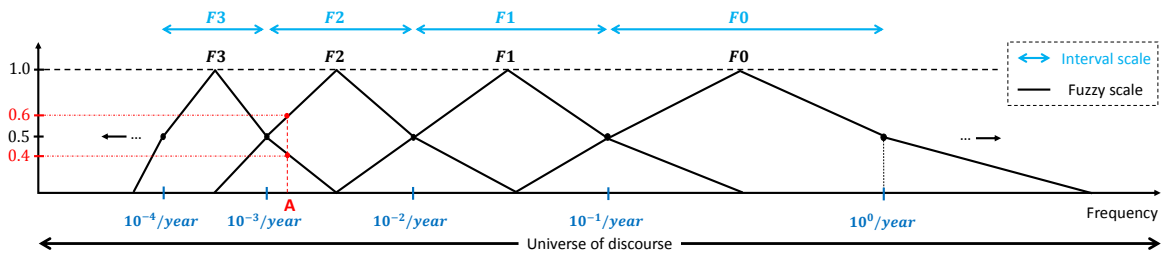
sides once recede from the middle point of the interval class  $FX$  until it gets 0 at the middle of the interval classes  $FX - 1$  and  $FX + 1$ .

Because the frequency classes follow a logarithmic scale, the grades  $FX$ ,  $X \in \mathbb{N}$  in the fuzzy scale are not triangular fuzzy numbers.  $FX$  is divided in three parts (see Figure 4.3) and derived with its membership function using Equations 4.1 and 4.2, respectively.

$$FX = [a, b, c, d, e] = \left[ \frac{10^{-(X+2)} + 10^{-(X+1)}}{2}; 10^{-(X+1)}; \frac{10^{-(X+1)} + 10^{-X}}{2}; 10^{-X}; \frac{10^{-X} + 10^{-X+1}}{2} \right] \quad (4.1)$$

$$\mu_{FX}(x) = \begin{cases} \frac{x - a}{b - a} \times 0.5 & a \leq x \leq b \\ 1 - \frac{x - c}{b - c} \times 0.5 & b \leq x \leq c \\ 1 + \frac{x - c}{c - d} \times 0.5 & c \leq x \leq d \\ \frac{e - x}{e - d} \times 0.5 & d \leq x \leq e \end{cases} \quad (4.2)$$

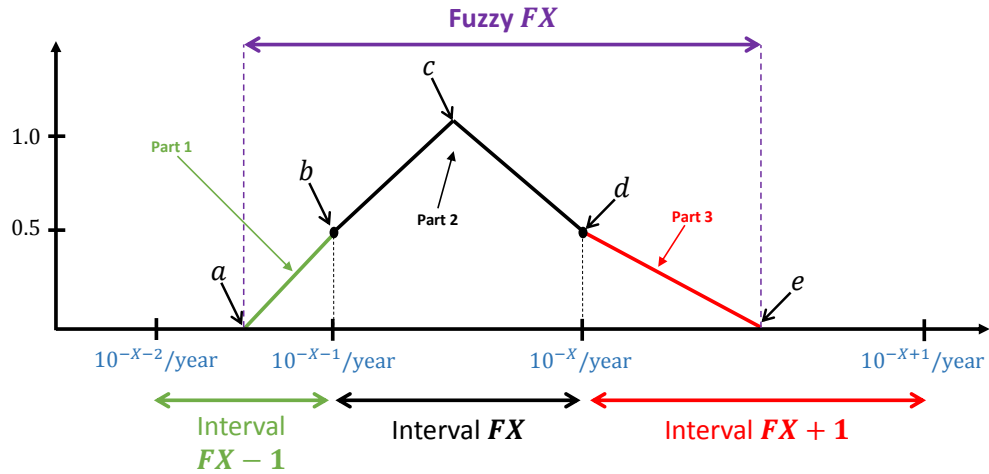
In the fuzzy scale, each value has its own possibility degrees to which classes it belongs (see Section 4.2.5 for more details). In Section 4.2.2, we will explain how to represent each type of input data based on this fuzzy semi-quantitative scale.



**Figure 4.2** – Mapping event frequencies on fuzzy scale.

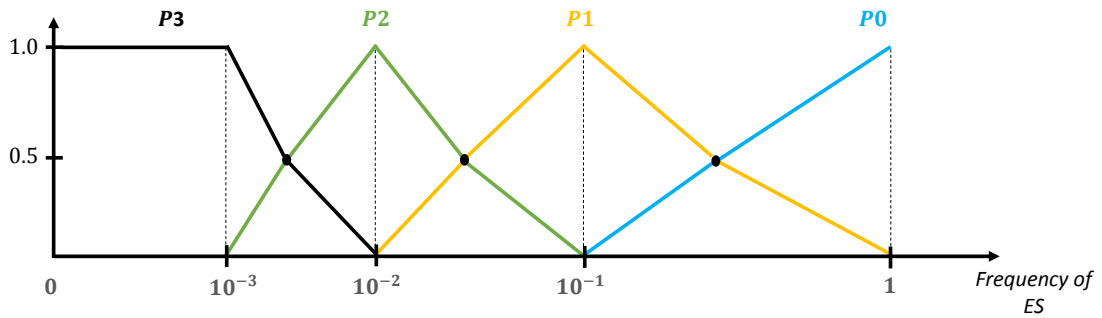
#### 4.2.1.2 Define secondary events probability of occurrence using a fuzzy scale

The used semi-quantitative approach to characterize the probability of ESs may lead to data lost due to the rounding-up. Therefore, because of this limitation, fuzzy scale is proposed to characterize the input probability of failure for secondary events. Figure 4.4 maps the probability of occurrences of secondary events on a fuzzy scale. The same type of fuzzy number used in the previous section is used here, but the values used in the



**Figure 4.3** – Fuzzy frequency class  $FX$ .

interval approach (the rounded up values) to represent probability of occurrence of ESs are considered to be the most possible values.



**Figure 4.4** – Mapping probability of secondary events on fuzzy scale.

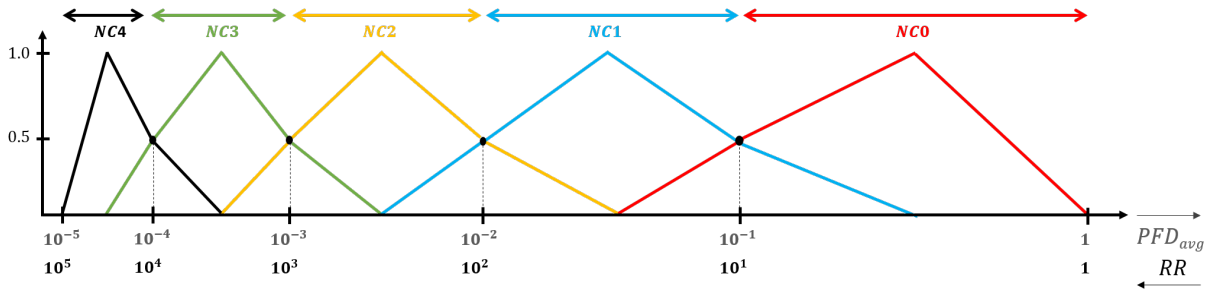
#### 4.2.1.3 Define risk barriers confidence levels using a fuzzy scale

For the same issues discussed in Section 2.6.1, fuzzy numbers are used to provide a fuzzy scale for determining NCs of risk barriers. The same type of fuzzy number utilized in scaling the frequencies of input events is used to scale the NCs. This fuzzy scale is depicted in Figure 4.5.

### 4.2.2 Representing (fuzzifying) input data based on the proposed fuzzy scales

Statistical accident data and expert elicitations are both used in the fuzzy semi-quantitative approach. Thus, quantitative crisp values and qualitative verbal expressions represent the input of the analysis. These inputs from statistical data or provided by experts and represented as follows:





**Figure 4.5** – Mapping confidence levels on fuzzy scale.

- The crisp values derived from statistical data are mapped on the universe of discourse. This process is called fuzzification. This fuzzification gives two membership degrees to each crisp value. See the second step (characterizing input data) in the framework in Figure 4.1 the quantitative part. Figure 4.2 shows an example of event A of frequency that belongs to classes F3 and F2 with membership degrees equal 0.4 and 0.6 respectively. It should be noted that uncertainties attached to statistical values are considered to be covered by the fuzzy classes.
- Experts are asked to give verbal expressions in terms of frequency classes to such an event if statistical data is not provided. The provided class or classes are taken to be the input for the event. If the expert describes the event using one class that means the elicited frequency class is of membership 1. See Figure 4.1 the second step the qualitative input. The case of risk barriers can be an illustrative example. Experts' elicitations regarding a risk barrier can be:
  - ✓ the frequency is of one class: this barrier is of *NC2* class and thus membership is equal to 1;
  - ✓ the frequency is of two classes: this barrier is a good known technology, thus it is a good *NC2*. Or it is a not known and new technology, thus it is a bad *NC2*. Using the fuzzy scales, good or bad *NC2* means *NC2* : 0.8 and *NC3* : 0.2 (0.8 and 0.2 are is the membership degree) or *NC2* : 0.8 and *NC3* : 0.2, respectively.

Thus, a fuzzy frequency class (fuzzy number) or two membership degrees for two different fuzzy classes represent the frequency of an event as input. The same process is used to the CLs of risk barriers and the occurrence probability of ESs.

The propagation process of this representation of input data is described in the next section.

### 4.2.3 Propagating fuzzy frequencies through the Bow-Tie

This section aims to set the fuzzy rules for propagating the fuzzy inputs through the Bow-Tie analysis. Propagating inputs is achieved by solving the gates between the

events, aggregating the confidence levels of risk barriers and the occurrence probabilities secondary events. The Fuzzy semi-quantitative rules are presented as follows:

- Treatment of OR and AND gates (section 4.2.3.1 and 4.2.3.2 respectively);
- Treatment of secondary events (section 4.2.3.3);
- Treatment of security barriers (section 4.2.3.4).

### 4.2.3.1 Treatment of OR gate

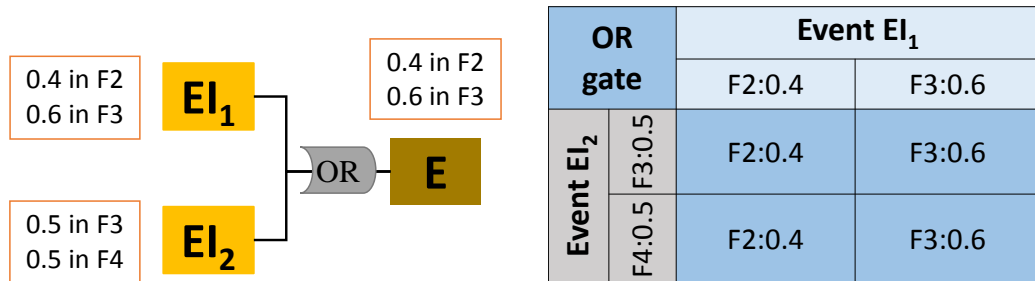
OR gate signifies that the output event occurs if either of the input events has occurred. Figure 4.6(a) presents an OR gate with two input events  $EI_1$  and  $EI_2$ ,  $E$  occurs after the occurrence of  $EI_1$  or  $EI_2$ . Based on the fuzzy approach, each one of  $EI_1$  and  $EI_2$  may be attached to one or two classes depending of the type of available data. Then, to generate all the possible combinations, the frequency of the OR gate is calculated using the Cartesian product where the Class and the possibility degree of the table cases are calculated based on the two equations below:

$$Class(E) = \min[Class(EI_1), Class(EI_2)] \quad (4.3)$$

$$Degree(Class(E)) = Degree(\min[Class(EI_1), Class(EI_2)]). \quad (4.4)$$

We based this on fuzzy rules to generate this equation [4]

For example, let  $EI_1$  and  $EI_2$  belong to [F2, F3] and [F3, F4] with possibility degrees equal [0.4, 0.6] and [0.5, 0.5] respectively. The Cartesian product for the output is presented in Table 4.6(b). If the same class  $F_i$  has different possibility degrees in the Cartesian table, then the maximum degree between them is selected for the  $F_i$ . The output frequency is [F2, F3] with possibility degrees of [0.4, 0.6].



(a) The output of the OR gate based on the fuzzy semi-quantitative approach.

(b) Cartesian product table for the OR gate example.

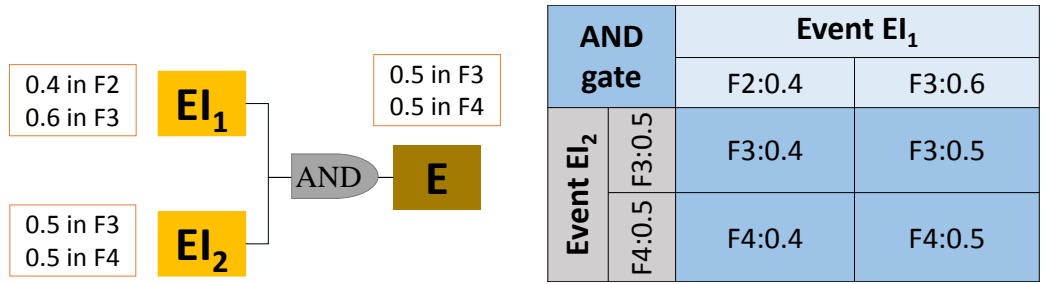
**Figure 4.6** – OR gate example.

### 4.2.3.2 Treatment of AND gate

AND gate signifies that the output event  $E$  occurs if the input events occur simultaneously. The classes and membership degrees of the output event are determined based on Eq 4.5 and Eq 4.6, respectively. An example is presented in Figure 4.7. The red rectangle beside each event in Figure 4.7(a) contains its fuzzy frequency. The output fuzzy frequency of event  $E$  is generated using the Cartesian product presented in Table 4.7(b).

$$Class(E) = \max[Class(EI_1), Class(EI_2)] \quad (4.5)$$

$$Degree(Class(E)) = \left( \min[Degree(Class(EI_1)), Degree(Class(EI_2))] \right) \quad (4.6)$$



(a) The output of the AND gate based on the fuzzy semi-quantitative approach. (b) Cartesian product table for the AND gate example.

**Figure 4.7** – AND gate example.

### 4.2.3.3 Treatment of secondary events

Figure 4.8 presents how secondary event is modeled in the bow-tie. PhD1 occurs if ERC occurs and ES occurs conditionally after ERC. PhD2 occurs after the occurrence of ERC and no occurrence of ES. The input data here are the frequency classes of the ERC and the conditional probability classes of ES (noted  $p$ ). Where  $p$  is a real number between 0 and 1. The output frequencies and degrees are calculated based on the equations below:

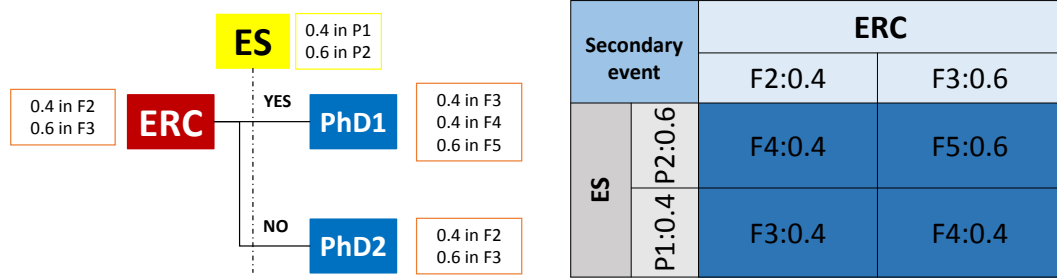
$$Class(PhD1) = [Class(ERC) + Class(ES)]; \quad [86] \quad (4.7)$$

$$Degree(Class(PhD1)) = \left( \min[Degree(Class(ERC)), Degree(Class(ES))] \right) \quad (4.8)$$

$$Class(PhD2) = [Class(ERC)]; [86] \quad (4.9)$$

$$Degree(Class(PhD2)) = Degree([Class(ERC)]) \quad (4.10)$$

The Cartesian product table and the output fuzzy frequencies of PhD1 and PhD2 are respectively presented in Table 4.8(b) and Figure 4.8(a).



(a) The fuzzy output of dangerous phenomena after the occurrence of an ES.

(b) The output fuzzy frequency of PhD1 obtained by means of the Cartesian product.

**Figure 4.8** – Consideration of an ES within the fuzzy semi quantitative approach.

#### 4.2.3.4 Treatment of security barriers

The INERIS approach consists firstly in verifying, on the basis of certain criteria, whether the security barrier can be used for the studied scenario. A security barrier operates after the occurrence of the event that this barrier is attached to. A proper functioning of a risk barrier will lead to a less dangerous complementary event ( $\bar{E}$ ), see Figure 4.9(a). In the other case (if the security barriers does not operate), another more dangerous but less probable complementary event  $E$  will occur.

The fuzzy classes of the output events  $E$  and  $\bar{E}$  are calculated based on the Cartesian product. The Frequency class and degree of each case in the Cartesian table is determined using the equations below:

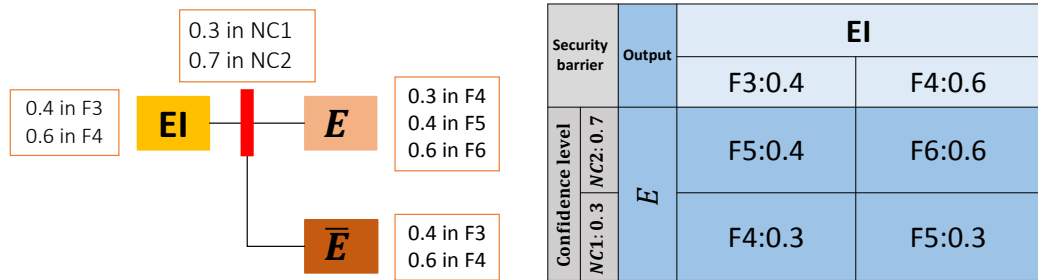
$$Class(E) = Class(EI) + NC = F(x + NC); \text{ where } x \text{ is the class of EI} \quad (4.11)$$

$$Degree(Class(E)) = \min[Degree(Class(EI)), Degree(NC)] \quad (4.12)$$

$$Class(\bar{E}) = Class(EI) \quad (4.13)$$

$$Degree(Class(\bar{E})) = Degree([Class(EI)]) \tag{4.14}$$

Figure 4.9 presents an example of a risk barrier treatment. The fuzzy inputs are plotted in the rectangles beside the EI and above the risk barrier in Figure 4.9(a). The Cartesian product of the event  $E$  is shown in Figure 4.9(b), where the fuzzy frequency of  $\bar{E}$  is the same as  $EI$ .



(a) The fuzzy outputs of events  $E$  and  $\bar{E}$ . (b) The output fuzzy frequency of  $E$  and  $\bar{E}$  obtained by means of the Cartesian product.

**Figure 4.9** – Consideration of a security barrier within the fuzzy semi quantitative approach.

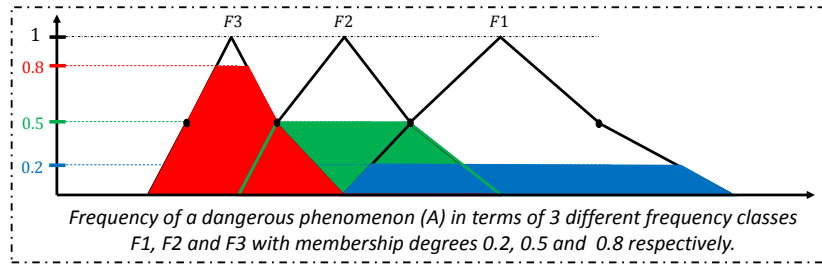
#### 4.2.4 Decision making under fuzzy environment

The result obtained from the propagation is a fuzzy variable, which may cover different frequency classes with different membership degrees. In this section we will explain the different ways of how the decision can be made based on the fuzzy output. It should be noted that the decision is also related to engineers who are making the decision and the type of the dangerous phenomenon being studied (if the phenomenon is high risk or not).

Here are the different options that are possible:

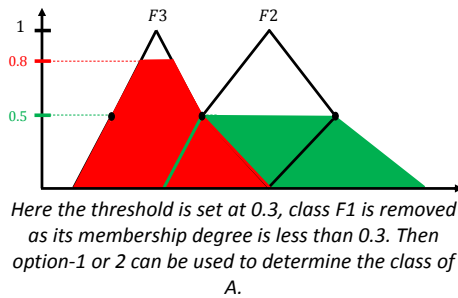
- Choosing a membership threshold - this threshold is set from 0 to 1 where frequency classes with membership degrees less than the chosen threshold will be eliminated. Then, decision-makers can return to choices 1 or 2 below to make their decision.
  1. being conservative - in this case decision-makers decide to be conservative (pessimistic) and take the highest frequency class;
  2. being realistic - here decision-makers take the frequency class with the highest membership degree as it represents the most likely class that the true frequency value lies in.

It should be noted that membership threshold can be set to zero if decision makers want to be conservative 100%. Figure 4.10 shows an example of the different decision options that can be used.

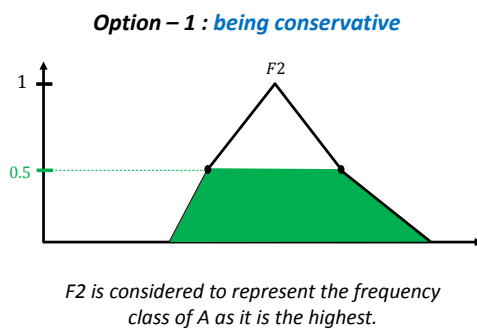


Set up a membership degree threshold

Step-1



Step-2



Option – 2: most possible

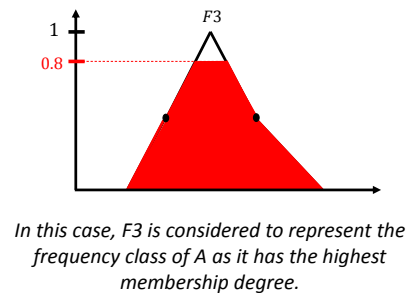


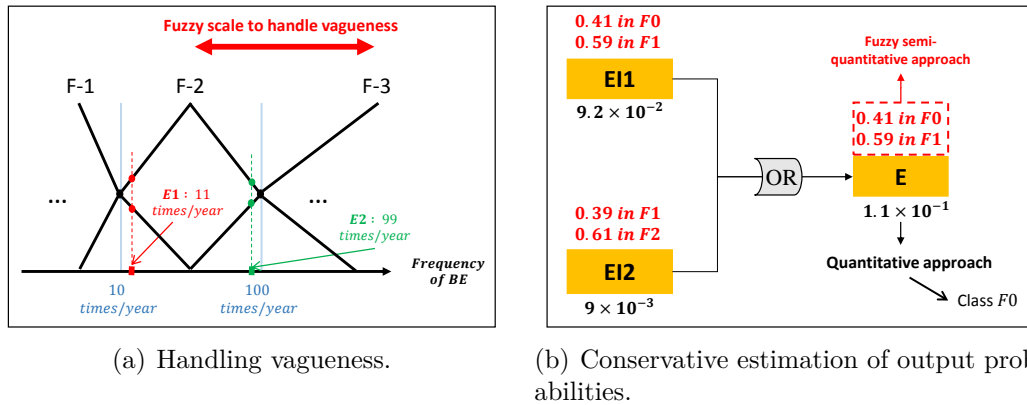
Figure 4.10 – Decision making under fuzzy environment.

### 4.2.5 Handling the existing limitations of the interval approach

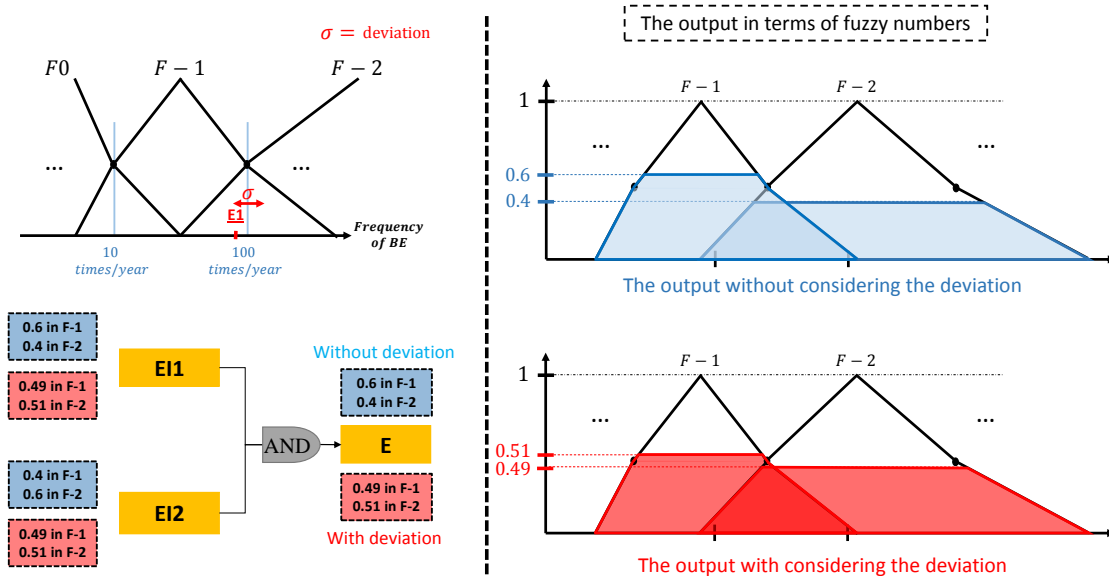
This section presents how the introducing of fuzzy concept deals with the limitations 1, 2 and 3 of the interval semi-quantitative approach discussed in Section 2.6.1. The improvements that solve problems 1, 2 and 3 are presented below in Solutions 1, 2 and 3, respectively.

- *Solution-1*: for the discreteness issue, two events with different frequencies now belong to different classes with different degrees (see events  $E1$  and  $E2$  in Figure 4.11(a)).
- *Solution-2*: Figure 4.11(b) presents how the fuzzy approach solves the probability underestimation problem. The output frequency is of classes  $F0$  and  $F1$ , which cover the output of the quantitative approach.

— *Solution-3*: a deviation due to error in the determination of input frequencies will not affect the result nor the decision. The same example taken in Section 2.6.1 is depicted in Figure 4.12. A small change in the possibility degrees is generated due to the deviation where the classes are the same. Here, lies the effectiveness of fuzzy theory in handling uncertainty.



**Figure 4.11** – How the fuzzy semi-quantitative approach deals with the limitations mentioned in Section 2.6.1.



**Figure 4.12** – A deviation will not lead to different result that affects the decision.

This methodology will be illustrated in the next section and applied to a loss of containment scenario.

### 4.3 Case study

In this section, a bow-tie analysis for a LOC scenario is utilized to prove the utility and effectiveness of the proposed methodology. This bow-tie is obtained from a risk analysis study done by the INERIS for a legislation demand. The risk analysis is affiliated to a tank farm of flammable solvents. Several solvents are implemented (methanol, iso-octane, acetone, etc.). The main risk event is the spreading of these solvents in the atmosphere. Causes and consequences of this main event were identified and the bow-tie was constructed as presented in Figure 4.13. The goal is to determine the probability of dangerous phenomena and major accidents.

The input data were collected from two different data bases (GT-DLT - [64] and BEVI - [29]) and an expert in the field. Input for the analysis are extracted and calculated as follows:

- frequencies of basic events: Five basic events were identified as presented below.
  - ✓ basic events 1 & 2 - overfilling containers: This event may occur during unloading the trucks to fill the tanks, eight tanks are concerned. It is caused from the no re-circulation of substance from tank to tank due to industrial waste from workshops. Two basic events are distinguished here depending on the presence or not of persons (see Figure 4.13, the first two basic events). Based on GT-DLI, the frequency of overfilling is  $5.10^{-4}/\text{container}/\text{year}$ . Eight containers are concerned, then frequency is equal to  $4.0 \times 10^{-3}$ ;
  - ✓ basic event 3 - leak in tanks: the value suggested by the BEVI for leak in a tank ( $1.1 \times 10^{-4}/\text{tank}/\text{year}$ ) is used. Eight vessels are concerned, then the frequency value of this basic event is equal to  $8.8 \times 10^{-4}$ ;
  - ✓ basic event 4 - leak in pumps: The value presented in BEVI ( $4.5 \times 10^{-3}$ ) for leak in a pump is used. Four pumps are presented in the farm, two are used continuously for recycling ( $8760 \text{ hours}/\text{year}$ ). The other two are used to transfer solvents where one is used  $190 \text{ hours}/\text{year}$  and the other  $395 \text{ hours}/\text{year}$ . Thus, the input frequency value for this basic event is as calculated below:

$$\text{Frequency} = (4.5 \times 10^{-3}) * (1 * 190 + 1 * 395 + 2 * 8760) = 9.3 \times 10^{-3} \quad (4.15)$$

- ✓ basic event 5 - leak in pipes: The value suggest by the BEVI depends on the length of the pipe ( $6.0 \times 10^{-6}$ ). Different pipes for different missions are used. Pipes for recycling of length  $20m$  that are full time used ( $8760\text{hours}/\text{year}$ ) and



others are less in use with length of  $20m$ . The frequency value is:

$$Frequency = (6.0 \times 10^{-6}) * (20 * 8760 + 20 * 1620) = 1.4 \times 10^{-4} \quad (4.16)$$

- probabilities of secondary events: probability of ignitions are determined based on data from GT-DLI and of values equal 0.1;
- CLs of safety barriers: CLs of safety barriers are either elicited from the expert or from BEVI. Four risk barrier are presented:
  - ✓ barrier 1 - a human barrier on the overflowing event with presence of persons that concerns of closing the valve if anything wrong happens. The CL of this barrier is elicited from the expert who suggested CL1;
  - ✓ barriers 2 & 3 - sensors that automatically close the valve in the case of overflowing with probability of failure equal to  $9.1 \times 10^{-1}$ ;
  - ✓ barrier 4 - a gas detector with a probability of failure equals  $1.0 \times 10^{-1}$ .

These inputs are translated into fuzzy classes (the dashed rectangle beside each event or risk barrier). These fuzzy classes are propagated through the Bow-Tie using the fuzzy rules set in Section 4.2.3. The output fuzzy frequencies of the ERC and outcomes are written on the bow-tie (the red dashed rectangles in Figure 4.13).

In addition to the proposed approach, quantitative and interval semi-quantitative analyses were also performed for the same bow-tie. In order to compare these approaches, the outputs are presented in Table 4.1. The output probabilities of the quantitative approach are translated into classes for comparison purposes. The quantitative approach does not consider uncertainty in the analysis. As there is no consideration of uncertainty, this may lead to risk underestimation in some cases. However, the fuzzy approach presents more accuracy than the interval approach where the output fuzzy classes cover the class obtained using the quantitative approach. Fuzzy approach is more conservative than the quantitative approach as uncertainty is considered (fuzzy numbers are used instead of crisp values). Again the result from the quantitative approach lies within the result obtained by the proposed approach, which makes the later more conservative.

For more clarity, let us make a decision about the **formation of toxic gas dangerous phenomenon**. If we want to be conservative based on the fuzzy semi-quantitative approach, then class  $F1$  is considered which is of probability level equal to A (see Table 2.6). While  $F2$  (probability level equals B ) and  $F3$  (probability level equals C) are respectively represent the frequency using the quantitative and the interval-based approach. This example explain how fuzzy approach is more conservative than the two other approaches.

It should be noted that this difference between those three approaches is not often the case. The fuzzy-based approach has been applied to different bow-ties were the same

Outcomes	Quantitative approach	Interval semi-quantitative approach	Fuzzy semi-quantitative approach
Spreading solvents in a large quantity	$2.3 \times 10^{-3}$ :F2	F3	0.42 in F1 0.5 in F2 0.63 in F3
Explosion UVCE	$2.3 \times 10^{-4}$ :F3	F4	0.42 in F2 0.5 in F3 0.63 in F4
Formation of toxic gas	$2.3 \times 10^{-3}$ :F2	F3	0.42 in F1 0.5 in F2 0.63 in F3
Pool fire	$2.3 \times 10^{-4}$ :F3	F4	0.42 in F2 0.5 in F3 0.63 in F4
Toxic effect	$2.3 \times 10^{-3}$ :F2	F3	0.42 in F1 0.5 in F2 0.63 in F3

Probability levels, see Table 5

	D: Unlikely		C: Moderate		B: Likely		A: Very Likely
--	-------------	--	-------------	--	-----------	--	----------------

**Tableau 4.1** – Estimated results obtained using quantitative, interval semi-quantitative and fuzzy semi quantitative approaches.

result as the quantitative and the interval-based approach in term of final probability level were obtained.

## 4.4 Conclusion

Probability analysis of dangerous phenomena has become a necessary step in risk analysis. Qualitative or quantitative probability analysis can be performed depending on the type of data available. This data is derived from different sources (historical accident data or expert judgments in terms of numerical values or linguistic variables, respectively). Quantitative information for a quantitative analysis is expensive and not always provided. Qualitative analysis is subjective and may lead to loss of quantitative information if it exists. In addition, the accuracy of the analysis based on these approaches still a major issue since uncertainty is not taken into consideration. That is why this chapter proposes a fuzzy-semi quantitative approach relying on the available information from historical data or experts if the former is not available. Fuzzy theory is introduced to handle uncertainty due to imprecision and vagueness in defining the frequency scale. Fuzzy rules are set to propagate the fuzzy input classes through the Bow-Tie analysis.

This methodology is applied to a Bow-Tie case study for a LOC scenario. A comparison with the quantitative and the interval semi-quantitative approaches is discussed.

The results show that the proposed methodology provides more simplicity and accuracy in the quantification, in addition to the consideration of uncertainty.

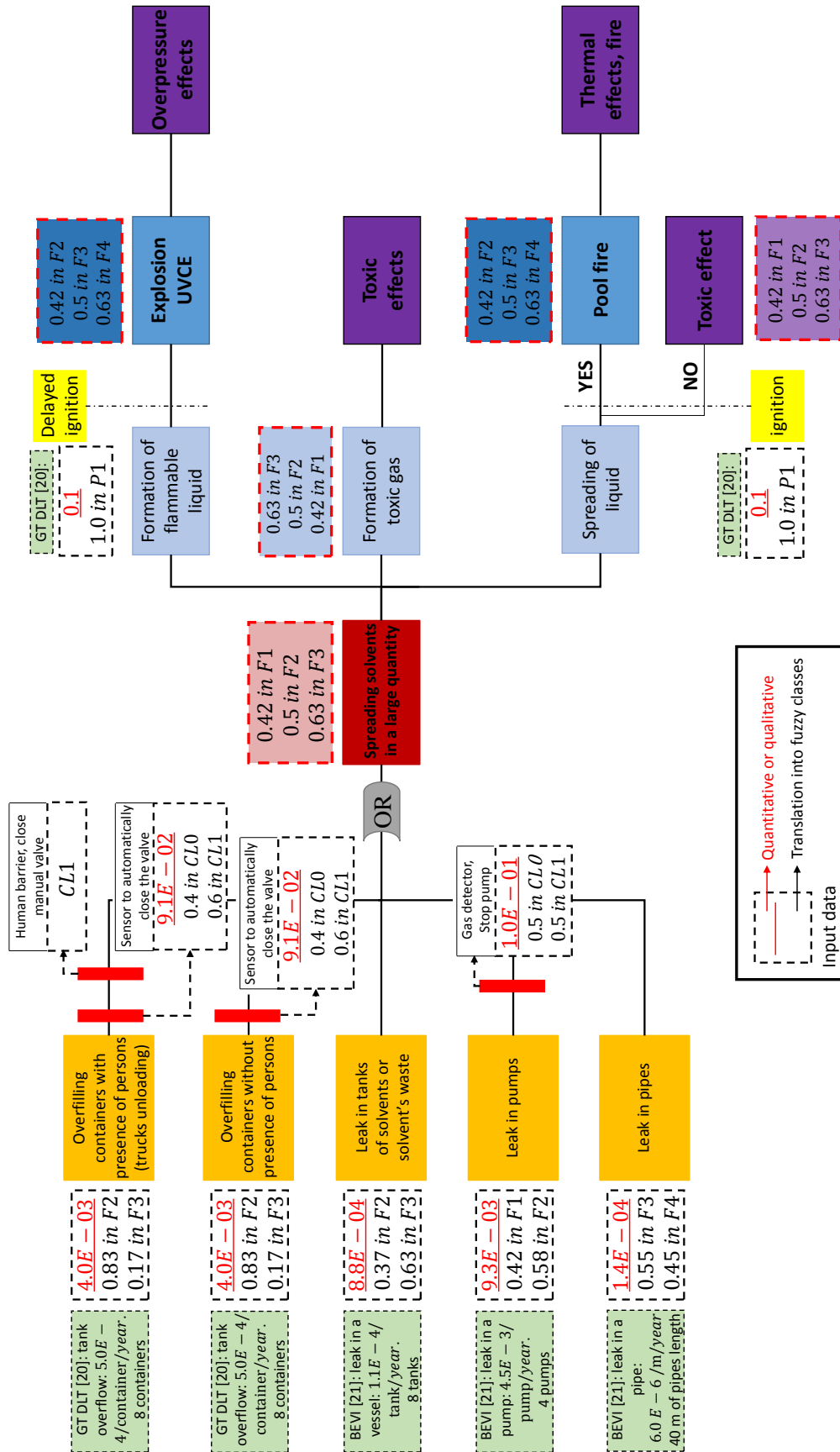


Figure 4.13 – The bow-tie diagram of the LOC scenario under study.



# 5

## Treatment of aleatory and epistemic uncertainties in analyzing the effect of risks

**Summary:** Quantifying uncertainty during risk analysis has become an important part of effective decision-making. In this chapter, we will present global approaches to treat uncertainty in effect analysis. A fuzzy-probabilistic approach to treat imprecision and variability separately is proposed. A global approach to treat all causes of uncertainty with the best theory is developed.

### Summary

---

<b>5.1</b>	<b>Introduction</b>	<b>104</b>
<b>5.2</b>	<b>A new uncertainty analysis approach with randomness and fuzzy theory to deal with variability and imprecision</b>	<b>105</b>
5.2.1	Introduction	105
5.2.2	Problem statement	105
5.2.3	Proposed fuzzy-probabilistic approach	107
5.2.4	Uncertainty representation	110
5.2.5	Uncertainty propagation	111
5.2.6	Case study - Application of the proposed hybrid approach to effect analysis with considering of parameter uncertainties	116
5.2.7	Description	116
<b>5.3</b>	<b>Comparison of the proposed hybrid approach with the existing approaches: applying uncertainty analysis to a Loss Of Containment scenario</b>	<b>131</b>
5.3.1	Models used in the analysis	132
5.3.2	Calculation of the concentration without considering uncertainty	133

5.3.3	Uncertainty modeling . . . . .	135
5.3.4	Uncertainty analysis using interval analysis . . . . .	136
5.3.5	Uncertainty analysis using the fuzzy approach . . . . .	136
5.3.6	Uncertainty analysis using the probabilistic approach . . . . .	139
5.3.7	Uncertainty analysis using Evidence theory . . . . .	140
5.3.8	Uncertainty analysis using the probabilistic-fuzzy approach . . . . .	142
5.3.9	Comparison of all approaches . . . . .	144
<b>5.4</b>	<b>A global approach to treat all types and causes of uncertainty in effect analysis: the ALI-Aggregated Likelihood Index . . . . .</b>	<b>148</b>
5.4.1	Uncertainty representation: characterizing each uncertain parameter based on the available information . . . . .	149
5.4.2	Uncertainty propagation . . . . .	151
5.4.3	Decision Making under uncertain environment . . . . .	153
5.4.4	Case study - Applying of the developed approach to a loss of containment scenario . . . . .	154
<b>5.5</b>	<b>Conclusion . . . . .</b>	<b>159</b>

---

---

This work was carried out in collaboration with the INERIS. The results developed in this chapter have been presented in the following articles:



[4] ABDO, HOUSSEIN AND FLAUS, JEAN-MARIE. *Uncertainty quantification in dynamic system risk assessment: a new approach with randomness and fuzzy theory*. International Journal of Production Research 54, 5862-5885 (2016).



[7] H. ABDO, J-M. FLAUS, AND F. MASSE. *Uncertainty quantification in risk assessment - Representation, propagation and treatment approaches: Application to atmospheric dispersion modeling*. Journal of Loss Prevention in the Process Industries.



[2] H. ABDO, AND J-M. FLAUS. *A mixed fuzzy probabilistic approach for risk assessment of dynamic systems*. In *15th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2015*, Ottawa, Canada (2015).



[2] H. ABDO, AND J-M. FLAUS. *Uncertainty quantification in bow-tie analysis: A mixed approach of fuzzy theory with Dempster-Shafer theory of evidence*. In *Risk, Reliability and Safety: Innovating Theory and Practice: Proceedings of ESREL*, Glasgow, Scotland (2016).



## 5.1 Introduction

In this chapter we are going to propose uncertainty analysis methodologies to treat the two types of parameter uncertainty separately. The first methodology is a hybrid fuzzy-probabilistic approach that deal with the two most faced causes of parameter uncertainty: variability and imprecision. The second methodology is an extension of the first methodology where Dempster-Shafer theory of evidence is introduced to deal with ignorance, incomplete information and lack of consensus. This methodology represents a global, exhaustive approach that can treat all causes of parameter uncertainty if affect the same analysis with the best representations.

Section 5.2 presents the hybrid fuzzy-probabilistic approach. In Section 5.2, first, we will start by a review on existing hybrid approaches that separately deal with epistemic and aleatoric uncertainties. Then we present the limits of these existing approaches and why a new hybrid approach with a new propagation algorithm is needed. The proposed methodology uses probability theory, fuzzy numbers and fuzzy random variables to represent variability, imprecision and if these two causes of uncertainty affect the same input parameter, respectively. Monte Carlo simulations are performed to propagate these representations through the risk models. The proposed hybrid approach is applied to an over-pressure scenario in a propylene oxide polymerisation reactor to prove its utility and effectiveness.

In Section 5.3, the proposed hybrid approach and the approaches reviewed in Chapter 3, Section 3.1 (not-hybrid) are applied to a case study of an atmospheric dispersion scenario. This application aims to compare between the hybrid and the existing approaches and to prove why different causes should be treated separately with different representations. Based on this comparison, a guidance on how parameter uncertainty should be treated is provided.

The last part of this Chapter develops a methodology that cover an overall uncertainty analysis approach and a decision making framework under parameter uncertainty (Section 5.4). This methodology aims to treat each cause of parameter uncertainty with the best suitable mathematical tool. In this methodology, probability theory is used to represent variability, fuzzy numbers are used to represent imprecision and evidence theory is used to represent ignorance, incompleteness and the lack of consensus. The represented parameters are propagated using a 2-stages MC simulation. The proposed methodology is applied on a loss of containment scenario (LOC). The objective is to calculate the toxic effect at a specific end point with consideration of parameter uncertainty. Finally, the developed methodology is discussed and compared against the pure probabilistic approach.

It should be noted that the proposed approaches can be used in any field to solve any

problem regarding parametric uncertainty. The provided guidelines will help risk analysts at any domain to treat parameter uncertainty that affect the input parameters with the best representation, the least time consuming and the most precise calculation based on the available data at hand.

## 5.2 A new uncertainty analysis approach with randomness and fuzzy theory to deal with variability and imprecision

### 5.2.1 Introduction

In this section, we propose a hybrid random-fuzzy approach for capturing uncertainty during risk analysis. The approach proposed here is not the first hybrid fuzzy-probabilistic approach. The hybrid approach is a new approach proposed in recent studies by ([168]; [24]; [74]; [95]). It uses a combination of probability and possibility theories to address uncertainty related to model parameters. In [14], the authors argue that probabilistic methods are not sufficient to represent epistemic uncertainty. Therefore, this approach aims to represent variability due to aleatory uncertainty and the vagueness and imprecision related to epistemic uncertainty using probability distributions and fuzzy numbers respectively. The method presented in this study is considerably different from the existing mixed approaches in that random fuzzy numbers are introduced and a new uncertainty propagation algorithm adapted to all cases and to dynamic systems is proposed. The algorithms of existing methods can produce results that are wrong. A detailed explanation of the advantages of this approach will be given later.

In order to present the possibilities offered by this mixed approach: in section 2, we present the problem of not treating epistemic and aleatory uncertainties separately; in section 3, we present the proposed hybrid approach; in section 4, we apply the proposed approach to a case study in a chemical reactor.

### 5.2.2 Problem statement

We shall now introduce the problem and discuss the overall methodology behind the proposed solution. The problem has been crafted to enable us to focus on the issues of risk analysis when calculating the severity of risk taking into account parameter uncertainty.

Very often, the parameters influencing the risk of exposure to failure in industry are fraught with uncertainty. One method commonly used to address this uncertainty is the Monte Carlo method, which applies probability theory and relies on a statistical repre-

presentation of available information. Our challenge is how to represent uncertainty relating to parameters when there is insufficient information available for statistical identification (epistemic uncertainty). This is the problem we shall study when the risk model input values are imprecise.

As mentioned in Chapter 1, calculating the effect of risk is most often described by a mathematical model. This mathematical representation is affected by both types of uncertainty. Uncertainty can occur in the parameters of the mathematical model. The particular context of our study is shown in Figure.5.1.

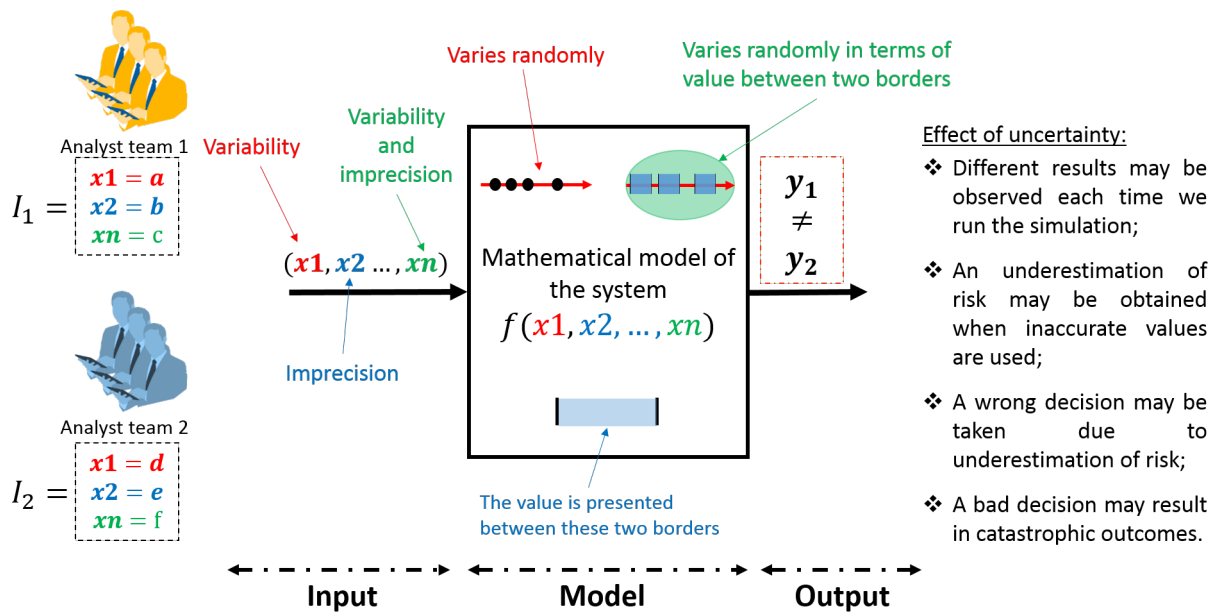


Figure 5.1 – Uncertainty presented in the analysis and its effects.

This figure shows a mathematical model for calculating the effect of a risk. We consider that the system model is made up of: input parameters  $(x_1, x_2, \dots, x_n)$  derived from two analyst teams, the model equation (the mathematical model used by the simulation), and outputs. The challenge is to address the issues relating to the representation of parameter-related information (see Figure 5.1). However, this information may not necessarily be accurate. It may be based on experimental data and measurement, and therefore be known statistically, or based on expert opinion and therefore be known with a certain degree of inaccuracy. Both types of uncertainty can influence parameters. In addition, while some parameter values are known statistically, the mean or standard deviation of their distribution is uncertain, i.e. it is known in terms of intervals. This means that these parameters are affected by aleatory and epistemic uncertainty.

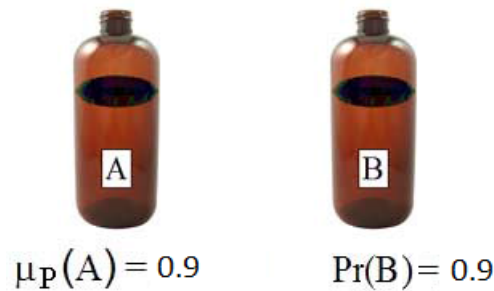
### 5.2.3 Proposed fuzzy-probabilistic approach

In this section we present the structure of the proposed approach. This approach will be applied in the next section on a chemical process example. Before proceeding, it is significant to note that, as we mention in the introduction, this approach is not the first to mix probability and fuzzy numbers. However, existing approaches are based on the Monte Carlo method combined with fuzzy calculus to propagate parameter uncertainties through the risk model. Fuzzy calculus using the  $\alpha$  – *cut* method is the same as interval analysis but for different levels of likelihoods. The existing methods can lead to incorrect results in some cases. The first is when computer programs are used for the analysis, they can act as black boxes (model equations inside these programs are not known). Knowledge relating to the input parameters of these black boxes can be uncertain. Interval analysis can be problematic when the program’s models are hidden. For example, suppose the program hides the expression  $F = 1/y$ , then the lower bound of the input must be used to calculate the upper bound of result. A second issue which is often presented in the dynamic models when they are not-monotonic [59]. For example,  $f(y) = (y - 1)^2$  and  $y = [0, 4]$ ; using interval analysis the result of this equation is calculated using the two bounds of the inputs interval ( $f(0) = (0 - 1)^2 = 1$  and  $f(4) = (4 - 1)^2 = 9$ ). Hence, the result is equal to  $[1, 9]$ . The true lower bound of the output is obtained from the input value  $y = 1$  ( $f(1) = (1 - 1)^2 = 0$ ) which is presented inside the input interval. Thus, the non-monotonicity requires the entire interval to be considered. In this case, simulation methods such as the Monte Carlo method would be useful. Other techniques also exist, such as the method proposed in [117].

#### 5.2.3.1 Difference between probability and possibility

Before going into detail, it is important to explain the difference between fuzzy logic and randomness. We will do this with the help of the bottle example shown in Figure 5.2. The figure shows two bottles of water, A and B, where ‘A’ has a membership value of 0.9 in the set of all potable liquids, while bottle ‘B’ has a 0.9 probability of containing potable liquid. Let us assume that a thirsty person happens upon these two bottles of water with their uncertainty in relation to their suitability for drinking. The question is which one should the person choose?

The rational decision is to choose bottle ‘A’, since a 0.9 membership means that 0.9 of its content is potable liquid, making it almost perfectly potable. Since ‘B’ has a 0.9 probability of containing potable liquid, this means there is a 10% chance that the contents are swamp water or lethal. In other words, there is a 1 in 10 chance that anybody who drinks it may be poisoned. Let us assume that after sampling, A contains a drinkable



**Figure 5.2** – Two bottles of liquid with their uncertainty representations.

liquid (juice) while B contains a toxic liquid. This means that the membership value remains the same, whereas the probability value of B is resolved from 0.9 to 0. The example therefore shows that a probability value is not the same as a fuzzy value.

### 5.2.3.2 Why a different new hybrid approach is needed

As seen, different mathematical theories are used in this approach when most uncertainty analysis focus on pure probabilistic approaches. Recently, several researchers have agreed that epistemic and aleatory uncertainty must be treated separately ([16]; [168]; [17]). [60] have demonstrated that, due to the vagueness, the treatment of epistemic uncertainty using probability theory may result in an unconservative estimation (minimization) of risk when additional and unjustified information is provided for the output result.

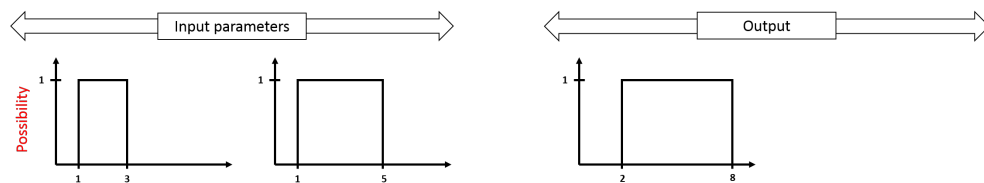
To understand this further, let us consider the example taken by [60] adding a number of modifications to pinpoint the importance of using a mixed or hybrid approach for dealing with uncertainties in the risk assessment process. Suppose that  $f = A + B$ , where  $A$  and  $B$  are two numbers somewhere between  $\{1, 3\}$  and  $\{1, 5\}$ . This represents the only available information about these two parameters. At first, the treatment of uncertainties related to these two parameters using the possibility theory aims to give a possibility distribution in terms of fuzzy numbers for  $A$  and  $B$ . When there is no information about the shape of the distributions, the same degree of possibility equal to one is given to each element inside the two intervals (see Figure 5.3(a)). The addition of the fuzzy numbers  $A$  and  $B$  is the interval  $[2, 8]$  with a constant membership function (possibility distribution) equal to one for all elements within this interval.

On the other hand, a probabilistic representation of these parameters is given in terms of uniform probability distributions for both variables. Choosing any other distribution would involve additional unjustified information about the parameters. The probability density function and the cumulative distribution of the addition are shown in Figure 5.3(b). The result of this case tells us that the addition must lie between 2 and 8, as in

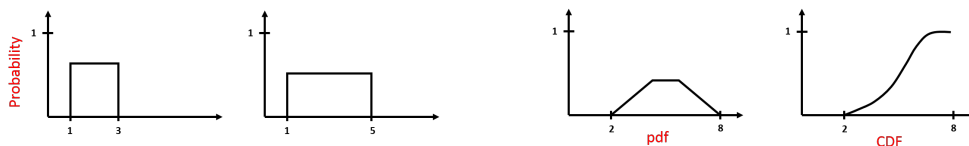
the previous case, but has a greater chance of being a value close to the central tendency. At 95 percentile of probability, the addition would be less than or equal to 7 (risk decisions are mostly made at 95 of probability).

In the third case (see Figure 5.3(c)), we represent  $A$  using a uniform probability distribution while  $B$  is represented using a possibility distribution. The result is depicted as a bunch of fuzzy numbers (a fuzzy number is generated for each sample from the probability distribution). The fuzzy cumulative distribution is also shown in the same Figure. In this case, at 95 percentile of probability would be the fuzzy number  $[3.9, 7.9]$  with the possibility equal to 1 in the entire interval (the values obtained in the three methods are accurate obtained from a real simulation).

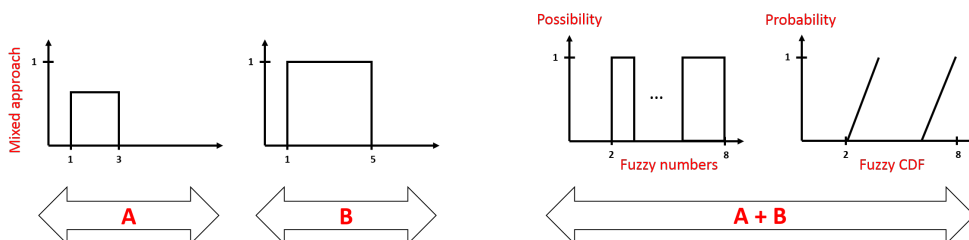
Figure 5.3 clearly shows that handling uncertainty using a purely probabilistic approach may result in an underestimation of the risk when insufficient information is available to build the distribution and assumptions must be made about the missing information. On the other hand, if all the parameters were represented by fuzzy numbers, despite the fact that some could be justifiably represented by PDFs, the range of results would be too conservative. Thus, a mixed approach offers an effective alternative between the missing information and the risk underestimation.



(a) Representing and propagating uncertainty in  $A$  and  $B$  using the fuzzy approach.



(b) Representing and propagating uncertainty in  $A$  and  $B$  using the probabilistic approach.



(c) Representing and propagating uncertainty in  $A$  and  $B$  using the mixed probabilistic-fuzzy approach.

**Figure 5.3** – Addition of  $A$  and  $B$  with the consideration of uncertainty.

As highlighted above, the mixed approach uses all the available information and ensures that the risk calculation is conservative. This ensures that the true value of the risk is between the calculated bounds. It should be noted that in risk analysis and in addition to the precise parameters that may be provided, others vary randomly when we can observe the frequency of their values. Furthermore, knowledge relating to other parameters is either poor or inexistent due to imprecision and vagueness. Frequencies, if available, provide a structure on which to build probability distributions, while a lack of information often restricts the performance of the analysis. In this situation, expert elicitation provides an alternative to collecting more information for the analysis ([35]; [57]; [53]; [54]). Several techniques have been proposed to address uncertainty in relation to expert judgment and the fuzzy representation of expert elicitation is one of the most popular [158]. [19] propose a chart to define the elicitation of experts in terms of triangular fuzzy numbers where the only required data are the most likely values ( see related articles by [19]; [107]; [136]). Thus, all these reasons approve why an alternative mixed approach will result in a better estimating and understanding of uncertainty in risk analysis.

#### 5.2.4 Uncertainty representation

The first step is to represent uncertainty. The types of uncertainty can be found and represented as follows:

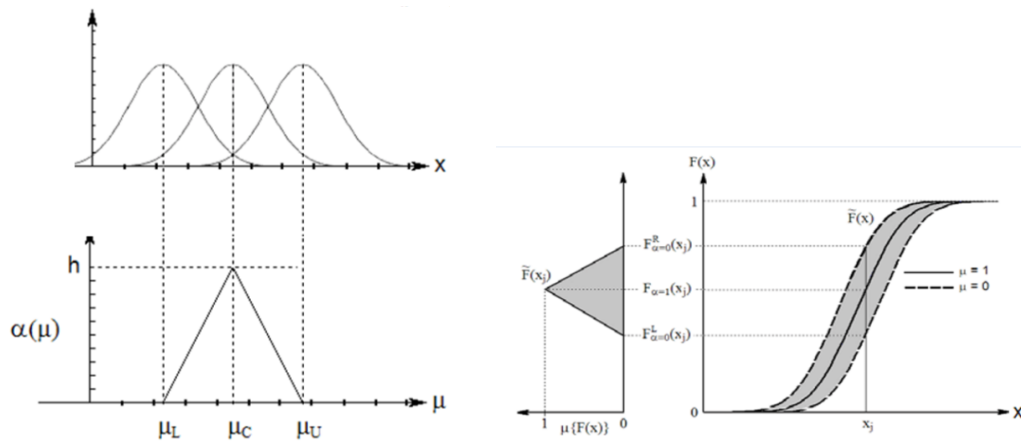
1. aleatory uncertainty will be represented by random variables based on probability theory;
2. epistemic uncertainty due to imprecision will be represented in terms of fuzzy numbers;
3. the third case can be observed when both types of uncertainty affect the same parameters. In this special case, fuzzy random variables are used to represent the mixture of both types. The concept of Fuzzy Random Variables is detailed in the rest of this section.

Randomness and fuzziness can be merged to formulate a fuzzy random variable (FRV). An “FRV can be seen as a random variable taking fuzzy values”. An FRV can be represented by its set of membership functions and its associated probabilities. For example, for two alternatives with probabilities  $p_1$  and  $p_2$ :

$$FRV = \begin{cases} (h_1, h_2, h_3) & \text{with probability } p_1 \\ (b_1, b_2, b_3) & \text{with probability } p_2 \end{cases} \quad (5.1)$$

FRVs are used when there are sources of uncertainty that random variables or fuzzy numbers cannot accommodate alone. This is the case, for example, when a parameter is known but with a certain degree of imprecision. In other words, it has a range of values and that range of values can have a random variation.

A detailed presentation of concepts and definitions relating to the theory of fuzzy random variables is given by [144]. A fuzzy random variable  $X$  might be described mathematically by a fuzzy probability distribution function  $\tilde{F}(x)$ . Figure 5.4(a) presents an example of a PDF (Probability Density Function), where the distribution mean is a fuzzy number (Triangular Fuzzy Number) which gives a PDF for each  $\mu$  between  $\mu_L$  and  $\mu_U$ . The lower, upper and middle PDF (greater membership value) are represented in the same figure. The cumulative fuzzy distribution function is presented in figure 5.4(b). The figure also shows the fuzzy version of  $F(x)$  while the solid line represents  $F(x)$  for  $\mu = 1$ , and the dashed lines  $\mu = 0$ . The figure also shows the membership associated with  $\tilde{F}(x_j)$  for a given  $X_j$ .



(a) The fuzzy mean of the distribution. (b) The fuzzy cumulative distribution function.

**Figure 5.4** – Representation of a fuzzy random variable.

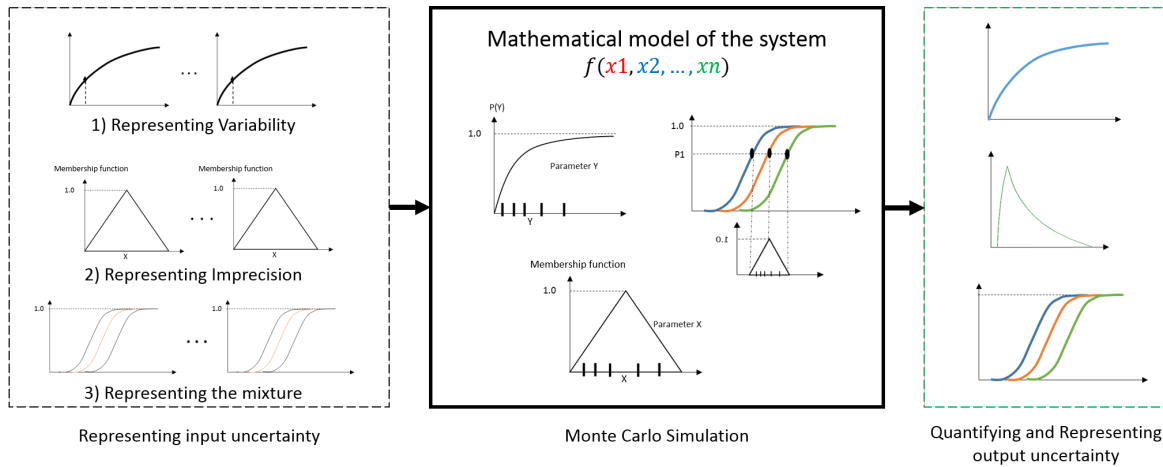
Thus, probability distributions, fuzzy numbers and fuzzy random variables can be used to represent uncertainty related to inputs as well as to the output. A simulation approach based on the Monte Carlo analysis can be used to propagate uncertainty in the input as described in what follows.

### 5.2.5 Uncertainty propagation

Figure 5.5 shows how we incorporate the three types of uncertainty in the risk analysis of dynamic systems and the general algorithm used for the simulations. The dynamic system is described by a mathematical model. The parameters of this model are described



by probability distribution, fuzzy numbers or fuzzy random variables according to the type of uncertainty, which may be epistemic or aleatory or a mixture of both types.

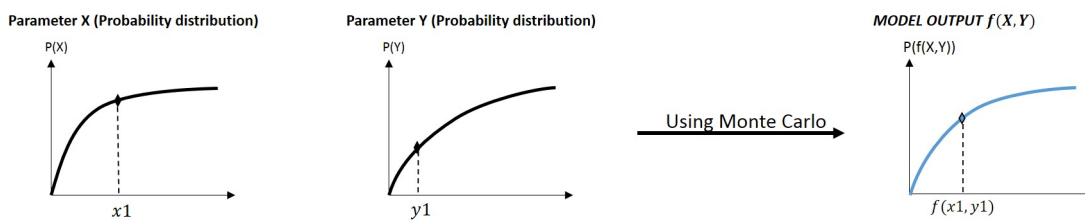


**Figure 5.5** – Representation and propagation of uncertainty.

In order to explain our algorithm, we shall start by presenting the MC simulation for random variables. Following this, we shall show how this is extended to propagate the fuzzy numbers. We shall then use the same approach in a 2-stages scheme to propagate the fuzzy random variables. A detailed description of these propagation techniques is presented in the next three sections.

### 5.2.5.1 Monte Carlo to propagate Random Variables

In this case, where we take into consideration the existing aleatory uncertainty, the system's parameters are described by random variables in terms of probability distributions (see Figure 5.6). The main simulation steps are listed in Chapter 3, Section 3.1.2.



**Figure 5.6** – Use of Monte Carlo to estimate the output probability distribution.

### 5.2.5.2 Monte Carlo to propagate uncertainty described by fuzzy numbers

The principle of the simulation is based on the Monte Carlo method that relies on repeated random sampling. For the parameters described by a fuzzy number, sampling is performed according to a uniform law or a Gaussian law to give more or less weight to

the extremities. Then the membership function is associated with the sampled value (see Figure 5.7). To understand the basics of this propagation, let us assume that  $Z = f(Y) = f(X_1, X_2, \dots, X_n)$ . This represents the model of a system together with its uncertain variables, which are represented by fuzzy numbers. The propagation steps are listed below:

- set  $i = 1$  and  $N =$  number of samples;
- for each imprecise parameter represented by fuzzy numbers, take a value with its membership degree using MC (these values and the related membership values represent the input parameters in the model system equations) where for each sample  $y_i = (x_1, x_2, \dots, x_n)$  and  $\alpha(y_i) = (\alpha(x_1), \alpha(x_2), \dots, \alpha(x_n))$  are taken and  $\alpha$  is the membership degree;
- calculate the value  $z_i = f(y_i) = f(x_1, x_2, \dots, x_n)$ , where the membership of  $z_i$  is obtained using the extension principle of fuzzy numbers ( $\alpha(z_i) =$  minimum of  $(\alpha(x_1), \alpha(x_2), \dots, \alpha(x_n))$ , if  $z_i$  has been obtained earlier from another sample then the maximum membership between it and the old membership is given to  $z_i$ );
- if  $i < N$  set  $i = i + 1$  then return to step 2. Otherwise, go to 5;
- after n MC samples, n values are obtained with their possibility degrees that generate the fuzzy result, see Figure 5.7.

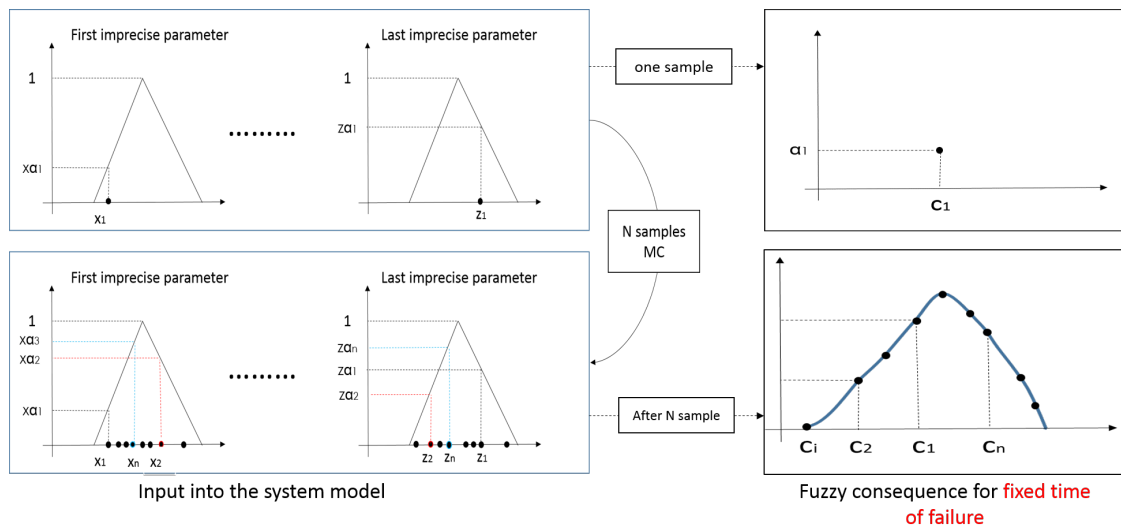


Figure 5.7 – Use of Monte Carlo to estimate the fuzzy result.

### 5.2.5.3 2D Monte Carlo to propagate Fuzzy Random Variables

Existing approaches cannot be used on this complex case. Instead, a two-dimensional Monte Carlo simulation is performed using the system's parameters in order to estimate

the random fuzzy output. Figure 5.8 and the framework in Figure 6.2 show the steps required to apply the proposed method: let us assume that  $Z = f(Y) = f(A_i, B_j, C_k) = f(A_1, \dots, A_n, B_1, \dots, B_m, C_1, \dots, C_q)$  representing the model of the system where  $A_i, i = 1, \dots, n$  are affected by aleatory uncertainty,  $B_j, j = 1 \dots m$  by epistemic and  $C_k, k = 1, \dots, q$  by a mixture of both.

1. set  $v = 0, w = 0$  and  $M, N$  are the numbers of samples for the first and second loop MC respectively;
2. from each stochastic and mixed uncertainty (represented by fuzzy random variables), a value (crisp value from each probability distribution) and a triangular fuzzy number (from each fuzzy random variable) are obtained using the first Monte Carlo sampling:  $a_v = (a_{v1}, \dots, a_{vn})$  and  $c = (c_{v1}, \dots, c_{vq})$ , where  $c_{vk}$  are triangular fuzzy numbers;
3. a second Monte Carlo loop is performed to take samples from the fuzzy numbers (those taken in the first loop and those reflecting epistemic uncertainty) as explained above; this means that  $y = f(a_{v1}, \dots, a_{vn}, b_{w1}, \dots, b_{wm}, c_{vw1}, \dots, c_{vwq})$  where  $b_{wj}$  and  $c_{vwk}$  are crisp values from the triangular fuzzy numbers;
4. calculate  $z_{vw} = y = f(a_{v1}, \dots, a_{vn}, b_{w1}, \dots, b_{wm}, c_{vw1}, \dots, c_{vwq})$  to obtain a crisp value with its membership degree. The membership of  $z_{vw}$  is obtained using the extension principle of fuzzy numbers as follows:

$$(\alpha(z_{vw}) = \text{minimum of } (\alpha(a_{v1}), \dots, \alpha(a_{vn}), \dots, \alpha(c_{vw1}), \dots, \alpha(c_{vwq}))$$

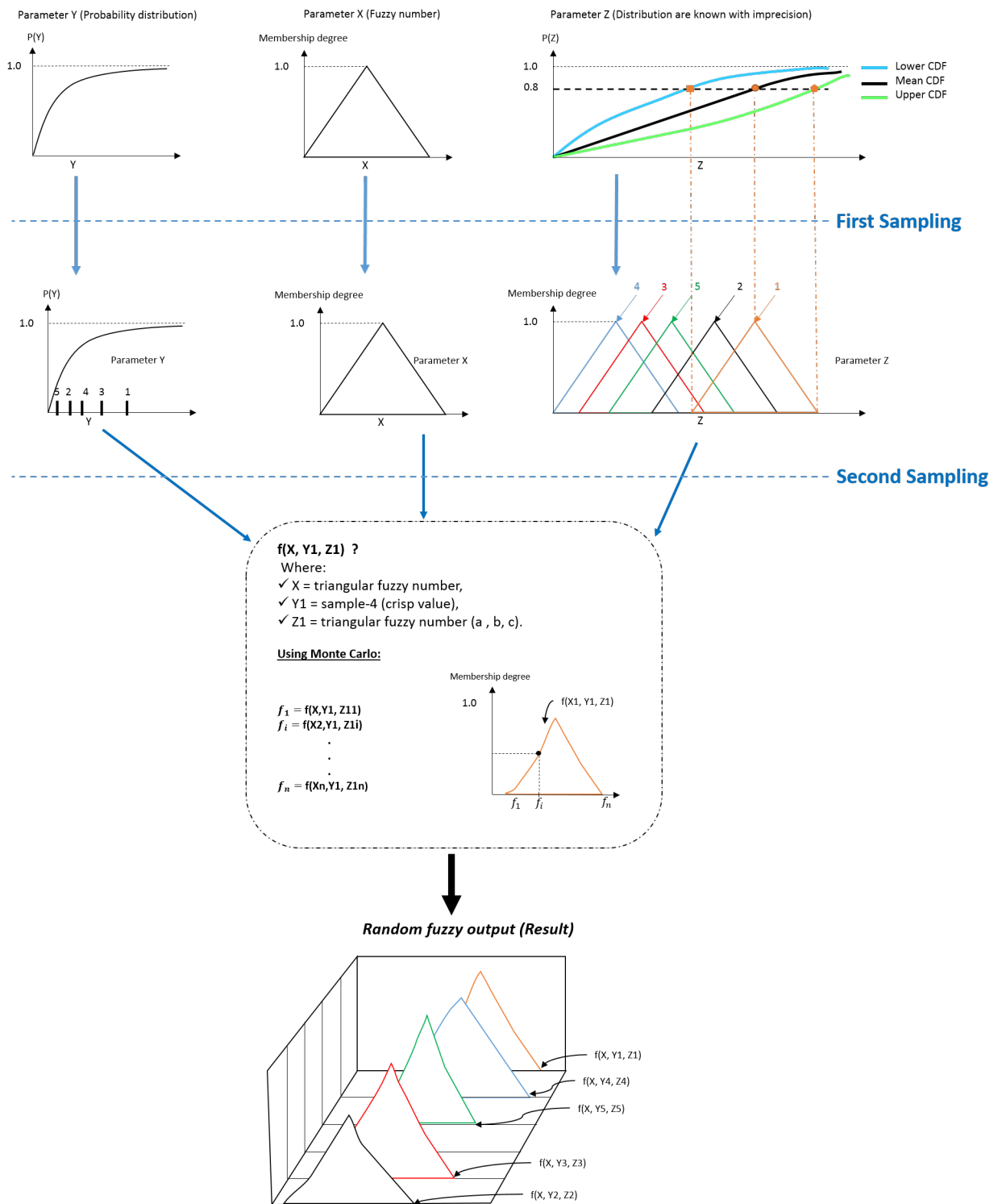
if  $z_{vw}$  has been obtained earlier from another sample then the maximum membership between it and the old membership is given to  $z_{vw}$ );

5. if  $w < N$ , return to step 3, otherwise go to 6;
6. generate the fuzzy number  $z_v$  from the  $N$  calculated samples;
7. if  $v < M$ , go back to step 2, otherwise go to 7;
8. after the simulations,  $M$  triangular fuzzy numbers are obtained. The CDFs of these results are plotted.

This algorithm will be illustrated in the next section and applied to the case of a chemical reactor.

## 5.2 A new uncertainty analysis approach with randomness and fuzzy theory to deal with variability and imprecision

Parameter X: source term parameter that known with imprecision  
 Parameter Y: parameter that known statistically represented by a probability distribution  
 Parameter Z: fuzzy random variable is used to handle the two type of uncertainty that presented in the same parameter



**Figure 5.8** – Use of a 2D Monte Carlo simulation to handle aleatory, epistemic and mixed uncertainties in risk assessment for dynamic systems.

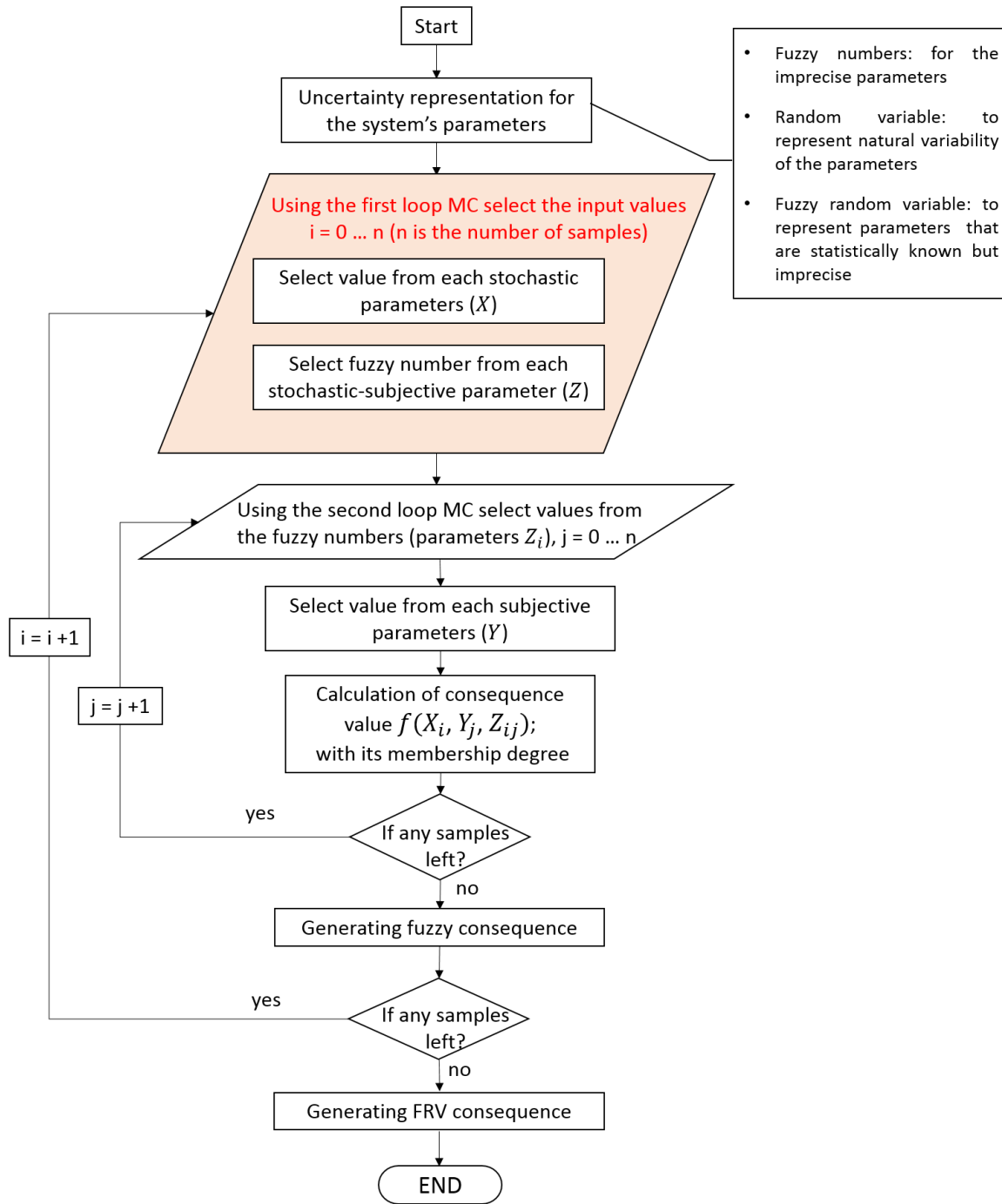


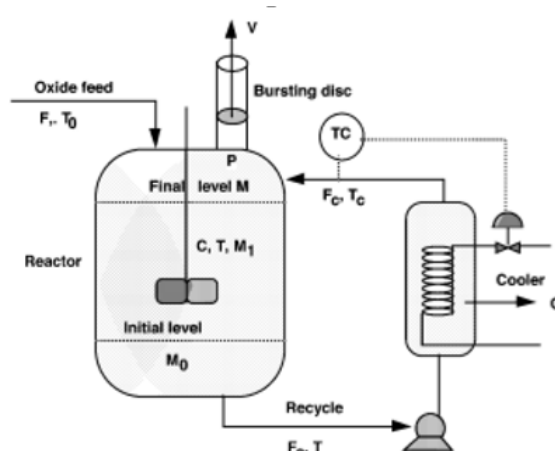
Figure 5.9 – Modeling simulations of the system using fuzzy input.

## 5.2.6 Case study - Application of the proposed hybrid approach to effect analysis with considering of parameter uncertainties

### 5.2.7 Description

The methodology developed is applied to an over-pressure scenario in a Propylene Oxide Polymerization Reactor. The aim is to calculate the intensity of a runaway reac-

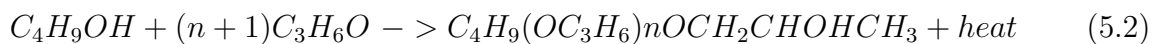
tion caused by a failure in the cooling water system taking uncertainty into consideration. The severity associated with this cooling failure is very high. Java software was developed to implement the differential equations of the reactor, the theories of uncertainty representation and the computation algorithms.



**Figure 5.10** – Polymerization reactor with its cooling unit and bursting disk.

### 5.2.7.1 Reactor model

Propylene oxide polymerization is a highly exothermic process performed at high pressures. An almost isothermal operation is required in order to prevent runaway conditions and the build-up of pressure exceeding the reactor’s design pressure. The safety problems associated with the operation of such a reactor are described by [66]. The reactor, which includes a bursting disk for pressure relief in the event of excessive pressure build-up, was mathematically modeled and simulated by [112]. These researchers considered the manufacture of a polyol lubricant by step-wise condensation of the propylene oxide with butanol:



The catalyzed alcohol is loaded into the reactor up to the “initial” level. The oxide is fed into the reactor at a constant rate until the batch is ready and the reactor is full. Excess heat from the reaction is removed via an external heat removal system. The reaction must be completed at the highest possible rate for economical reasons. The reaction rate is a function of the temperature, catalyst concentration, and the liquid phase oxide concentration (which is a function of the pressure). The limits in relation to the reactor temperature and the catalyst concentration are set according to thermal degradation and purification difficulties. To maximize the reaction rate, the pressure must

be kept as high as possible for the entire duration of the batch. The higher limits of the pressure and reaction rate are dictated by the reactor system pressure suitability and the feasible heat removal rate.

The mathematical model of the reactor, the heat removal system and the bursting disk orifice, as proposed by [112], is shown in Table 5.1. In the system model, Eq. A1-1 corresponds to the variation of the total mass in the reactor ( $M$ ) over time. The concentration of oxides is given by Eq. A1-19, where  $C$  is the oxide mass (Eq. A1-2) per total mass in the reactor. The mass of oxide reacted is calculated by Eq. A1-3, where  $X$  is the mass of oxide reacted at time  $t$ . Eq. A1-4 enables the calculation of the temperature in the system. Under normal operating conditions, the reacting mass is re-circulated through the external heat removal system at a flow rate of  $F_c$  and cooled to temperature  $T_0$  (see equations A1-4 and A1-15 in Table 5.1). The bursting disk is intact ( $\text{Open} = 0$ , see equation A1-5) and the vapor discharge rate through orifice  $V$  is zero (see eq. A1-7). If, for some reason, the pressure exceeds the limit of  $P_{\text{burst}}$ , the bursting disk ruptures. In this case, the variable ‘Open’ becomes greater than zero (eq. A1-5) and vapor discharge is initiated (eq. A1-7) at either a sonic (eq. A1-9) or subsonic (eq. A1-10) discharge rate. The latent heat of vaporization of the discharging oxide cools down the reactor (see eqs. A1-4 and A1-13) and the reaction essentially stops. When the disk ruptures, the feed to the reactor is stopped (eq. A1-6).

### 5.2.7.2 Computation of exposed area

In our study we propose to use the TNT equivalent method to calculate the zone affected by the explosion in the case of failure (runaway reaction). For a given TNT mass, the distance at which a pressure is reached is given by the following relation:

$$d_i = \lambda_i (M_{TNT})^{(1/3)} \quad (5.3)$$

The  $\lambda$  value is determined based on the pressure using the TM5 1300 abacus, Table 5.2. The equivalent TNT mass of a product is determined by the equation:

$$M_{TNT} = n \frac{M \cdot \Delta H}{4690} \quad (5.4)$$

where

- $M$  is the mass released;
- $n$  in Eq. 5.4 is the performance coefficient with an order of 10 %;
- $\Delta H$  is the combustion enthalpy of product (kJ/kg), in our system  $\Delta H = 45790$ .

5.2 A new uncertainty analysis approach with randomness and fuzzy theory to deal with variability and imprecision

**Tableau 5.1** – Model equations and output variable description for the system.

No.	Name	Definition	Initial value	Model equation - Runaway polymerization reaction
A1-1	M	Total mass in the reactor(Kg)	M(0)=4400	$d(M/d(t))=F-V$
A1-2	MC	Oxide mass in the reactor(Kg)	MC(0)=0	$d(MC/d(t))=F-V-r$
A1-3	X	The mass of oxide reacted (kg)	X(0)=0	$d(X)/d(t) = r$
A1-4	TR	Temperature in the reactor(C)	TR(0)=80	$d(TR)/dt=(Hc-Hv-Qg-Qr)/(M*Cp)$
A1-5	Open	Status of the burst disk: 0 closed, >0 open	Open(0)=0	$d(Open)/d(t) = \text{if } (P < P_{burst}) \text{ then } (0) \text{ else } (0.001)$
A1-6	F	Oxide feed rate (kg/min)		$F = \text{if } (Open > 0) \text{ then } (100) \text{ else } (0)$
A1-7	V	Vapor discharge rate (kg/min)		$V = \text{if } ((P \leq 1) \text{ or } (Open = 0)) \text{ then } (0) \text{ else } (V1)$
A1-8	V1	Vapor discharge rate (kg/min)		$V1 = \text{if } (P < 1.9) \text{ then } (V_{subs}) \text{ else } (V_s)$
A1-9	Vs	Sonic vapor discharge rate (kg/min)		$V_s = 0.85 * K_v * P / \sqrt{TR + 273}$
A1-10	Vsubs	Sub-sonic - vapor discharge rate (kg/min)		$V_{subs} = K_v * P / \sqrt{((TR + 273)) * \sqrt{1 + 1/P^2}}$
A1-11	r	Reaction rate (kg oxide/min)		$r = k * MC$
A1-12	Hc	Feed enthalpy change (kJ/min)		$Hc = F * C_p * (T_0 - TR)$
A1-13	Hv	Latent heat of vapor discharge (kJ/min)		$Hv = V * \text{Lamda}$
A1-14	Qg	Heat of reaction (kJ/min)		$Qg = r * HR$
A1-15	Qr	Heat removal (kJ/min)		$Qr = F_c * C_p * (TR - T_0)$
A1-16	P	Oxide vapor pressure (bar)		$P = \text{if } (P1 < 1) \text{ then } (1) \text{ else } (P1)$
A1-17	P1	Oxide vapor pressure (bar)		$P1 = (\exp(-3430 / (TR + 273) + 11.7) + 1.45e-3 * MW) * C$
A1-18	k	Reaction rate coefficient		$k = 9e9 * \exp(-E / (R * (TR + 273)))$
A1-19	C	Oxide concentration (kg/kg)		$C = MC / M$
A1-20	MW	Molecular weight of the polymer (kg/mol)		$MW = (M_0 + X) / (M_0 / 74)$
A1-21	T0	Feed temperature (°C)		$T_0 = 80$
A1-22	Lamda	Lamda Heat of vaporization of the oxide (kj/kg)		$Lamda = 670$
A1-23	Cp	Spec. heat of feed reacting mass (kJ/kg-°C)		$C_p = 3.5$
A1-24	HR	Heat of reaction (kJ/ kg oxide)		$HR = -1660$
A1-25	Fc	Re-circulation mass flow rate (kg/min)		$F_c = 3300$
A1-26	Pburst	Disk rupture pressure (bar)		$P_{burst} = 8$
A1-27	R	Gas constant		$R = 1.987$
A1-28	E	Activation energy		$E = 21000$
A1-29	M0	Initial alcohol charge (kg)		$M_0 = 4400$
A1-30	Kv	Valve discharge coefficient		$K_v = 100$

### 5.2.7.3 Uncertainty modeling

This section will define the uncertain parameters presented in the system under study. A suitable representation is then given for each one based on the type of uncertainty



Over-pressure(mbar)	Reduced Distance $\lambda$ (TM5-1300 abacus)(m)
50	22
140	10.1
170	8.9
200	7.6

**Tableau 5.2** – Lamda using the TM5 1300 abacus.

affecting it. These uncertain parameters are presented as follows:

- the time of failure of the cooling system is modeled using random variables based on an exponential probability distribution - this is ***an aleatory uncertainty***;
- the initial mass and temperature of the reactor are known but with a certain degree of imprecision (based on an expert elicitation) - **epistemic uncertainty** - represented by triangular fuzzy numbers (see Table 5.3);
- the time response of the operator in an emergency situation is more complex when it comes to modeling and we shall therefore use fuzzy random variables since it involves both - **aleatory and epistemic uncertainty**.

Fuzzy numbers	Lower value	Mean value	Upper value
Initial mass in the reactor [Kg]	4400	5000	6000
Temperature in the reactor [ $^{\circ}C$ ]	80	90	100

**Tableau 5.3** – Fuzzy numbers for uncertain parameters.

#### 5.2.7.4 Simulation conditions

The model of the system under study includes six differential equations: mass balance in the reactor (yields the total mass); component balance (yields the mass of the oxide component); enthalpy balance (yields the temperature in the reactor); reaction rate (yields the mass of oxide reacted); the bursting disk (open or closed); and the vapor discharge rate. These equations are solved using an ODE solver. The simulations of this system are developed for a batch duration of 800 min. The first simulation is under normal conditions (i.e. no breakdown) to see how the system behaves, how the temperature varies over time, etc. The second is under abnormal conditions. The aim is again to see how the system behaves but also to calculate the severity of the consequences (areas affected), but without taking into account input data uncertainty. The third simulation takes into account input data uncertainty to see what the effects of this uncertainty are.

#### 5.2.7.4.1 *Simulation under normal conditions*

We first simulate the system under normal conditions without considering uncertainty. All the system's parameters are crisp values and are presented in Table 5.1.

The changes to the system are computed while the total mass in the reactor ( $M$  in kg), the Oxide mass in the reactor ( $M_c$  in kg), the temperature, the mass of the oxide reacted (Kg), the status of the bursting disk (Open), and the Vapor discharge rate ( $\text{kg}/\text{min}^{-1}$ ) are plotted over time (see Figure 5.11). This simulation is used as a reference.

#### 5.2.7.4.2 *Abnormal condition without uncertainty*

Let us consider a situation where there is a cooling water failure lasting 12 minutes and occurring 700 minutes from the start of the batch. In order to simulate the reactor's operation under such abnormal conditions, some necessary changes are introduced into the model's equations.

$$if ( (t > 700) \& (t < 712) ) \{Fc = 0\}$$

The results for the output parameters ( $M$ ,  $MC$ ,  $TR$ ,  $X$ ,  $Open$ ,  $V$ ) for normal and abnormal conditions are presented in Figure 5.11 and Figure 5.12 respectively. Using this new model, we can make a comparison between the output parameters under normal conditions and in the case where there is a cooling failure. It is clear that in the abnormal case the temperature reaches a maximal value of  $310\text{ }^\circ\text{C}$  (see Figure C5.12(c)), whereas under normal conditions the maximal value is  $112\text{ }^\circ\text{C}$  (see Figure C5.11(c)). This abnormal increase in temperature ruptures the bursting disk resulting in a release of toxins (Figures C5.12(e) and C5.12(f)). The mass released (mass of vapor discharged after the failure, see Figure C5.12(f)) following the cooling failure is used to calculate the area affected when the explosion takes place (see Section 5.2.7.2). After estimating the mass released, the affected area for this mass is calculated. The distance is represented by a crisp value since uncertainty is not taken into account in this case.

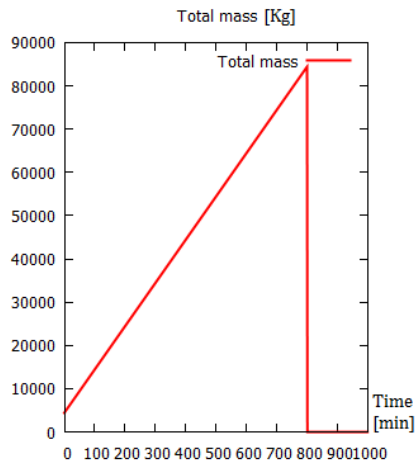
The mass released and the radius of the affected area in the case of failure are:

$$\text{Mass released} = 318.92$$

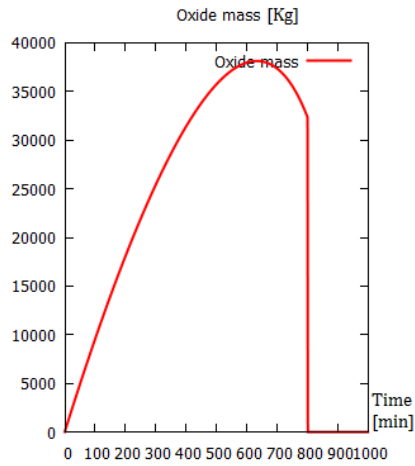
Affected area represented by a crisp value

$$\text{Affected area} = 217.23$$

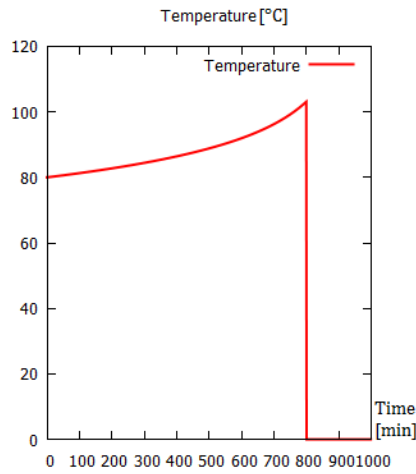
we will use these values as reference in the next section.



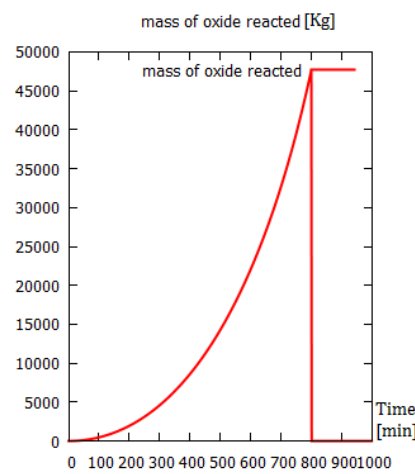
(a) Total mass of the reactor.



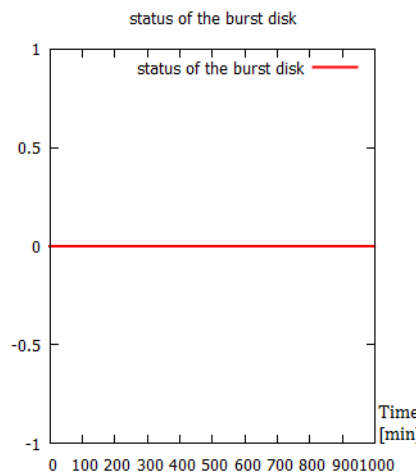
(b) Oxide mass in the reactor.



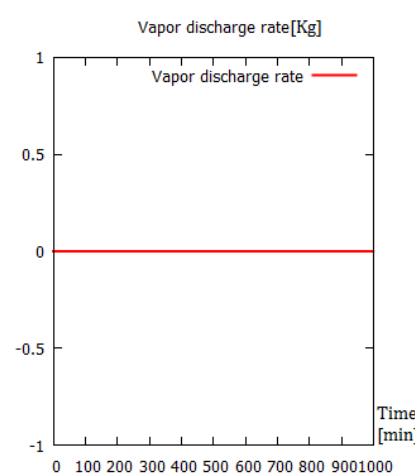
(c) Temperature in the reactor.



(d) The mass of oxide reactor.



(e) Status of the burst disk.



(f) Vapor discharge rate.

**Figure 5.11** – Simulation under normal conditions with normal numbers.

5.2 A new uncertainty analysis approach with randomness and fuzzy theory to deal with variability and imprecision

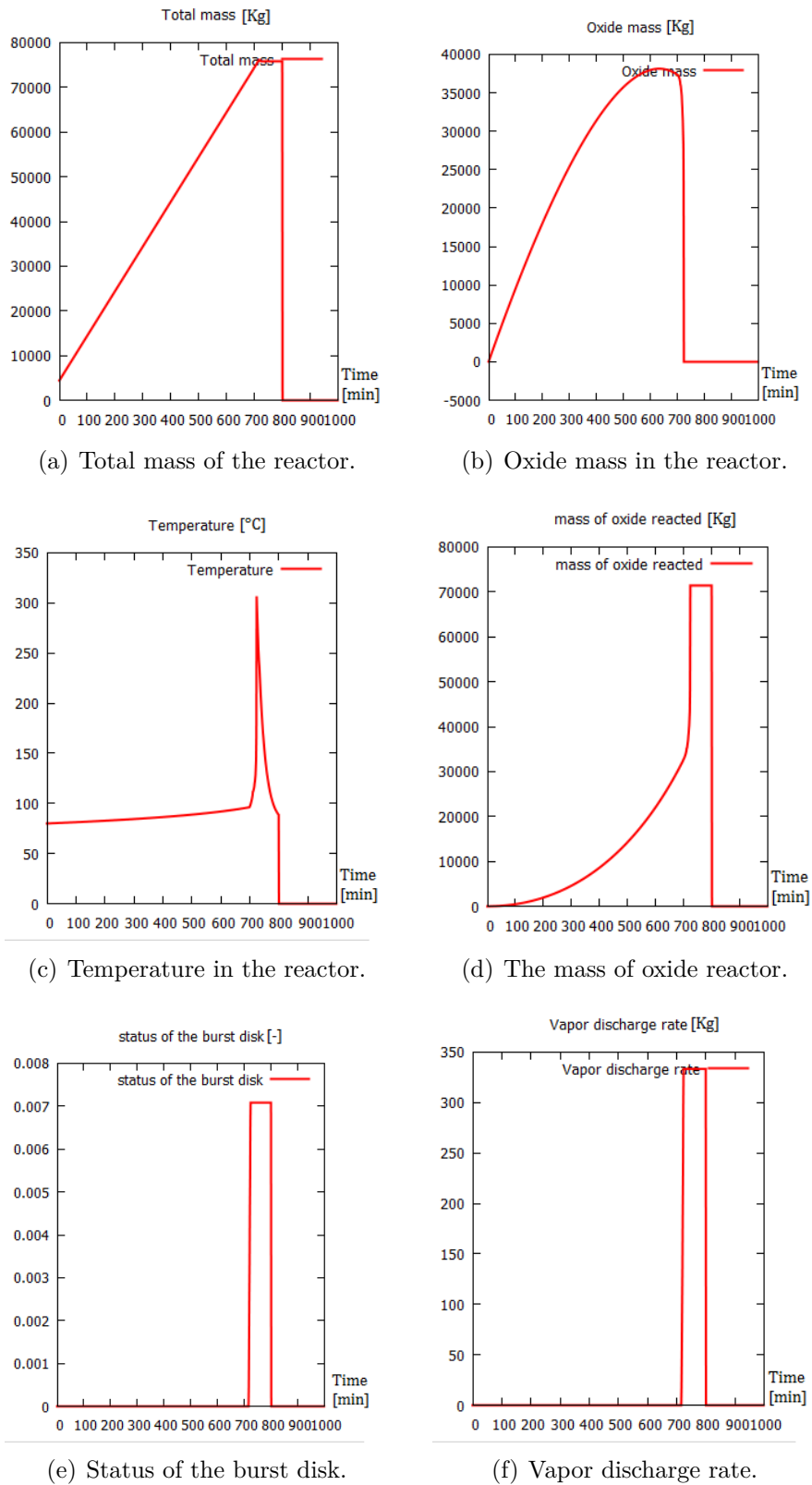
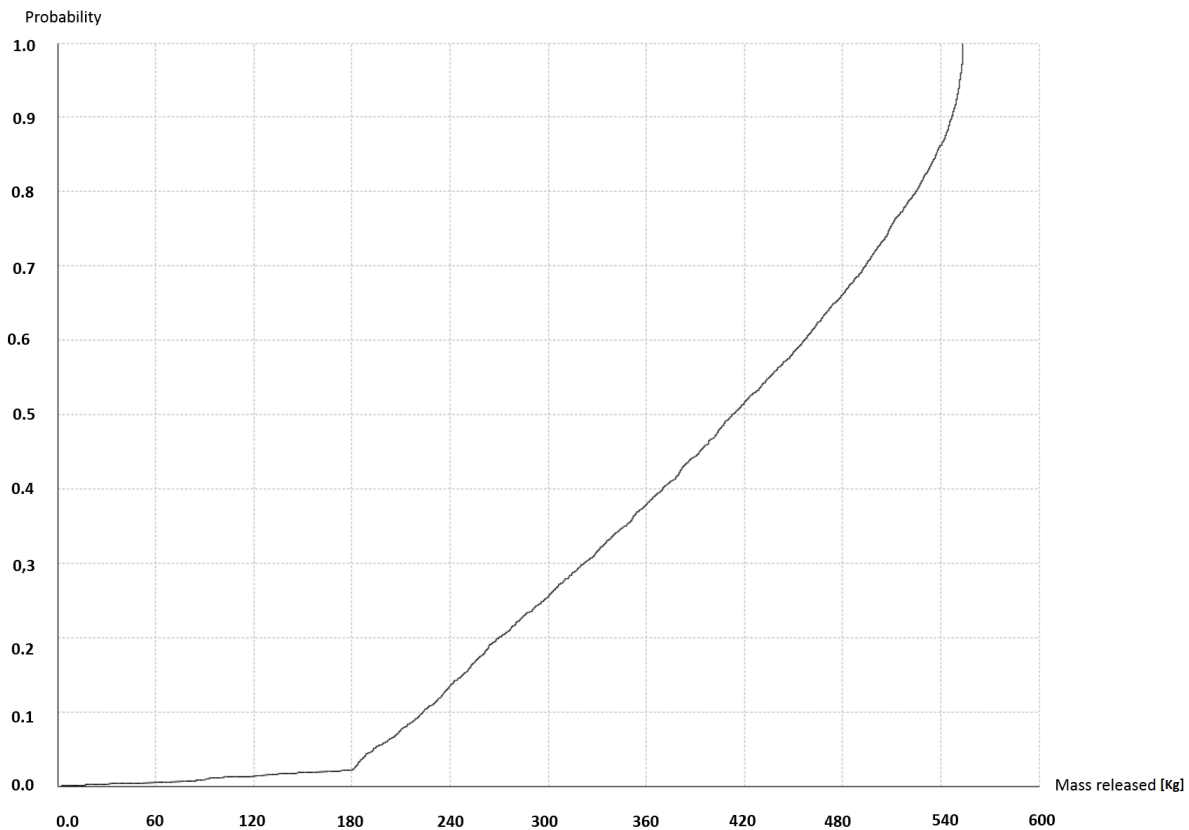


Figure 5.12 – Simulation under abnormal conditions with normal numbers.

### 5.2.7.4.3 Abnormal conditions with uncertainty

To provide a better understanding of how to handle each type of uncertainty, three kinds of simulation are performed. The first simulation aims to represent and propagate aleatory uncertainty presented in the time-to-failure of the cooling system using the Monte Carlo method, as explained in Section 5.2.5.1. The second aims to propagate epistemic uncertainty relating to certain input parameters and stemming from the inability to determine exact values (see Section 5.2.5.2). The last aims to represent all uncertainty types affecting the system under study (aleatory, epistemic and mixed), see Section 5.2.5.3.



**Figure 5.13** – CDF of mass released after consideration of aleatory uncertainty.

**Uncertainty represented by random variables** In this simulation, only the aleatory uncertainty is taken into account. The time of failure of the cooling system is the only uncertain parameter and is represented by a random variable. Instead of being represented by a crisp value, the time of failure is randomly chosen as in Section 5.2.7.4.2. 12 minutes is taken as the duration of each failure. The simulation is performed as detailed in Section 5.2.5.1 with  $10^4$  samples and depicted in Figure 5.13.

The result shows that a variation in failure time leads to a variation in the mass released. Failing to take this uncertainty into account may result in an underestimation

of the outputs or other issues in the analysis.

**Uncertainty represented by fuzzy numbers** The initial inputs are represented by crisp and triangular fuzzy numbers to reflect the imprecise parameters. As mentioned in Section 5.2.7.3, the imprecise parameters are as follows:

- $M=(4400, 5000, 6000)$ ; // Triangular fuzzy number
- $TR=(80, 90, 100)$ ; // Triangular fuzzy number

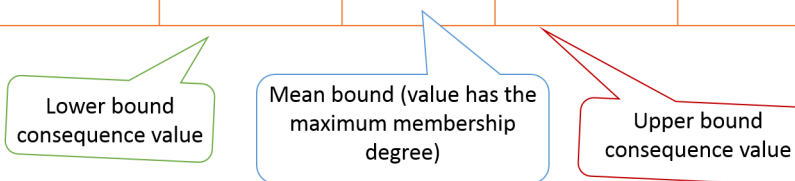
The principle of the simulation is presented in Figure 5.7.

To understand the extent of the imprecision, the time of cooling system failure is taken to be the same as in Section 5.2.7.4.2. In order to know the sample numbers required to perform an efficient Monte Carlo simulation, the simulations are performed with 5,000 and 10,000 simulations according to Uniform and Gaussian distribution laws. Then, the fuzzy mass released is generated and the fuzzy distance is calculated.

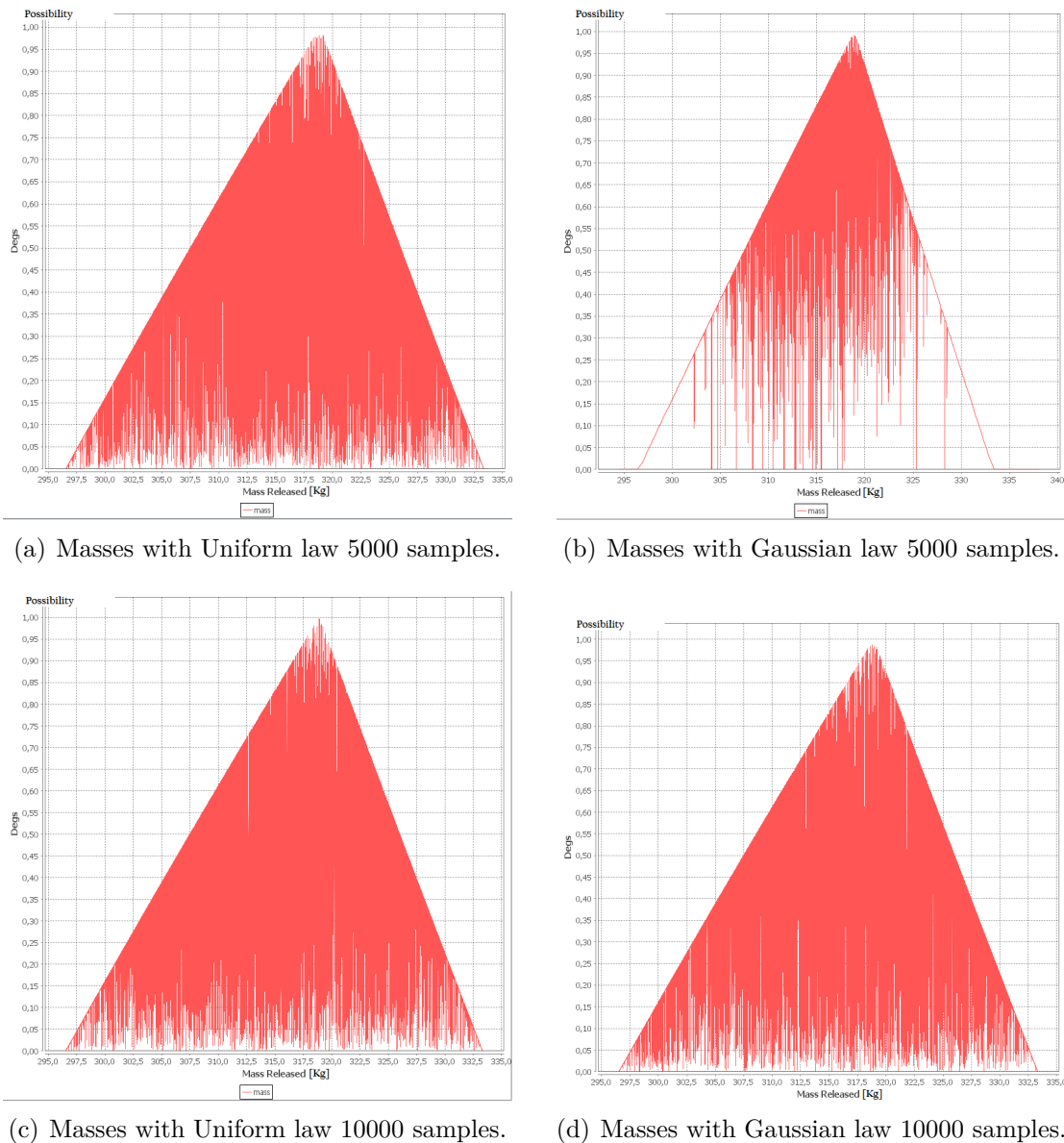
The results of these simulations are presented in Figure 5.14. We can see that 5,000 samples are sufficient to generate a correct estimation as are 10,000. In addition, the laws used to generate the samples do not affect the output since both laws show the same result.

After estimating the fuzzy mass released, we generate the minimum and maximum mass released to restrict the distances affected by this fuzzy mass. The values for the fuzzy mass released and the fuzzy zone affected are depicted in Table 5.4. The outputs of the simulation show that the mass released and the distance affected, calculated in Section 5.2.7.4.2, are between the two borders of the fuzzy outputs (see Table 5.4).

	5000 samples			10000 samples		
	Lower	Mean	Upper	Lower	Mean	Upper
Fuzzy mass released (Uniform law) [Kg]	296.54	318.39	333.32	296.54	318.91	333.33
Fuzzy distance [m]	185.22	213.19	242.64	185.22	212.14	242.64
Fuzzy mass released (Gaussian law) [Kg]	294.48	308.84	338.06	296.60	318.80	333.33
Fuzzy distance [m]	184.79	214.92	243.78	184.79	218.50	243.78



**Tableau 5.4** – Fuzzy distance and fuzzy mass released for the simulations.



**Figure 5.14** – Masses released after the simulations.

**Uncertainty represented by fuzzy random variable** In this case, as well as representing the variability of the cooling system time of failure and parameter imprecision, an operator acting on a sprinkler system has been added to the model. This makes it possible to adjust the failure of the cooling system if presented. Statistically, the response of this operator is imprecisely known, meaning that both uncertainty types apply. Thus, the system model is affected by all three types of uncertainty: aleatory, epistemic and mixed. The response of the sprinkler is represented using a fuzzy random variable because it is manually turned on in an abnormal situation. Several operators (human beings) are responsible for doing this although each uses his/her own personal experience when re-

acting to such a situation. Their behavior can thus differ. These different “experts” may therefore respond differently (variability between the different experts can be observed). The imprecision relating to each operator’s response can be presented according to their experience and what they are doing when the abnormal situation happens.

To represent the operator in the model, a new term  $Q_{deluge}$  will be added to the relation  $dTR/dt$  in Table 5.1 eq. 3.

$$dTR = (Hc - Hv - Qg - Qr - Q_{deluge}) / (M * Cp) \quad (5.5)$$

where

$$Q_{deluge} = F_{deluge} * A * Cp * (TR - T0) \quad (5.6)$$

$F_{deluge}$  must be set by test error to obtain a cooling time response able to adjust the failure and prevent the runaway reaction (details are presented in the next section). This cooling helps to minimize the pressure and thus limits the release.

The simulation is performed as explained in Section 5.2.5.3. Figure 5.15 shows the fuzzy cumulative distribution function of mass released, the first loop MC is repeated 5,000 times (5,000 fuzzy numbers are generated). With each repetition, an intern loop of 5,000 samples is also performed. However, decision making based on this result will be effective, precise and conservative. Degree of conservativeness from 0% to 100% can be chosen from the output FRV. This degree of conservativeness will be discussed in more details in Section 5.3 when we compare this mixed approach with the existing approaches.

**Efficiency analysis of the sprinkler system** This section aims to evaluate the effectiveness of the sprinkler. The evaluation is based on monitoring of the same simulation performed in the previous section with  $1,000 \times 1,000$  as the dimension of the first and second loop.

After running the first simulation, the random failure value obtained is equal to 487.6 and (4.5, 6.2, 10.3) is the triangular fuzzy number of the response time presented in Figure 5.16. After finishing the simulations, we observe 308 unresolved failures from the 1,000 samples (1,000 samples were selected from the triangular response time of the sprinkler out of which 308 were unable to prevent the release).

We then take the same sprinkler’s response time (to see the relation between the response time and the time of failure of the cooling system). The failure time observed is 752.5. The output shows that the number of unresolved failures is 1,000, which is greater than before (308 at a failure time of 487.6). The third simulation with the same fuzzy response and a failure time of 334.0 shows zero unresolved failures. The results are



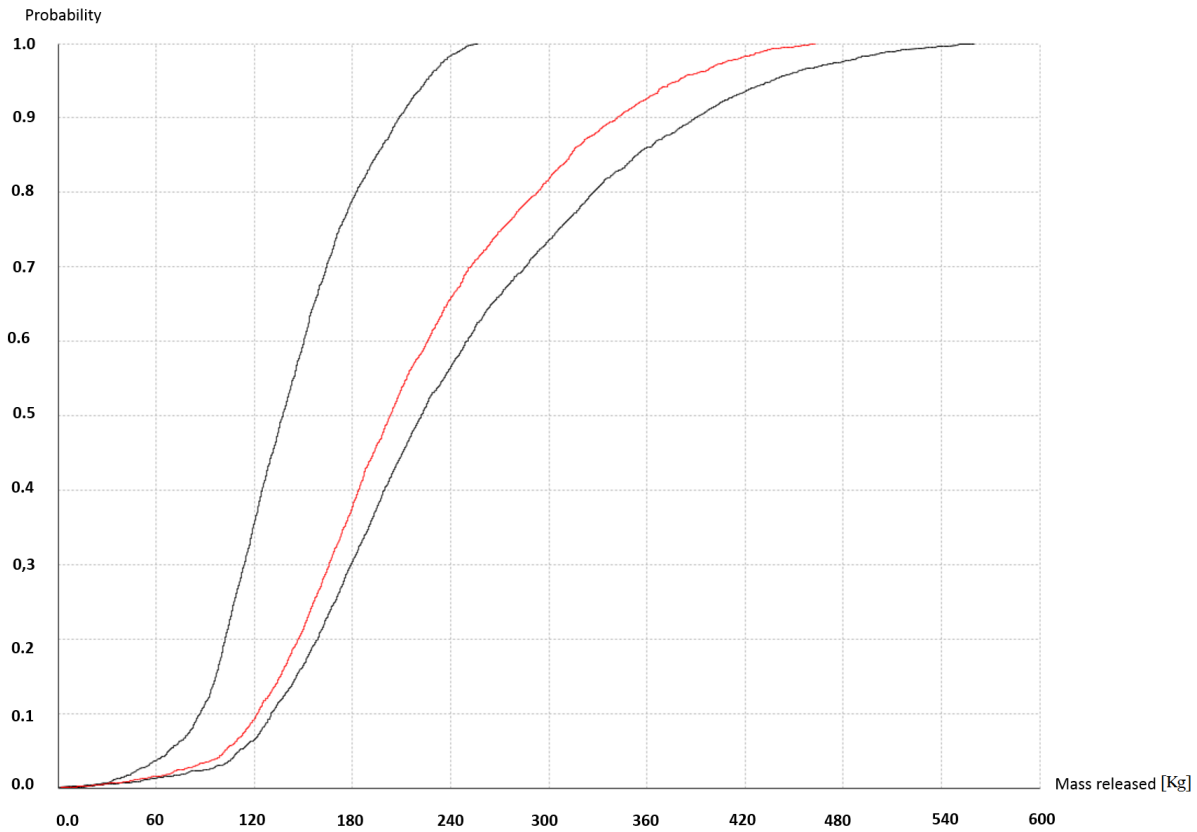


Figure 5.15 – CDF of results after taking into consideration all uncertainty types.

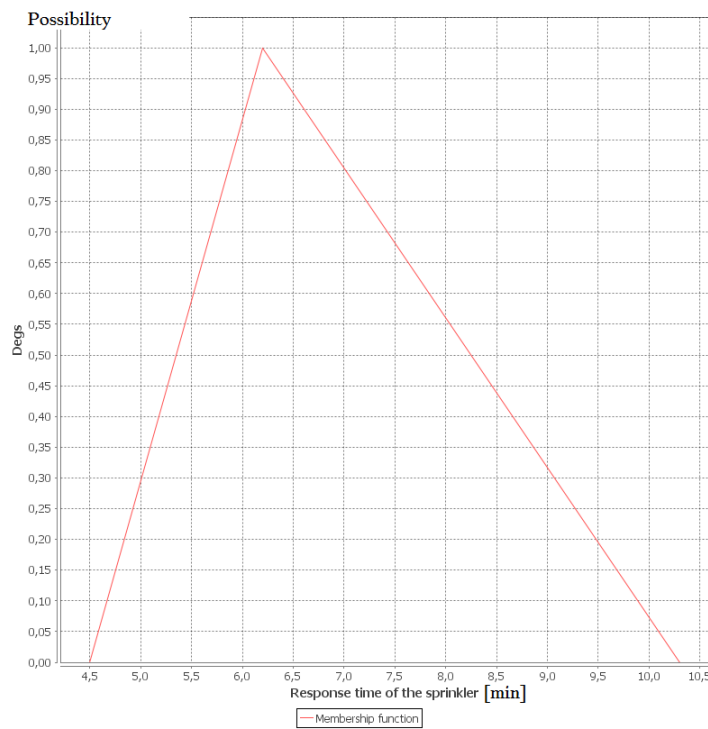
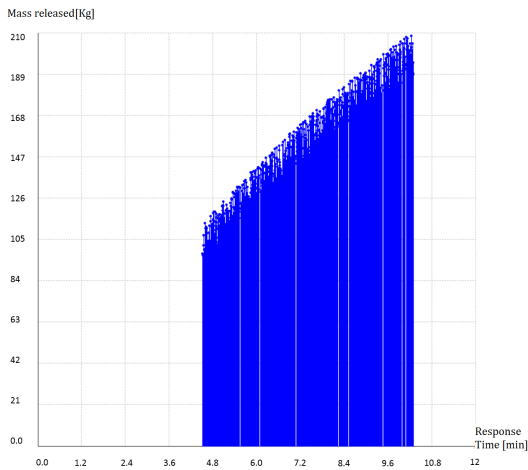
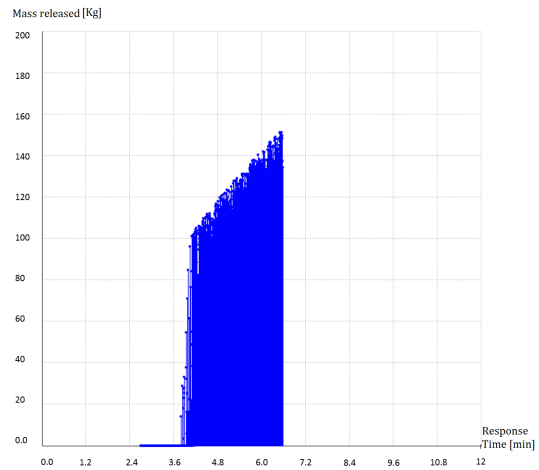


Figure 5.16 – Triangular fuzzy number of the response time.

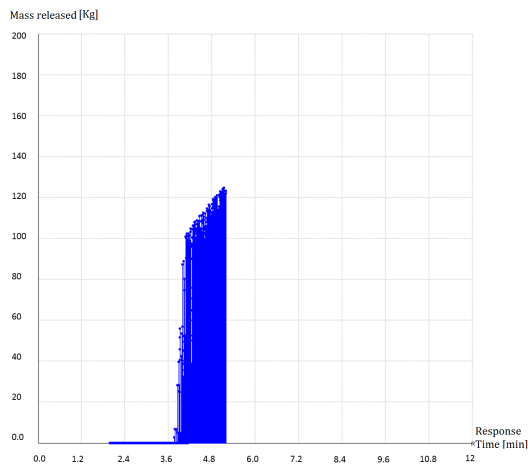
## 5.2 A new uncertainty analysis approach with randomness and fuzzy theory to deal with variability and imprecision



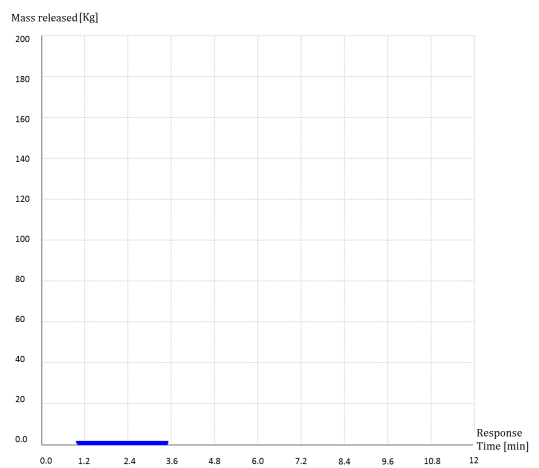
(a)  $TTF(\text{time to failure}) = 752.5$  and Response time = (4.5, 6.2, 10.3).



(b)  $TTF(\text{time to failure}) = 752.5$  and Response time = (2.7, 4.0, 6.6).



(c)  $TTF(\text{time to failure}) = 752.5$  and Response time = (2.0, 3.1, 5.2).



(d)  $TTF(\text{time to failure}) = 752.5$  and Response time = (1.0, 1.8, 3.5).

**Figure 5.17** – Estimation of the fuzzy mass released for  $TTF = 752.5$  and different response times.

presented in Table 5.5.

Failure time [min]	Failure not resolved
752.5	1000
487.6	308
334.0	0

**Tableau 5.5** – Results for fixed fuzzy time response (4.5, 6.2, 10.3)

The next step aims to fix the failure time and simulate the system at different fuzzy response times. The results for the simulations of the system at a failure time of 752.5 are presented in Table 5.6) and Figure 5.17.

Triangular fuzzy response [ <i>min</i> ]	Failure not resolved
(4.5, 6.2, 10.3)	1000
(2.7, 4.0, 6.6)	684
(2.0, 3.1, 5.2)	393
(1.0, 1.8, 3.5)	0

**Tableau 5.6** – Result at failure time(752.5)

The solution for the cooling water failure at time 752.5 is resolved by the sprinkler system with a quick response time:

- (1.0, 1.8, 3.5) fuzzy time response.

Fuzzy time response [ <i>min</i> ]	$F_{deluge}$ [ <i>Kg/min</i> ]	Failure not resolved
(4.5, 6.2, 10.3)	3000	1000
	4000	463
	5000	0
(2.0, 3.1, 5.2)	3000	1000
	4000	0
	5000	0

**Tableau 5.7** – Result at time of failure equal to 752.5.

Furthermore, a third simulation, which consists in fixing the time of failure and the fuzzy response, is then run to change the flow rate value ( $F_{deluge}$ ). The results obtained are represented in Table 5.7. The output of the simulation presented in Table 5.7 shows that at a failure time equal to 752.5 we have 1,000 unresolved failures for a flow value equal to 3,000 ( $F_{deluge}$ ) for both fuzzy time response (4.5, 6.2, 10.3) and (2.0, 3.1, 5.2), while for  $F_{deluge}$  equal to 4,000, the number of unresolved failures is 463 for (4.5, 6.2, 10.3) and 0 for (2.0, 3.1, 5.2). These results prove that the  $F_{deluge}$  value can be lower when the sprinkler time response is fast enough.

The results for this section show that introducing the sprinkler system prevents the temperature runaway. With the right strategy parameters for the proposed solution to the cooling water failure (sprinkler system), the level of vulnerable variables (temperature and pressure) stays well below the disk rupture threshold value. Thus, the batch is successfully completed.

Therefore, the system’s response depends on several conditions to prevent the runaway reaction. A healthy response depends on the failure, response times and the flow value ( $F_{deluge}$ ). It should be emphasized that failure after a long reactor run time requires a rapid response time and a high flow value. Finally, in spite of all the uncertainties presented in the system under study, we manage to maintain its stability thanks to an

efficient sprinkler response and the only remaining risk is linked to the failure of the sprinkler system.

### 5.3 Comparison of the proposed hybrid approach with the existing approaches: applying uncertainty analysis to a Loss Of Containment scenario

The purpose of this section is to compare the proposed approach of the previous section with the approaches reviewed in Chapter 3, Section 3.1 to quantify uncertainty, checking how they differ and sizing up the advantages and disadvantages of each one and, finally, establishing why one is more suitable than another in certain cases. The interval, fuzzy (possibilistic), probabilistic, evidence and probabilistic-fuzzy approaches will be compared by applying them to a real LOC scenario. According to the bibliographic research performed, this will be the first time that all the above-mentioned approaches are systematically compared with regard to uncertainty in risk analysis.

The case involves the total loss of containment of a pressurized vessel holding Hydrochloric acid (HCL). The assessment will consider the release of the HCL following the appearance of a crack in the storage vessel and aims to calculate the concentration at end-points  $(x, y, z)$ , as explained in figure 5.18. Two mathematical models, one for discharge and one for dispersion, will be used to calculate the concentrations as presented in the next section.

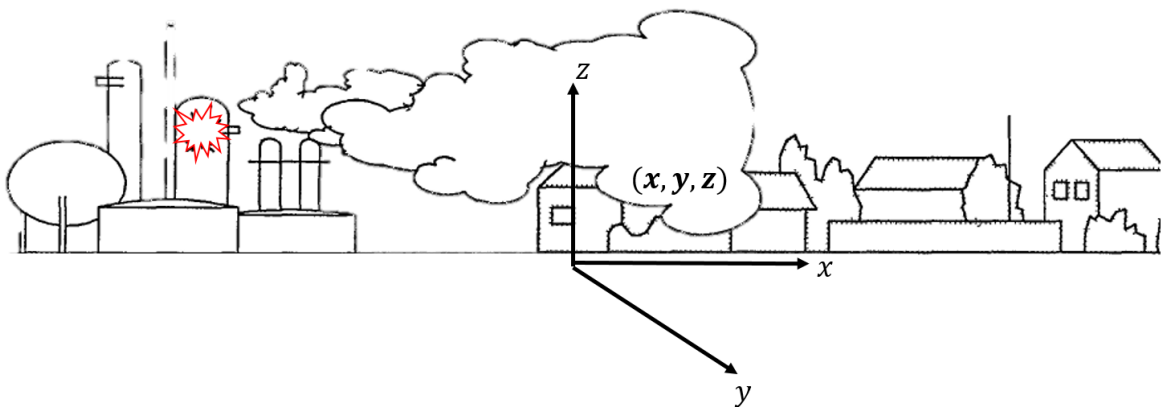


Figure 5.18 – Loss of containment scenario

### 5.3.1 Models used in the analysis

This section highlights the characteristics of the two mathematical models used in the assessment: (1) a discharge model for estimating the source terms and (2) a dispersion model for estimating the concentration of HCL at a specific end-point as presented in Eqs 5.7 and 5.8 respectively.

#### Discharge model

The equation used for the mass discharge rate is taken from [34] and refers to discharge from a pressurized tank.

$$Q = AC_d \sqrt{\frac{2(P_0 - P_a)}{V_f}} \quad (5.7)$$

where

- $Q$  is the mass discharge rate [ $Kg/s$ ];
- $A$  is the area of the hole [ $m^2$ ];
- $C_d$  is the discharge coefficient;
- $V_f$  is the specific volume of liquid [ $m^3/kg$ ];
- $P_0$  is the storage pressure [ $N/m^2$ ];
- $P_a$  is the ambient pressure [ $N/m^2$ ].

#### Dispersion model

$$C(x, y, z) = \frac{Q}{2\pi U \sigma_y \sigma_z} \left\{ \exp\left(\frac{-(z-h)^2}{2\sigma_z^2}\right) + \exp\left(\frac{-(z+h)^2}{2\sigma_z^2}\right) \right\} \left\{ \exp\left(\frac{-(y)^2}{2\sigma_y^2}\right) \right\}. \quad (5.8)$$

where

- $x, y, z$  are the downwind, crosswind and vertical distances respectively;
- $C(x, y, z)$  is the average concentration [ $kg/m^3$ ];
- $Q$  is the continuous release rate;
- $\sigma_y, \sigma_z$  are the dispersion coefficients in the  $y$  and  $z$  directions [ $m$ ];
- $U$  is the wind speed [ $m/s$ ].

### Dispersion coefficients and stability classes

The horizontal and vertical dispersion coefficients,  $\sigma_y$  and  $\sigma_z$ , are obtained after determining the atmospheric stability class ([153]; [45]). Six stability classes (Pasquill-Gifford-Turner classes) from *A* to *F* were defined by [122]. The list of these classes and their relationship to wind speed and cloud cover are given in table 5.8.

Wind speed (m/s)	Day time insolation			Night (Cloud Cover)	
	Strong	Moderate	Slight	> 4/8	< 3/8
<2	A	A-B	B	-	-
2-3	A-B	B	C	E	F
3-5	B	B-C	C	D	E
5-6	C	C-D	D	D	D
>6	C	D	D	D	D

**Tableau 5.8** – Pasquill stability classes

Many schemes have been proposed in different studies in order to estimate the horizontal and vertical dispersion coefficients through the stability classes ([22]; [77]; [152]; [146]; [162]). In this study, the estimation of these coefficients is based on the scheme proposed by [27]. It is characterized by a number of advantages when surface roughness is taken into account. The mathematical equation to calculate each  $\sigma$  is expressed as:

$$\sigma = ax(1 + bx)^{-1/2}$$

where *a* and *b* are the fitting constants given in table 5.9.

#### 5.3.2 Calculation of the concentration without considering uncertainty

In this section, the concentration of HCL at grid point  $(x, y, z) = (500, 0, 0)$  is calculated without considering input parameter uncertainty. The output of this section will be used as a reference to see the effect of uncertainty on the output. The stability class used in the calculation is assumed to be *B*, and two wind speed values are taken into consideration, 3 and 5[m/s]. The input parameters for the discharge and dispersion models are presented in table 5.10. The resulting concentrations for the two different wind speed values are:

Fitting Constant	ATMOSPHERIC STABILITY CLASSES					
	A	B	C	D	E	F
Open-Country Conditions						
$a_y$	0.22	0.16	0.11	0.08	0.06	0.04
$a_z$	0.20	0.12	0.08	0.06	0.03	0.016
$b_y$	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
$b_z$	0	0	0.0002	0.0015	0.0003	0.0003
Urban Conditions						
$a_y$	0.32	0.32	0.22	0.16	0.11	0.11
$a_z$	0.24	0.24	0.20	0.14	0.08	0.08
$b_y$	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
$b_z$	0.001	0.001	0	0.0003	0.00015	0.00015

**Tableau 5.9** – Briggs’ Fitting Constants for  $\sigma_y$  and  $\sigma_z$

Concentration for  $U = 3$  [m/s]:

$$C(500, 0, 0) = 518 \text{ [mg/m}^3\text{];}$$

Concentration for  $U = 5$  [m/s]:

$$C(500, 0, 0) = 310 \text{ [mg/m}^3\text{];}$$

Input variables	Values	
A, hole area	$3.5E - 3$ [m <sup>2</sup> ]	
$C_d$ , discharge coefficient	0.7	
$V_f$ , specific volume	$8.4 \times 10^{-4}$ [kg/m <sup>3</sup> ]	
$P_0$ , storage pressure	$2.1 \times 10^5$ [N/m <sup>2</sup> ]	
$P_1$ , ambient pressure	$1 \times 10^5$ [N/m <sup>2</sup> ]	
$U$ , wind speed	3	5 [m/s]

**Tableau 5.10** – Input values used to calculate the concentration at a grid point without considering uncertainty

5.3 Comparison of the proposed hybrid approach with the existing approaches: applying uncertainty analysis to a Loss Of Containment scenario

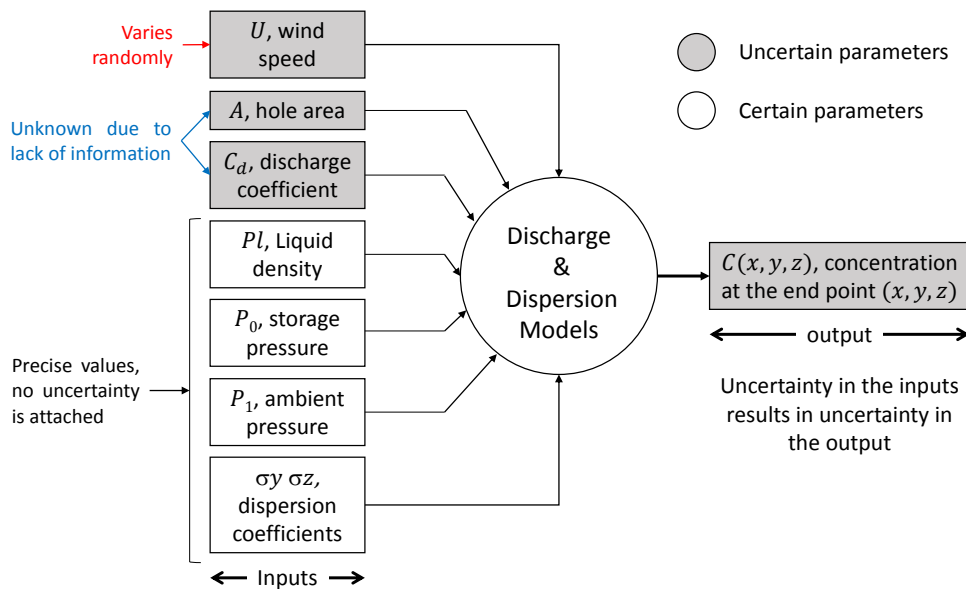
Uncertain parameters	Type of uncertainty	Source of data
A, hole area	Epistemic	expert elicitation
$C_d$ , discharge coefficient	Epistemic	expert elicitation
U, wind speed	Aleatory	statistical data

**Tableau 5.11** – Types of uncertainty and sources of information about the uncertain parameters

### 5.3.3 Uncertainty modeling

This section aims to identify all the uncertain parameters presented in the models used, the type of uncertainty that affects each parameter, and the available data pertaining to these parameters together with their source (see table 5.11 and figure 5.19). Information relating to these uncertain parameters is either derived from statistical data or based on expert judgment. It is presented below:

- A: somewhere between  $[20E - 3, 50E - 3]$  but more likely to be close to the central tendency;
- $C_d$ : between the interval  $[0.7, 0.9]$  but more likely to be close to the central tendency;
- U: based on statistical data, u follows a triangular distribution of inputs (3, 4, 5).



**Figure 5.19** – Uncertain and precise input parameters used to calculate the concentration at a specific end-point



In the following sections, the different approaches to modeling uncertainty will be used to represent these uncertain variables as well as to propagate them through the models in order to achieve a representation of uncertainty for the output concentration.

### 5.3.4 Uncertainty analysis using interval analysis

This analysis aims to represent all the uncertain variables stated in Section 5.3.3 in terms of intervals. Each interval should represent a conservative bound of the uncertain parameter. The chosen bounds of the intervals are the minimum and maximum values provided by experts or based on statistical data. They are presented in table 5.12. Note that the stability class is taken as equal to B as in the previous section and will remain so in the following sections. The propagation of these intervals results in a concentration interval equal to:

$$C(500, 0, 0) = [177, 951] \text{ mg/m}^3$$

Input variables	Values
A, hole area	$[2E - 3, 5E - 3] \text{ [m}^2\text{]}$
$C_d$ , discharge coefficient	$[0.7, 0.9]$
$U$ , wind speed	$[3, 5] \text{ [m/s]}$

**Tableau 5.12** – Uncertain variables represented as intervals

It is clear that the interval of the output concentration includes the values obtained in the case where uncertainty is not considered. This is due to the way in which the imprecision and subjectivity in hole area determination are treated, as well as to the discharge coefficient and wind speed variability.

### 5.3.5 Uncertainty analysis using the fuzzy approach

With this method for treating uncertainty, all uncertain variables are specified as triangular fuzzy numbers (i.e. the best representation of the information provided in paragraph 5.3.3 in order to produce the closest likelihood when approaching the central interval tendency). The lower and upper bounds for the input parameters at  $\alpha = 0$  are the same as the minimum and maximum values taken in the analysis using the interval approach. Table 5.13 presents the fuzzy numbers for all the uncertain variables. The

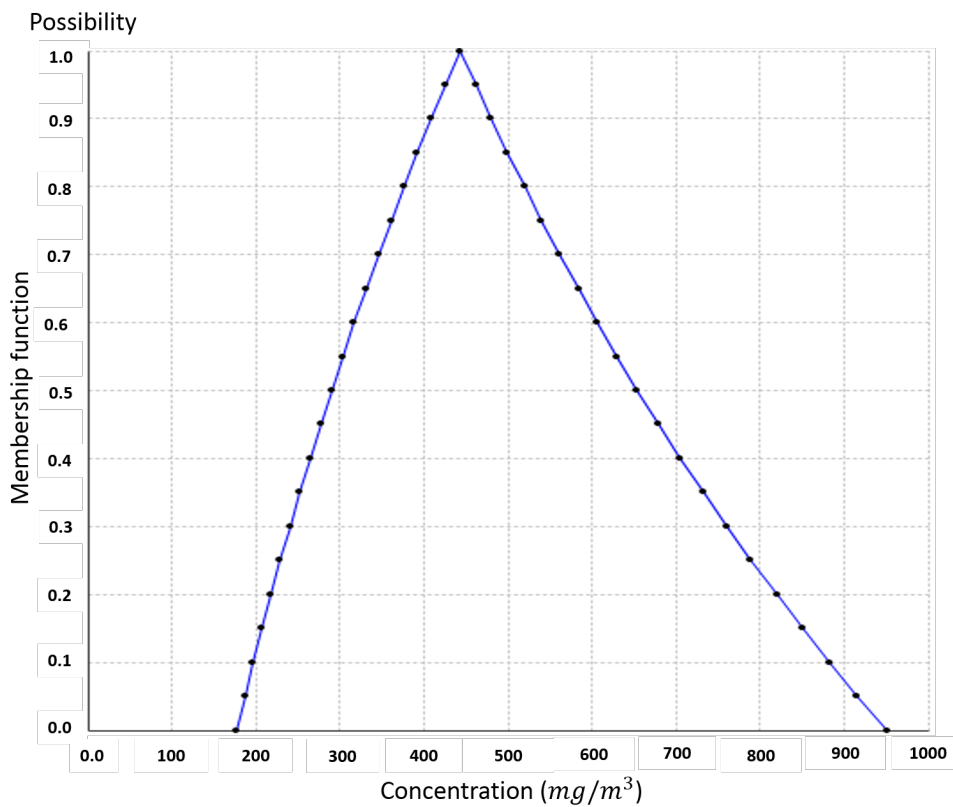
5.3 Comparison of the proposed hybrid approach with the existing approaches:  
applying uncertainty analysis to a Loss Of Containment scenario

distribution of the resulting fuzzy concentration at grid point  $C(500, 0, 0)$  is shown in figure 5.20. The lower, most likely and upper values are as follows:

$$C(500, 0, 0) = (177, 444, 951) \text{ mg/m}^3$$

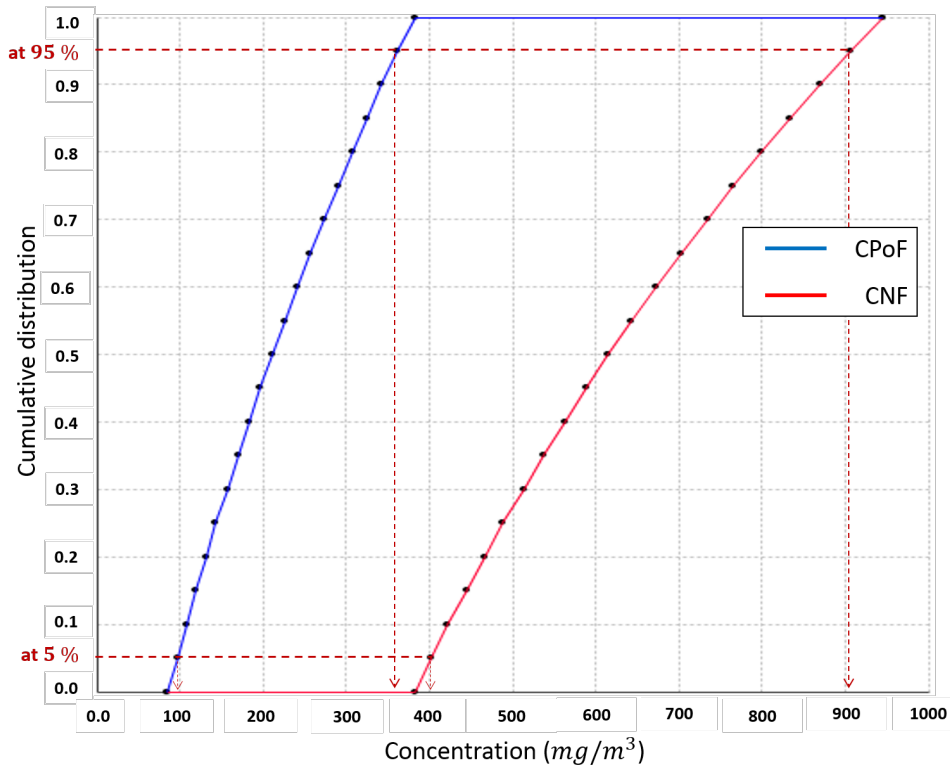
Input variables	Values
A, hole area	<i>TFN</i> $(2E - 3, 3.5E - 3, 5E - 3)$ [ $m^2$ ]
$C_d$ , discharge coefficient	<i>TFN</i> $(0.7, 0.8, 0.9)$
$U$ , wind speed	<i>TFN</i> $(3, 4, 5)$ [ $m/s$ ]

**Tableau 5.13** – Uncertain variables specified as triangular fuzzy numbers



**Figure 5.20** – Fuzzy concentration at grid point  $C(500, 0, 0)$

Not surprisingly, the range of the resulting concentration at  $\alpha = 0$  is the same as for interval analysis. This  $\alpha = 0$  range represents the most conservative range (to be sure that the true value is within the interval). A crisp value equal to 444 is obtained at  $\alpha = 1$ , which is not the middle of the interval at  $\alpha = 0$  as is the case for the inputs. Furthermore, a



**Figure 5.21** – Cumulative necessity function (CNF), and cumulative possibility function (CPoF) at grid point  $C(500, 0, 0)$

level of conservativeness is given by the different intermediate  $\alpha$  levels, which can render the decision-making more flexible. This conservativeness decreases when  $\alpha$  increases. The fuzzy cumulative distribution (cumulative possibility (CPoF) and necessity (CNF) functions) is presented in figure 5.21, where an interval is obtained at each percentile (%) (the minimum and maximum values are obtained from the CPoF and CNF respectively). The decision can either be optimistic or pessimistic based on either the CPoF or CNF respectively.

Moreover, this approach provides more information than interval analysis, which is useful both in terms of representing uncertainty and contributing to the decision-making process. Firstly, in terms of uncertainty representation, when experts believe that the true value is more likely to be somewhere inside the interval. Secondly, the possibility distribution shape can provide more information to help with the decision-making process (see figure 5.22). Figure 5.22 presents an example of how the distribution shape can be a useful tool for decision makers. The example represents decisions for the effect of risk (concentration) in a end-point using the fuzzy approach versus the interval approach. The red line specifies a decision criterion where the effect is acceptable if the concentration is lower than the criterion (left hand side). Using the interval approach only the maximum and minimum values provide information for representing uncertainty and it is clear that

### 5.3 Comparison of the proposed hybrid approach with the existing approaches: applying uncertainty analysis to a Loss Of Containment scenario

the maximum value exceeds the criterion. Therefore, the effect cannot be accepted since the likelihood of the concentration exceeding the criterion is not known. However, for clarification purposes, two cases based on the fuzzy approach can be considered. Two different decisions can be taken for these cases. According to the distribution shape, case 1 shows that it is highly likely that the concentration exceeds the criterion while case 2 shows the opposite. Hence the effect of risk cannot be accepted in case 1 while it can in case 2.

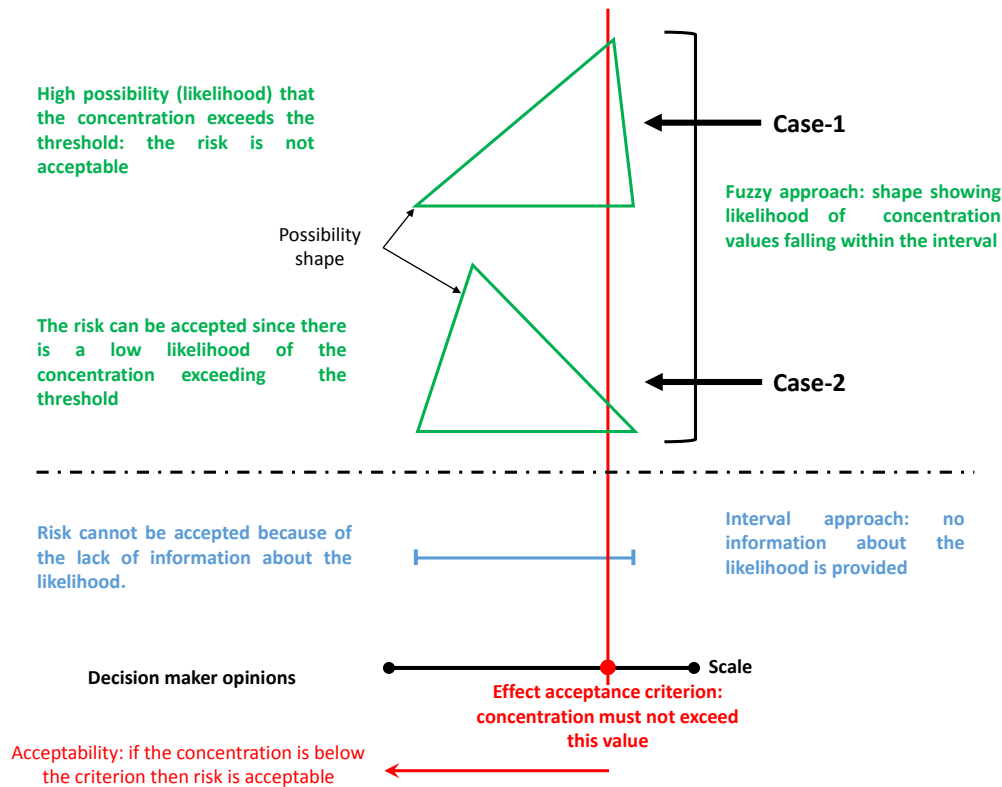


Figure 5.22 – How distribution shapes can contribute to decision making.

#### 5.3.6 Uncertainty analysis using the probabilistic approach

As mentioned in Section 3.1.2, the probabilistic approach uses probability distributions to represent uncertain parameters which are then propagated through an MC simulation. However, in order to build probability distributions, statistical data are required and these are often missing. This means that assumptions must be made in order to build the distributions. Thus, the assumptions made represent the uncertain parameters using probability distributions that have the same shape as the possibility distributions in the previous section. These assumptions show the difference between treating uncertainty using probability and fuzzy theories. In table 5.14, the distributions for the uncertain variables are all represented by triangular probability distribution (TPD).

Input variables	Values
A, hole area	<i>TPD</i> (2E - 3, 3.5E - 3, 5E - 3) [m <sup>2</sup> ]
<i>C<sub>d</sub></i> , discharge coefficient	<i>TPD</i> (0.7, 0.8, 0.9)
<i>U</i> , wind speed	<i>TPD</i> (3, 4, 5) [m/s]

**Tableau 5.14** – Uncertain variables represented using probability distributions.

Once the uncertain parameters have been specified using probability distribution,  $10^7$  iterations are conducted using the MC method in order to calculate the distribution of the output concentration. The output pdf and CDF at grid point  $C(500, 0, 0)$  are shown in figures 5.23 and 5.24 (the uncertain variables are assumed to be independent). Here, the decision can be made with a level of conservativeness or according to the compliance criteria on which the decision must be based. For instance, let us assume that according to the compliance criteria a concentration  $> x \text{ kg/m}^3$  should not have a probability higher than 5%. This information can easily be extracted and checked from the cumulative probability distribution at 95%. As shown in figure 5.24, there is a 95% certainty that the concentration will not exceed  $614 \text{ mg/m}^3$ .

### 5.3.7 Uncertainty analysis using Evidence theory

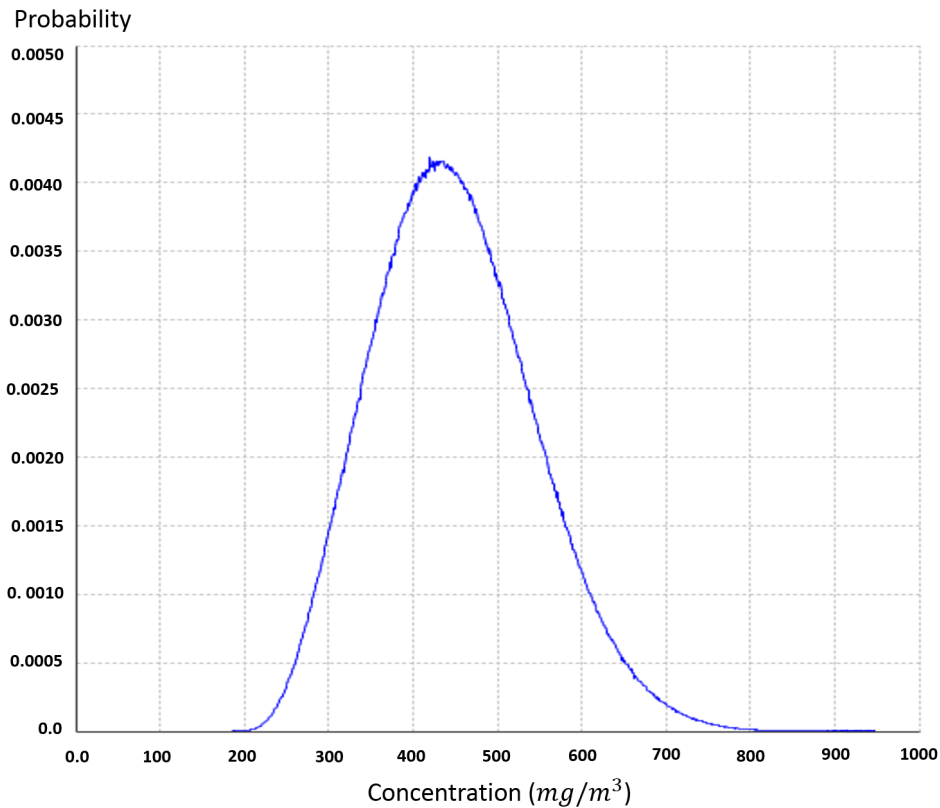
This section aims to represent the uncertain variables using a sets of focal elements with their bodies of evidence. The possibility distributions used in Section 5.3.5 are encoded into discrete focal elements as follows:

- generate  $q$  nested focal sets from the possibility distribution of the uncertain parameter  $X$  using  $\alpha$ -cuts  $X_{\alpha_i} = [\underline{x}_{\alpha_i}, \bar{x}_{\alpha_i}]$ , where  $i = 1, \dots, q$  and  $\alpha_1 = 1 > \dots > \alpha_q > \alpha_{q+1} > 0$ .
- build the bpa of the focal sets by assigning  $m_{\alpha_i} = \alpha_i - \alpha_{i+1}$

The focal elements are extracted from the fuzzy distribution in order to make comparisons between these two approaches (see Section 5.3.9). Table 5.15 provides the evidence spaces for the uncertain variables. These will be propagated through the model as explained in Section 3.1.4. Independence and total dependence between the input parameters are considered in order to see the effects of dependence on the output. The uncertainty in the output concentration will be represented in terms of an evidence space and summarized with the cumulative belief and plausibility functions.

The resulting CBF and CPF of the output evidence spaces at grid point  $C(500, 0, 0)$

5.3 Comparison of the proposed hybrid approach with the existing approaches:  
applying uncertainty analysis to a Loss Of Containment scenario

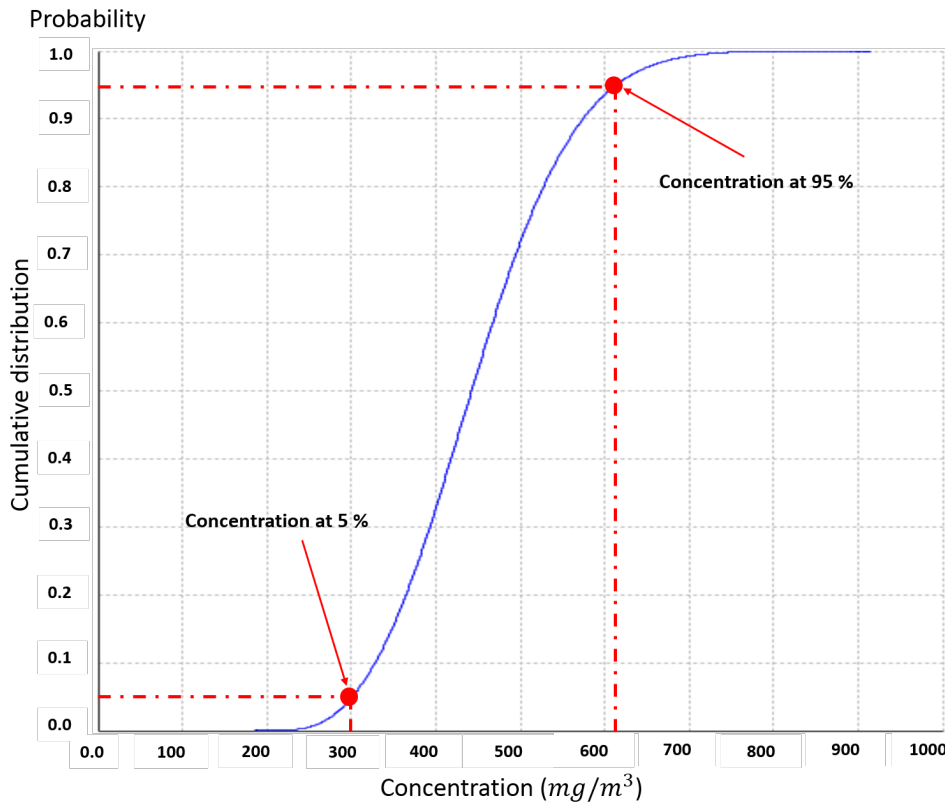


**Figure 5.23** – The pdf for the concentration at grid point  $C(500,0,0)$  using the Monte Carlo simulation

Input variables	Values
A, hole area	$(2E - 3, 3.5E - 3, 5E - 3), \quad q = 20 \quad m_{X_{\alpha_i}} = 0.05 \quad [m^2]$
$C_d$ , discharge coefficient	$(0.7, 0.8, 0.9), \quad q = 20 \quad m_{X_{\alpha_i}} = 0.05$
$U$ , wind speed	$(3, 4, 5), \quad q = 20 \quad m_{X_{\alpha_i}} = 0.05 \quad [m/s]$

**Tableau 5.15** – Uncertain variables specified as spaces of evidence

for both dependent and independent parameters are shown in figure 5.25. These two cumulative functions display the lowest and highest probabilities that may be assigned to an output value for each case. The lower and upper concentrations (minimum and maximum values) are the same as the two bounds obtained using interval analysis for the two cases of dependence. As with the probabilistic approach, a decision can be made with a level of conservativeness depending on the probability percentile taken but, as in the fuzzy approach, an interval is obtained instead of a single point. For example,  $[426, 916]$  and  $[370, 833]$  represent the concentration range at a likelihood percentile of 95



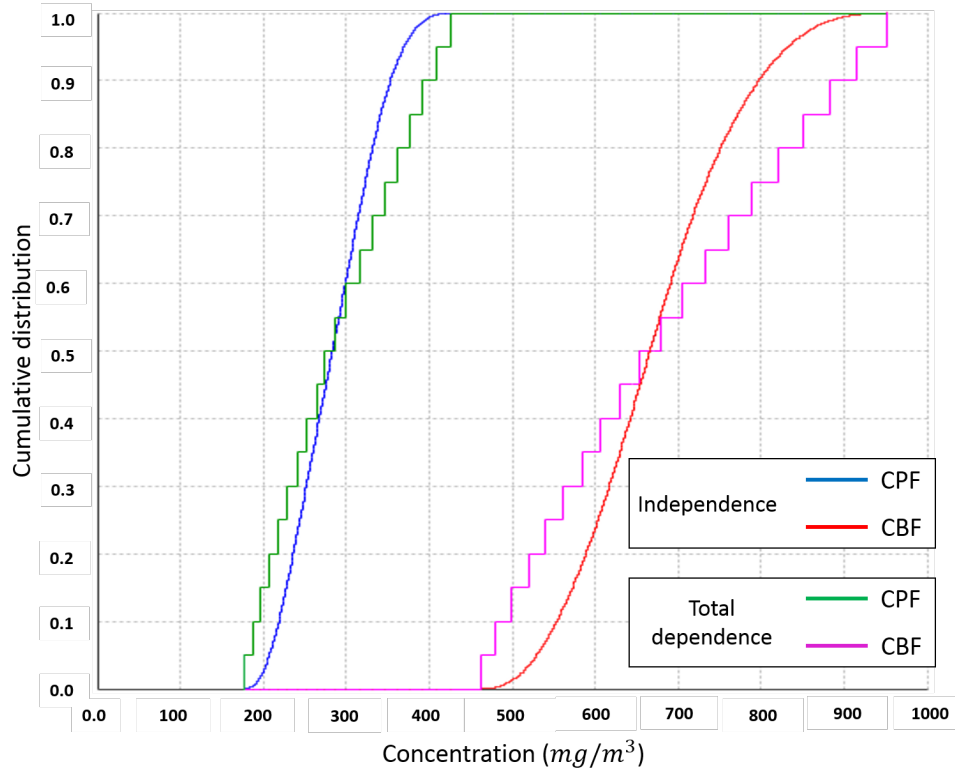
**Figure 5.24** – The CDF for the concentration at grid point  $C(500, 0, 0)$

for independence and total dependence respectively. However, there is clearly a difference between the two cases due to input dependence being considered. Furthermore, the outputs of the evidence approach, where total dependence is considered, and of the fuzzy approach are the same (see Figure 5.26 where the distributions do not totally coincide due to the discrete intervals chosen in the evidence approach, where  $q$  is equal to 20). This comparison between the evidence approach with independent input parameters and the fuzzy approach reveals that the cumulative distributions produced by the two approaches are different. This dissimilarity means that the fuzzy approach is significantly influenced by the total dependence of the uncertain input parameters.

### 5.3.8 Uncertainty analysis using the probabilistic-fuzzy approach

In this mixed approach, aleatory uncertainty is represented using probability distribution (if statistical information is available to build the distribution), while imprecision and subjectivity linked to expert elicitation are represented in terms of fuzzy numbers. It should be noted that FRVs are not used since the parameters affected by both types of uncertainty are not presented. Table 5.16 shows the representation of uncertain variables in terms of probability distributions and fuzzy numbers. As expected, this recently devel-

5.3 Comparison of the proposed hybrid approach with the existing approaches:  
applying uncertainty analysis to a Loss Of Containment scenario



**Figure 5.25** – The CBF and CPF for the concentration at grid point  $C(500, 0, 0)$

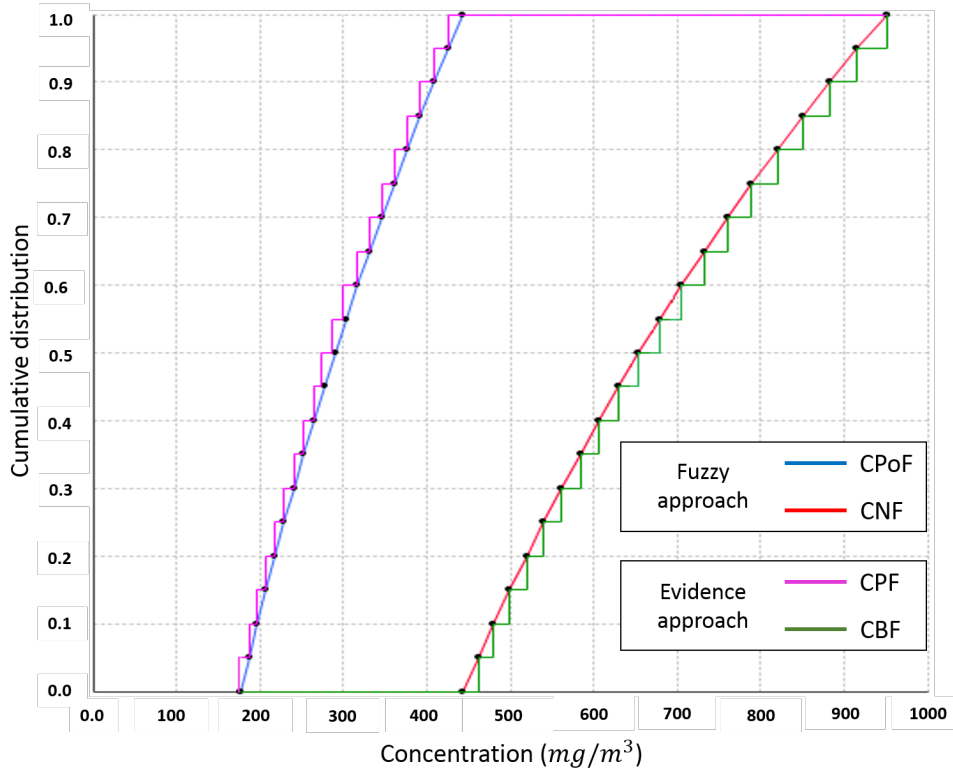
oped approach seems to be the most interesting and accurate to date. In order to build the distributions the probabilistic approach uses assumptions and guesswork when statistical data are not available while the fuzzy approach can neglect important information when experimental and statistical data are available. Hence, this recent approach offers an alternative to neglecting information and adding unjustified assumptions. Furthermore, it allows for separate treatment of both types of uncertainty (aleatory and epistemic).

Input variables	Values
$A$ , hole area	$TFN$ $(2E - 3, 3.5E - 3, 5E - 3)$ $[m^2]$
$C_d$ , discharge coefficient	$TFN$ $(0.7, 0.8, 0.9)$
$U$ , wind speed	$TPD$ $(3, 4, 5)$ $[m/s]$

**Tableau 5.16** – Uncertain variables represented using probability distributions

Based on the 2-stages MC method explained in Section 5.2.5.3, figure 5.27 shows the resulting fuzzy random variables for the concentration of HCL. The black lines represent the lower and upper distributions for  $\alpha = 0$ , whereas the red line represents the most likely distribution for  $\alpha = 1$ . Using this approach, a decision is based on a combination





**Figure 5.26** – Difference between the fuzzy approach and the evidence approach based on total dependence

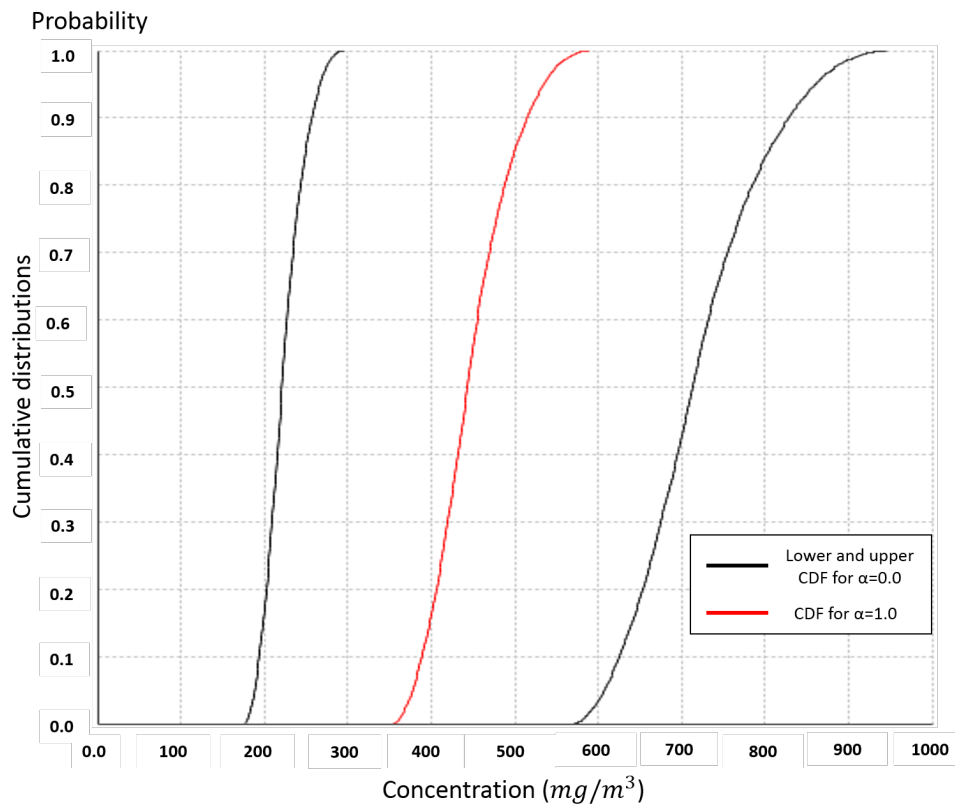
of probabilistic-fuzzy decisions where at each % of probability a fuzzy concentration is obtained. The minimum, maximum and most likely values for the fuzzy number at a confidence level of 95% are (267, 535, 861).

In what follows, the different approaches presented above shall be compared.

### 5.3.9 Comparison of all approaches

This section aims to compare the results obtained in the previous sections. Figure 5.28 displays the output derived from the interval approach, fuzzy approach, probabilistic approach, evidential approach and the probabilistic-fuzzy approach. The output distributions of the evidence approach with dependent input parameters are not depicted because they are the same as those obtained with the fuzzy approach.

For interval analysis there is no representation of likelihood. Instead, the minimum and maximum bounds are plotted on the Figure. These bounds represent the best and worst estimations (the most optimistic and pessimistic values), respectively. A decision based on this approach would be too conservative and may result in the risk being over-estimated since there is no distribution for the likelihood of the values falling within the interval. Thus, a distribution shape inside the interval may allow for greater decision-



**Figure 5.27** – Lower, most likely and upper distributions for the concentration at grid point  $C(500, 0, 0)$

making flexibility.

This shape is given by each of the other approaches. Furthermore, the minimum and maximum values provided by each one are exactly the same as for the interval approach. However, a difference between these shapes can be observed in spite of the fact that the initial information used is the same. To be able to compare these approaches, a confidence interval of 90% is chosen from each one. Table 5.17 shows the confidence interval for each approach. For the probabilistic approach, the two values at confidence levels equal to 5% and 95% represent the minimum and maximum bounds of the confidence interval. For the fuzzy approach, on the other hand, an  $\alpha$  – cut interval at  $\alpha = 0.05$  is shown (the same interval can be obtained from the CPoF and CNF, see table 5.17 where the minimum and maximum values of the confidence interval are the CPoF at 5% and the CNF at 95% respectively). For evidence theory, the confidence interval is derived in the same way as for the fuzzy approach where the CPF and CBF are used instead of the CPoF and CNF. For the mixed probabilistic-fuzzy approach, the confidence levels at 5% and 95% are fuzzy numbers. The minimum and maximum values at 5% and 95% for  $\alpha$  – cut = 0 are used to build the confidence.

Table 5.17 clearly shows that the probabilistic approach is the most optimistic ap-

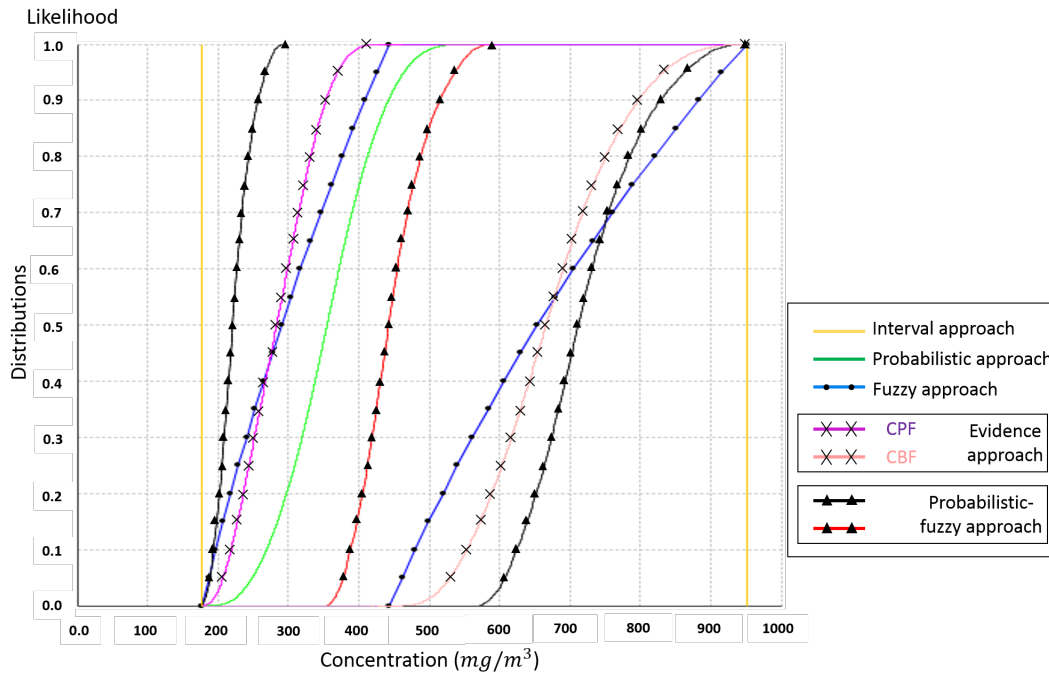


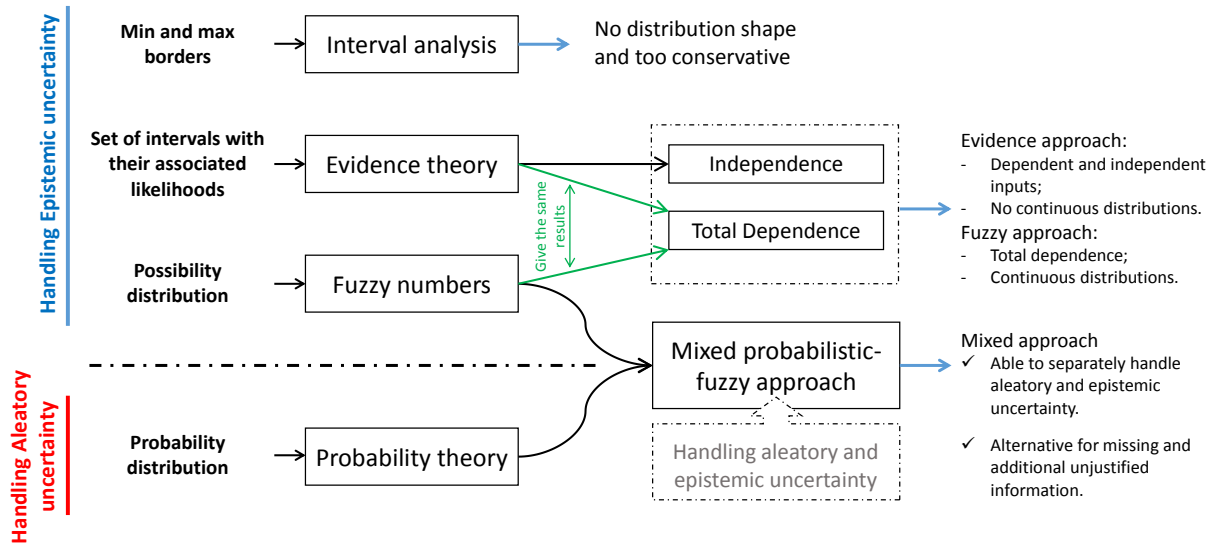
Figure 5.28 – Comparison of the results generated from the different approaches.

Approach	Confidence levels		Confidence Interval	Gap size
	5%	95%		
Probabilistic	302	614	[302, 614]	40.3%
Fuzzy	[187, 462]	[426, 916]	[187, 916]	94.2%
Evidence	[208, 530]	[370, 833]	[208, 833]	80.7%
Probabilistic-fuzzy	(188, 443, 608)	(267, 535, 861)	[188, 861] at $\alpha = 0$	86.9%

Tableau 5.17 – 90% confidence intervals for each of the approaches used.

proach as it presents the smallest gap between the confidence intervals used. The fuzzy and evidence approaches based on dependent input parameters, on the other hand, are the most pessimistic as they present the largest gap between the borders of the 90% confidence interval. The confidence interval for the fuzzy approach is more conservative than the evidence approach based on independent input variables. This is due to the fact that the fuzzy approach is automatically based on total dependence. The size of the gap produced by each approach is presented in the last column of table 5.17 (the largest interval obtained for the interval approach is assumed to have a gap size equal to 100%). Moreover, the confidence intervals obtained for the fuzzy, evidence and mixed approaches cover the confidence interval obtained for the probabilistic approach. However, these observations can be interpreted in two different ways: they are either due to (i)

### 5.3 Comparison of the proposed hybrid approach with the existing approaches: applying uncertainty analysis to a Loss Of Containment scenario



**Figure 5.29** – Summary of the characteristics of the approaches highlighted in this comparison

preciseness or (ii) non-conservativeness. The probabilistic approach is precise if the probability distributions used to represent uncertainty are based on justifiable statistical data (i.e. real data and not assumptions). The approach is non-conservative if imprecise or subjective information is used, as is the case with expert elicitation. Non-conservativeness may result in an underestimation of the risk and subsequently to wrong decision making. On the other hand, if all the parameters were represented by fuzzy numbers or evidence spaces, despite the fact that some could be justifiably represented by PDFs, the range of results would be too conservative and statistical information would be lost. This is why the mixed approach offers an effective compromise between lack of information and risk underestimation. The confidence interval obtained for the mixed approach lies between the intervals obtained for the probabilistic and fuzzy approaches and offers a more precise gap than the evidence approach.

Figure 5.29 summarizes the differences and similarities between the approaches described. Probability theory offers the best approach for dealing with aleatory uncertainty when statistical data are available. The evidence and fuzzy approaches, on the other hand, deal with epistemic uncertainty. Both approaches can provide the same or a different output according to whether or not input parameter dependence is taken into account. As already stipulated, the fuzzy approach automatically considers that the input parameters are totally dependent while the evidence approach offers this as an option. Hence, it is not desirable to use fuzzy numbers with independent parameters.

Based on these studies, the guidelines for choosing the best approach can be given as follows:

- use probability distributions if there are enough statistical data to build them;

- use fuzzy theory if the information is imprecise and there is total dependence between the input parameters;
- use evidence theory in the case of independent input parameters and if the available information can be interpreted as a set of non-nested intervals with different likelihoods (cannot be continuously represented). It should be noted that, unlike fuzzy theory, evidence theory uses non-continuous distribution for representing uncertainty. In addition, evidence theory is the best theory to deal with ignorance (see Section 3.1.4.3);
- if statistical information is available for some parameters whereas subjective information or minimum and maximum values are available for others, a mixed proper algorithm will produce a more credible and accurate analysis allowing the decision to be made with the best available information.

However, important research questions can be generated here: what if all these causes of parameter uncertainty affect the same analysis? How uncertainty should be represented? And how uncertainty can be propagated if fuzzy numbers, probability distributions, evidence focal elements and FRVs are used as inputs. The answer to these questions is presented in the next section. The next section proposes a global parameter uncertainty analysis approach that treat all causes of parametric uncertainty within the same analysis.

## **5.4 A global approach to treat all types and causes of uncertainty in effect analysis: the ALI-Aggregated Likelihood Index**

As we proved in the previous sections of this Chapter that the use of a wrong mathematical tool to represent data that are uncertain may lead to miss-evaluation of risk. From the literature review, a global approach that correctly deal with parameter uncertainty regarding the causes of this uncertainty does not exist. Epistemic uncertainty has different causes and each cause should be treated apart and with the right mathematical representation. Imprecision and vagueness if affect an analysis, they should not be treated with the same mathematical tool. Uncertainty analysts claim that the best theories to treat imprecision and vagueness are fuzzy and evidence theories, respectively. However, existing approaches for modeling and quantifying parameter uncertainty do not take this problem into consideration. Existing approaches model epistemic and aleatoric uncertainties within a single mathematical representation, or using a hybrid approach that treat epistemic and aleatoric uncertainties separately. Hybrid approaches model the epistemic

uncertain parameters by the same mathematical representation, even if they have different causes of uncertainty.

For these reasons, a new global approach is needed. The main questions that our approach is going to answer are: (i) How parameter uncertainty regarding their causes are going to be modeled or represented? Section 5.4.1 answers this question in details. (ii) How the represented data are going to be propagated through the risk models? A propagation algorithm is developed as presented in Section 5.4.2. And (iii) how decision is going to be made under uncertainty?. A decision making guidelines is presented in Section 5.4.3. We name this approach the ALI (Aggregated Likelihood Index) approach because it combines different likelihoods indexes (probability, belief and possibility) in a single framework.

This approach is used when all causes of uncertainty affect the same analysis. In other words, some input parameters for the analysis are random, others are imprecise and available information about some others is incomplete. In this case, probability theory, fuzzy numbers and evidence theory and FRVs are used together to best deal with uncertainty. The framework of the proposed approach is presented in Figure 5.30.

#### **5.4.1 Uncertainty representation: characterizing each uncertain parameter based on the available information**

For an effective uncertainty analysis, the types and causes of parameter uncertainty define the selection of mathematical theories to analyze uncertainty. This section aims to present how analysts can characterize uncertain parameters with best suitable mathematical theories regarding the available information. For example, if a given parameter is deterministic and known, it will be represented by a precise value, such as the pipe length or diameter. If a parameter varies randomly, then it will be represented by a probability distribution. If a parameter is affected by epistemic uncertainty but sufficient information to build the probability distribution is available, then probability distribution is the appropriate representation. If the information is qualitative and imprecise, it will be represented by fuzzy membership function or possibility distribution. If the input information is incomplete, then evidence theory is the best to use. Other types of data might need different representations. Figure 5.31 shows a flow chart on how uncertain parameters can be appropriately represented regarding the available information.

Moreover, the chart shows the appropriate representation for other types of data including conflictive data (lack of consensus between data sources if multiple sources of data are used); and imprecise statistical information (The distribution is known but its parameters are imprecise). Lack of consensus or if information regarding a parameter is imprecise and incomplete at the same time are represented using BOE (Body Of Ev-

idence) with fuzzy focal elements. Fuzzy focal elements are used instead of intervals in order to represent imprecision or subjectivity in data from different sources. Section 5.4.1.1 presents more details in terms of an example.

The developed flow chart in Figure 5.31 proves that there is no one knowledge representation theory that fits all, in other word that can characterize all causes of parameter uncertainty. As listed, uncertain parameters can be characterized by probability distributions, fuzzy numbers (includes interval analysis), fuzzy random variables, or evidence with fuzzy focal elements regarding the available information. A MC-based algorithm to propagate these representation together through the risk model is developed in the next section.

#### 5.4.1.1 Evidence theory with fuzzy focal elements

Focal elements of an evidence can be *fuzzy numbers* instead of normal intervals. This representation allows the analysts to give shapes inside the intervals in the case if they believe that some values are more likely than the others. This representation provides a mathematical tool to combine knowledge from different sources in the same formalism .

The use of input data from multiple sources provides better approximation of risk and makes the analysis more reliable. Whereas, it may lead in inconsistency in the inputs since different sources may give different elicitation regarding the same parameters. In the remainder of this section, we present how to use BOE (Body Of Evidence) with fuzzy focal elements to combine data derived from multiple sources. The knowledge from each source of data can be fuzzy numbers or evidence structures. The general fusing equations to combine all data in the same evidence formalism are presented as follows:

$$P = \cup p_{ij} \quad \text{and} \quad m_{ij}(P) = \frac{w_j}{\sum w_j} m_{ij} \quad (5.9)$$

Where  $P$  is the set of all focal elements for the output after combination,  $p_{ij}$  are the focal elements from the source of data (expert or database)  $j$ ,  $m_{ij}$  are the BPMs of focal elements  $p_{ij}$  and  $w_j$  is the weighting factor given to the source of data  $j$  (expert). A weighting factor represents the experience and the familiarity of the expert in the field under study.

To better understand the combination algorithm, let us take a parameter A which is uncertain. Regarding the uncertainty in A, two different expert knowledge are elicited about its value. Expert-1 said, A is most likely to be equal 0.5. While A is somewhere between  $[0.3, 0.4]$  or it is with more likelihood to be in  $[0.6, 0.7]$  based on the opinion of expert-2. In addition, it is more possibly to be near to the center of the interval in the both cases. We supposed that the two experts have different ratings, a weighting values

equal 0.4 and 0.6 are given to expert-1 and expert-2 respectively. However, it is clearly seen that a lack of consensus between these two experts is presented. Information from Expert-1 is represented in terms of a TFN, while a Fuzzy-DS structure is used to represent the knowledge from Expert-2 (see figure 5.32). Note that a *bpa* equal 1 is given to the TFN of expert-1 to perform the aggregation. Based on the combination equations explained above, the output fuzzy-DS structure is depicted in figure 5.32. The next section explains how to interpret this representation and construct the fuzzy belief and plausibility.

#### 5.4.1.2 Construction of the Fuzzy belief and plausibility for a BOE with fuzzy focal elements

The construction of cumulative belief and plausibility when the focal elements are fuzzy numbers can be performed using  $\alpha - cut$  principles. In the example (figure 5.32), at each  $\alpha - cut$ , an interval is generated from each fuzzy number (the blue and red lines represent the intervals for  $\alpha - cut = 0.0$  and  $0.5$  respectively where single points is obtained for  $\alpha - cut = 1.0$ ). Thus, at each  $\alpha$ , a CBF and CPF are generated. In figure 5.32, CBF and CPF for three different  $\alpha - cut$  are depicted (the cumulative belief and plausibility at  $\alpha - cut = 1.0$  coincide). Finally, a fuzzy CDF from these cumulative belief and plausibility functions is obtained. The fuzzy CDF can be discretized into percentiles where a fuzzy membership function is obtained at each percentile. Also figure 5.32 shows the fuzzy value of A at 60 percentile. Finally, the strength of this new representation exists in its capability of benefiting from all the available data and without conflicting, neglecting nor adding of unjustified knowledge.

### 5.4.2 Uncertainty propagation

In this section, input parameter are represented using probability distributions, fuzzy numbers, body of evidences and FRVs. In other words, we suppose that different causes of parameter uncertainty affect the same analysis. This step aims to propagate all these representations through the risk model. However, propagating uncertainty here can be complex and time consuming depending on the complexity and the type of the risk model.

In the following we will detail the steps on how to propagate uncertainties through the risk model using 2-stages (steps) MC simulation and all possible combinations. Figure 5.33 shows the steps required to apply this approach. Let us assume that

$$Z = f(Y) = f(A_i, B_j, C_k, D_l) = f(A_1, \dots, A_n, B_1, \dots, B_m, C_1, \dots, C_q, D_1, \dots, D_t)$$

representing the risk model where  $A_i, i = 1, \dots, n$  are affected by epistemic uncertainty due to ignorance or lack of consensus,  $B_j, j = 1 \dots m$  by aleatory uncertainty,  $C_k, k = 1, \dots, q$  by



epistemic uncertainty due to imprecision, and  $D_l, l = 1, \dots, t$  by a mixture of both types of uncertainty due to variability and imprecision.

1. set  $h = 0, v = 0, v = 0$  and L, M, N are the number of all possible combinations of fuzzy focal elements, and samples for the first and second loop MC respectively;
2. from each parameter represented using a body of evidence, a fuzzy number with its mass degree are taken from each body of evidence. Thus, at each h, we obtain  $a_{hi} = (a_{h1}, \dots, a_{hn})$  which are fuzzy numbers.
3. from each stochastic and mixed uncertainty (represented by fuzzy random variables), a value (crisp value from each probability distribution) and a triangular fuzzy number (from each fuzzy random variable) are obtained using the first Monte Carlo sampling:  $b_{vj} = (b_{v1}, \dots, b_{vm})$  and  $d_{vl} = (d_{v1}, \dots, d_{vt})$ , where  $d_{vl}$  are triangular fuzzy numbers;
4. a second Monte Carlo loop is performed to take samples from the fuzzy numbers (those obtained from step 2, those taken in the first MC loop (step 3) and those reflecting epistemic uncertainty due to imprecision) as explained above; this means that

$$y = f(a_{hw1}, \dots, a_{hwn}, b_{v1}, \dots, b_{vm}, c_{w1}, \dots, c_{wq}, d_{vw1}, \dots, d_{vwt})$$

where  $a_{hwi}, c_{wk}$  and  $d_{vwl}$  are crisp values from the triangular fuzzy numbers;

5. calculate  $z_{vw} = y = f(a_{hw1}, \dots, a_{hwn}, b_{v1}, \dots, b_{vm}, c_{w1}, \dots, c_{wq}, d_{vw1}, \dots, d_{vwt})$  to obtain a crisp value with its membership and mass degrees. The membership of  $z_{vw}$  is obtained using the extension principle of fuzzy numbers as follows:

$$(\alpha(z_{hvw})) = \text{minimum of } (\alpha(a_{hw1}), \dots, \alpha(a_{hwn}), \alpha(c_{w1}), \dots, \alpha(c_{wq}), \alpha(d_{vw1}), \dots, \alpha(d_{vwt}))$$

if  $z_{hvw}$  has been obtained earlier from another sample then the maximum membership between it and the old membership is given to  $z_{hvw}$ . The mass degree of  $z_{hvw}$  is obtained by multiplying the mass degree off all input  $a_{hw1}, \dots, a_{hwn}$  (the mass degree of  $a_{hwi}$  is the same degree as  $a_i$ );

6. if  $w < N$ , return to step 4, otherwise go to 7;
7. generate the fuzzy number  $z_v$  from the N calculated samples;
8. if  $v < M$ , go back to step 3, otherwise go to 8;
9. after the second MC is completed, M triangular fuzzy numbers are obtained that represent a FRV. The CDFs of these results are plotted.
10. if  $h < L$ , go back to step 2, otherwise go to 9;

11. after the simulations, L FRVs are obtained, where a mass degree is attached to each one. In other word, we can say that the output is a body of evidence where the focal elements are FRVs.

The output obtained after propagating all these representation of uncertainty are hard to be interpreted. A bunch of FRVs with their mass degrees is obtained. This forms a body of evidence where the focal elements are FRVs. But we believe that this is the best way to represent parameter uncertainty regarding its causes. For these reasons, the next section proposes a method that simplifies the shape of the obtained representation of uncertainty for the output. This method aims to combine the likelihoods in a single likelihood index to provide a more meaningful representation that is simpler to be interpreted.

However, this propagation algorithm is used when the both types of parameter uncertainty affect the same analysis and due to different causes. Simpler propagation algorithms can be used depending on the causes of uncertainty that affect the input data as summarized in the flowchart of Figure 5.34.

### 5.4.3 Decision Making under uncertain environment

The shape obtained as result from propagating uncertainty in terms of a body of evidence with FRV focal elements is quite complex to be interpreted. In this section, we propose a new simpler representation for the output obtained from the previous section.

This new representation aims to aggregate the likelihoods obtained in the previous section in a single likelihood index (ALI) as presented in Figure 5.35. From the previous section, at each 2-stages sample using MC, as crisp value  $y_{hij}$  with its  $(m, \mu, p)$  are obtained. Here we multiply  $m, \mu$  and  $p$  to generate a single index  $I$ . Thus, at the end of the propagation step, a single distribution is obtained instead of a BOE with FRV focal elements.

In Figure 5.35, if the cumulative ALI distribution does not reach 1,  $1 - I_{max}$  implies ignorance. If this ignorance is not accepted, analysts should gather more information to reduce it and re-perform the analysis.

An important and criticizing question can be asked here: why do not we use probability distributions to represent all types of uncertainties if the output of ALI methodology looks like a probability distribution. The answer of this question will be given in the next section after the application of the ALI approach. The next section applies the ALI method and the pure probabilistic approach on the LOC case study. The objective is to compare the output provide by the ALI methodology against the pure probabilistic approach.

It should be noted that decision making can be always based on the output of Section 5.4.2. But here we wanted to provide simpler information for decision makers.

#### 5.4.4 Case study - Applying of the developed approach to a loss of containment scenario

In this section, the proposed methodology is applied to the same LOC case study presented in Section 5.3. The same mathematical models for the discharge and the dispersion are used as presented in Eqs 5.7 and 5.8, respectively. The objective is to calculate the concentration at the end point (500, 0, 0) taking parameter uncertainty into consideration. However, uncertainty regarding the input parameters is studied in more details in this case study.

Figure 5.36 presents the uncertain parameters and the related causes of this uncertainty. The wind speed is statistically known in terms of a probability distribution, but the parameters of this distribution are imprecise. In addition, uncertainty in calculating the horizontal and vertical dispersion coefficients is considered. A new method based on fuzzy logic is developed to calculate  $\sigma_y$  and  $\sigma_z$ . Fuzzy logic is introduced to deal with imprecision and vagueness in determining the stability classes used for calculating the dispersion coefficients. Section 5.4.4.1 details the causes of parameter uncertainty in calculating the dispersion coefficients and how introducing of fuzzy logic is the solution for this issue.

##### 5.4.4.1 Introducing fuzzy logic to analyze parameter uncertainty in calculating $\sigma_y$ and $\sigma_z$

As explained in Section 5.3.1,  $\sigma_y$  and  $\sigma_z$  are calculated depending on the stability class. The atmospheric stability class is determined regarding two input parameters: (i) the surface wind speed (WS) and (ii) the daytime incoming solar radiation (SR), or cloud cover (CC) at night as presented in Section 5.3.1, Table 5.8. This scheme is extensively used due to the availability of data for the wind speed and the cloud cover provided by the national weather institutions [38]. But it has some limits because of uncertainty.

From Table 5.8, five different wind speed intervals are taken by Pasquill. For the incoming solar radiation during the day, three different linguistic terms are used. Strong refers that the solar elevation is greater than  $60^\circ$ , while moderate refers that the solar elevation lies between  $35^\circ$  and  $60^\circ$ . If the solar elevation is less than  $35^\circ$ , then the incoming solar radiation is referred to as slight. For the Cloud Cover at night, two ranges are used: thinly overcast (if it is  $\geq 4/8$  cloud) or clear overcast (if it is  $\leq 3/8$  cloud).

Thus, input parameters to determine the stability class are represented in terms of intervals (WS) or linguistic terms (SR and CC). However, because of discreteness between the input intervals and linguistic terms, conditions on the border between two intervals or two linguistic terms are vague and not well defined. However, replace intervals and

linguistic terms by fuzzy numbers will remove vagueness and imprecision.

Intervals of WS and linguistic terms of SR and CC are represented using fuzzy numbers as shown in Figure 5.37. Using fuzzy scales, uncertainty on the borders between two categories is removed. Precise values of inputs for WS, SI or CC can be used instead of just intervals or linguistic terms. Where each input value is belong to one or two categories. For example, a WS of value equals 1.9 is very low with a possibility degree equals 0.7 and low with a possibility degree equals 0.3 as depicted in Figure 5.37.

Determining the stability classes now is based on Table 5.18. In Table 5.18, the wind speed intervals are exchanged by linguistic terms that signify fuzzy numbers. Based on the proposed fuzzy scheme, if we have quantitative values for WS and SR, one or more stability classes with different possibility degrees can be obtained. After the fuzzy stability classes being determined, the Briggs' Fitting Constants  $X_i$  where  $X \in \{a, b\}$  and  $i \in \{y, z\}$  are calculated based on Eq 5.10 and Table 5.9.

Wind speed	Daytime incoming solar radiation			Night-time Cloud Cover	
(m/s)	Strong	Moderate	Slight	$\geq 4/8$	$\leq 3/8$
Very low	A	B	B	-	-
low	B	B	C	E	F
Moderate	B	C	C	D	E
High	C	D	D	D	D
Very high	C	D	D	D	D

**Tableau 5.18** – Determining the stability class based on the proposed fuzzy scale

$$X_i = \frac{\sum_{n=A}^F \mu_{SC}(n) \times X_{it}(n)}{\sum_{n=A}^F \mu_{SC}(n)} \quad (5.10)$$

where  $n$  is a stability class,  $\mu_{SC}(n)$  is the membership degree of being in class  $n$ ,  $X_{it}(n)$  is the old fitting constant determined from Table 5.9 of being in class  $n$ . At the end, the dispersion coefficients are calculated by placing the values  $X_i$  in Eq 5.10.

$$\sigma = ax(1 + bx)^{-1/2} \quad (5.11)$$

As an example, let us suppose that the WS is equal to 1.9 and the solar elevation angle is  $30^\circ$ . each of these two quantities returns to two different linguistic terms in the fuzzy scale (see Table 5.19). WS is VeryLow and Low with membership degrees equal 0.7

and 0.3, respectively. And SR is slight and moderate with membership degrees equal 0.6 and 0.4, respectively. Then based on the proposed fuzzy scheme, the stability classes are obtained as presented in Table 5.19. From Table 5.19, the output stability classes are B with a membership degree equals 0.6 and C with membership degree equals 0.3.  $a$  and  $b$  are calculated using Eq 5.10 and Table 5.9 as follows:

$$a_y = \frac{0.6 \times 0.16 + 0.3 \times 0.11}{0.6 + 0.3} = 0.143;$$

$$a_z = \frac{0.6 \times 0.12 + 0.3 \times 0.08}{0.6 + 0.3} = 0.106;$$

$$b_y = \frac{0.6 \times 0.00001 + 0.3 \times 0.00001}{0.6 + 0.3} = 1.0e^{-05};$$

$$b_z = \frac{0.6 \times 0 + 0.3 \times 0.0002}{0.6 + 0.3} = 6.666666666666667e^{-05};$$

$\sigma_y$  and  $\sigma_z$  are determined using Eq 5.11:

$$\sigma_y = a_y x (1 + b_y x)^{-1/2} = 0.143 * 500 (1 + 0.00001 * 500) = 71.4881$$

$$\sigma_z = a_z x (1 + b_z x)^{-1/2} = 0.106 * 500 (1 + 6.666666666666667e^{-05} * 500) = 52.466$$

Determining the stability classes based on the fuzzy scheme		Wind Speed (WS) = 1.9	
		Very low: 0.7	Low: 0.3
Solar Radiation (SR) = 30°	Moderate: 0.4	B: 0.4	B: 0.3
	Slight: 0.6	B: 0.6	C: 0.3

**Note**  
The class B appears in different cases in the table with different membership degrees. The highest membership degree is taken for the output.  
The output stability classes are: B and C with membership degrees equal 0.6 and 0.3, respectively.

**Tableau 5.19** – Example of determining the stability classes based on the fuzzy scheme

It should be noted that WS, SR or CC still can be represented in terms of linguistic terms if quantitative data is unavailable.

#### 5.4.4.2 Calculation of the concentration based on the proposed approach

In this section, calculating the concentration of HCL is conducted. Information relating to these uncertain parameters presented below:

- $A$ : somewhere between  $[20E - 3, 50E - 3]$  but more likely to be close to the central tendency;
- $C_d$ : between the interval  $[0.7, 0.9]$  but more likely to be close to the central tendency;
- $U$ : based on statistical data,  $u$  follows a uniform distribution of inputs  $(3, U_{mean}, 5)$ , where  $U_{mean}$  is imprecise.  $U_{mean}$ : measurements show that the mean of the distribution is an interval  $[3.8, 4.2]$ ;
- $SE$ : based on experts' opinions, during the day, incoming solar radiation moderate half of the daytime, strong 40% of the daytime and they do not know the rest of the time. The "do not know" verbal expression here represents the ignorance.

Based on this information, table 5.20 presents the used theory to represent each uncertain input parameter.

Input variables	Values
$A$ , hole area	$TFN (2E - 3, 3.5E - 3, 5E - 3) [m^2]$
$C_d$ , discharge coefficient	$TFN (0.7, 0.8, 0.9)$
$U$ , wind speed	$FRV = TPD(3, U_{mean}, 5) [m/s]$ , where $U_{mean} = TFN(3.8, 4, 4.2)$
$SE$ , solar elevation angle	$BOE = \{moderate : 0.5, strong : 0.4, ignorance : 0.1\}$

**Tableau 5.20** – Uncertain variables represented using either probability distributions, fuzzy numbers, FRV and body of evidence

Based on the 2-stages MC with evidence calculus method depicted in Figure 5.33, Figure 5.38 shows the resulting ALI distribution for the concentration of HCL at the grid point  $C(500, 0, 0)$ . The ALI distribution does not reach 1 as presented in Figure 5.38. This represents the ignorance about the calculation of the concentration of HCL. To present the effectiveness and simplicity of decision of the ALI approach, the second section presents a comparison between this proposed approach against the pure probabilistic approach.

#### 5.4.4.3 Comparing the proposed approach with the pure probabilistic approach

This section aims to compare the proposed ALI approach with the pure probabilistic approach. The pure probabilistic and the ALI approaches are performed on the same LOC case study presented in the previous section. But here, uncertainty about the

solar radiation is not considered since the pure probabilistic approach can not deal with ignorance as discussed in Section 5.4.1.1. The representation of uncertain input parameter for both the ALI and the pure probabilistic approaches are presented in Tables 5.21 and 5.22, respectively. It should be noted that a 2-stages MC simulation is to propagate uncertainty for the both ALI and the pure probabilistic approaches.

Input variables	Values
A, hole area	$TFN (2E - 3, 3.5E - 3, 5E - 3) [m^2]$
$C_d$ , discharge coefficient	$TFN (0.7, 0.8, 0.9)$
$U$ , wind speed	$FRV = TPD(3, U_{mean}, 5) [m/s]$ , where $U_{mean} = TFN(3.8, 4, 4.2)$
$SE$ , solar elevation angle	moderate

**Tableau 5.21** – Uncertain variables represented using probability distributions, Fuzzy numbers and FRVs

Input variables	Values
A, hole area	$TPD (2E - 3, 3.5E - 3, 5E - 3) [m^2]$
$C_d$ , discharge coefficient	$TPD (0.7, 0.8, 0.9)$
$U$ , wind speed	$TPD = TPD(3, U_{mean}, 5) [m/s]$ , where $U_{mean} = TPD(3.8, 4, 4.2)$
$SE$ , solar elevation angle	moderate

**Tableau 5.22** – Uncertain variables represented using probability distributions

In order to compare these approaches, a confidence interval of 90% is chosen from each approach as presented in Table 5.23. The minimum and maximum bounds of this interval are respectively the two values at confidence levels equal to 5% and 95% for the ALI or the probability distributions. From Table 5.23, the 90%-ALI confidence interval is more conservative than confidence interval extracted from the pure probabilistic approach. This is because we have chosen the right theory to represent each uncertain input parameter as discussed in Section 5.3.9.

Approach	Confidence levels		Confidence Interval
	5%	95%	
ALI approach	188	400	[188, 400]
Pure probabilistic approach	203	374	[203, 374]

**Tableau 5.23** – 90% confidence intervals for the ALI-approach vs the pure probabilistic approach.

## 5.5 Conclusion

Given the substantial development of high-risk industries today, and the expansion of populations around them, a risk analysis must be capable of building a certain amount of confidence in the results so as to support the decision making process. This level of confidence, or accuracy, is difficult to achieve when the analysis inputs suffer from imprecision, lack of information, vagueness and variability. In this case, handling parameter uncertainty is the most important and expensive part of the analysis if confidence in the results and accurate risk predictions are to be achieved.

Effect analysis is an important tool in the decision-making process. To be still more valuable, effect analysis must take into account parameter uncertainty. This chapter proposes two approaches to take into account the different types of parameter uncertainty in the model parameters commonly used for assessing the effects of risks. The objective of these approaches is to allow an uncertainty representation and propagation that is consistent with the information actually available. A growing number of researchers in the field of risk analysis promote the importance of separately representing information variability (aleatory) and lack of information in the evaluation of risk.

The first approach combines fuzzy theory and randomness to separately tread variability and imprecision. This approach also presents a propagation method based on a 2-stages Monte Carlo simulation developed for propagating all types of parameter uncertainty and illustrated through a case study of a chemical reactor. The results of this study can contribute to the decision-making process for the management of this chemical reactor. The approach described demonstrates that the system can be safe whatever the time of failure.

The proposed fuzzy-probabilistic approach is compared to the approaches presented in Chapter 3 Section 3.1. Five theories or approaches for dealing with uncertainty are compared: interval analysis, probability theory, evidence theory, the fuzzy approach and the mixed probabilistic-fuzzy approach.



These methods are applied to the atmospheric dispersion of a toxic cloud formed after a total loss of containment in a pressurized vessel of HCL. This scenario presents a major interest for two reasons: the HCL is a very common substance with many uses, and it is a toxic, corrosive and non-flammable substance. Two risk models are used to predict the toxic concentration at a specific end-point: a discharge model to calculate the source term and a Gaussian type dispersion model. Model input uncertainties are characterized and propagated to obtain a prediction.

The results show that the treatment of uncertainty using interval analysis is the easiest and also the least demanding in terms of information required (the interval borders are the only information needed). However, this method proves to be the most pessimistic and the least representative (only minimum and maximum values are provided). It can also result in missing information (if the data available exceed the two borders). Fuzzy analysis provides more information than just an interval. Based on possibility distribution, this method calculates the likelihood of the values being inside the interval. It is also very useful for representing imprecision (when conservativeness is needed) and subjective expert elicitation when historical and empirical data are not provided. However, it considers that the input parameters are totally dependent and this can influence the output. The probabilistic approach is easy to interpret but it may result in the risk being underestimated when no historical information is provided and assumptions must therefore be made in order to build the probability distributions. The evidence approach is less pessimistic than the interval approach (the distribution shape is provided) and less demanding in terms of information than probability theory. The evidence approach provides the same result as the fuzzy approach in the case of dependent input parameters. It is preferable to use the evidence approach in the case of independent input parameters. Finally, the treatment of variability in a non-probabilistic approach when empirical data are provided results in a large deviation in the confidence level and can lead to the risk being overestimated and a pessimistic decision being taken. The probabilistic-fuzzy approach offers a solution in this respect as it produces the most accurate results where imprecision, subjectivity and variability are treated separately. This mixed approach represents an alternative with regard to missing information and risk underestimation but assumes total dependence between epistemic uncertain parameters. However, as underlined in this chapter, parameter uncertainty characterization (representation) must be managed correctly so that it can then be propagated through risk models and produce a satisfactory output.

Based on this comparison, a new approach that treat each cause of parameter uncertainty with the right theory is needed. We handle this problem by combining randomness, possibility and belief. We extended the fuzzy-probabilistic approach by introducing evidence theory to represent ignorance and incompleteness of information. A new prop-

agation algorithm is developed that uses 2-stages MC simulation and evidence calculus. The approach is named the ALI-approach because we aggregated the different likelihood index (probability, possibility and belief) in a single one. The output of the ALI approach after aggregating the likelihood indexes is a single distribution. Aggregating likelihood indexes provides simpler output for decision makers.

The ALI approach is demonstrated on the same atmospheric dispersion case study mentioned above in this section. Then ALI approach is compared with the pure probabilistic approach. The ALI approach provides the best suitable representation of uncertainty regarding the available information. Moreover, guidelines on how representing and propagating parameter uncertainty regarding the available information and the types of risk models are also provided.

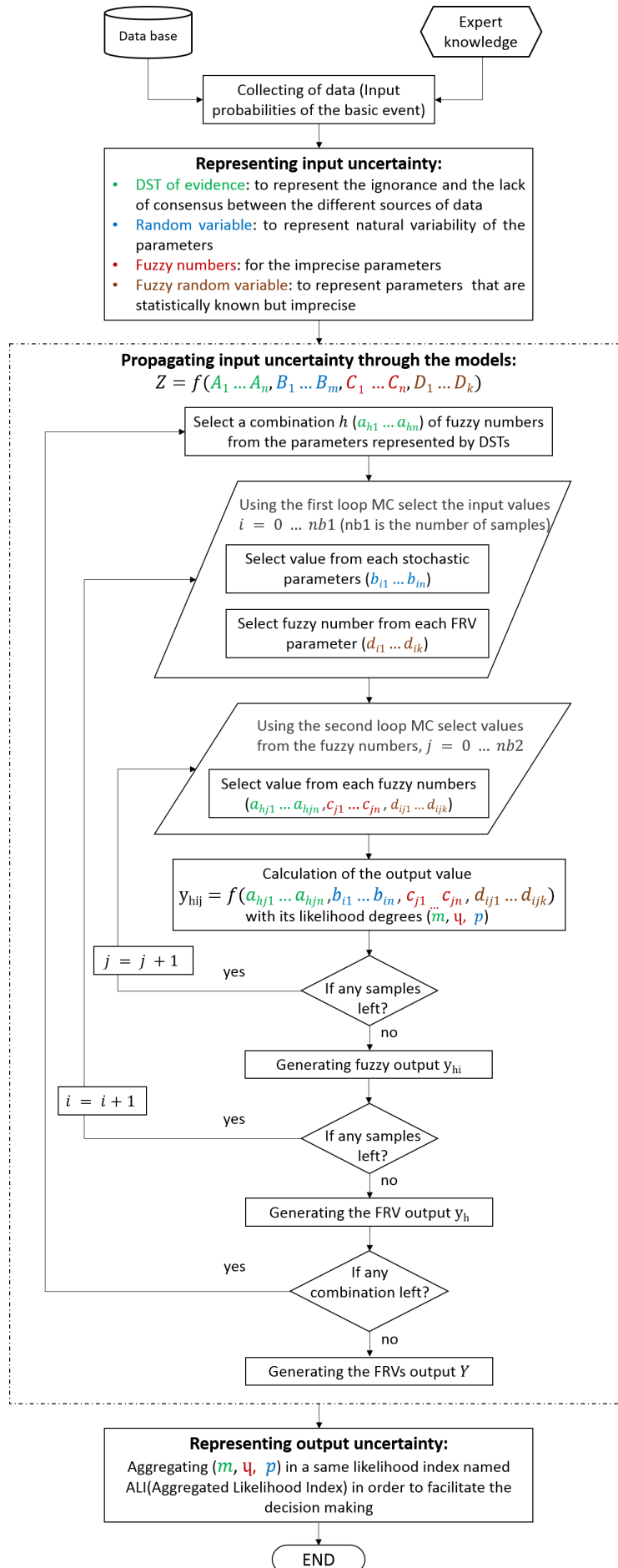


Figure 5.30 – The ALI framework to deal with uncertainties

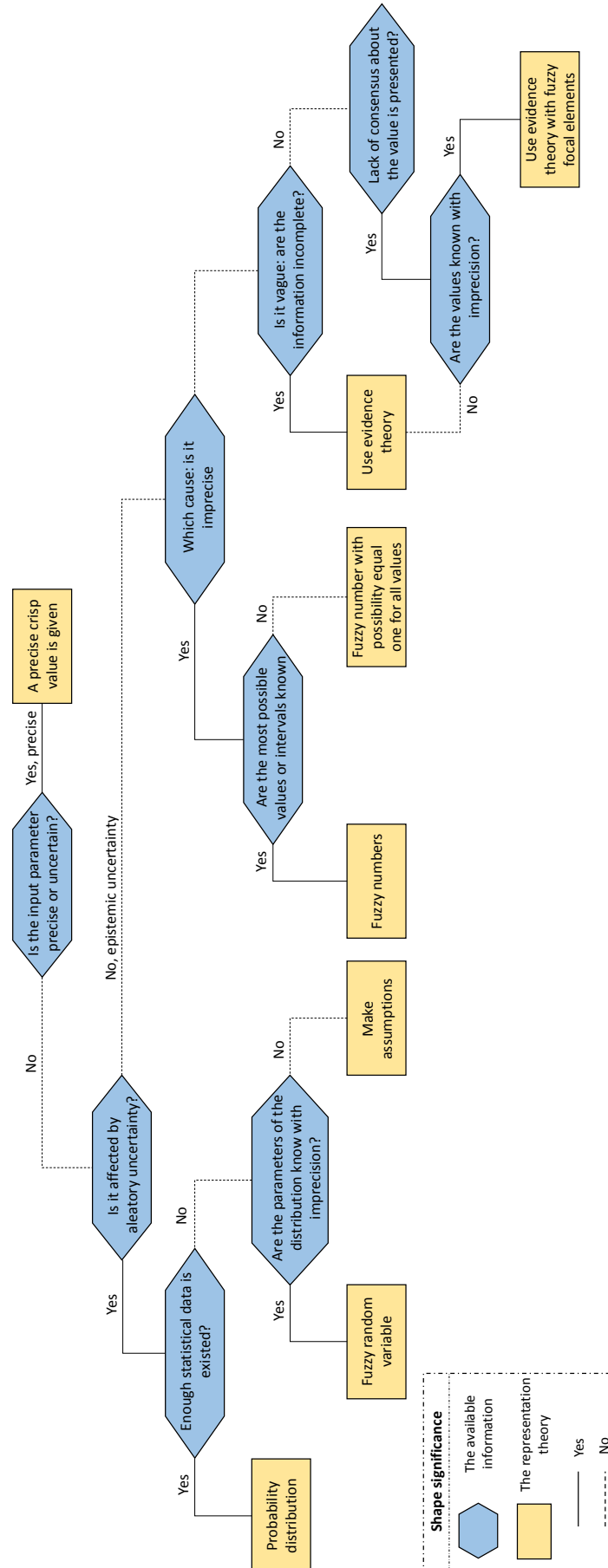


Figure 5.31 – Flow chart for representing parameter uncertainty based on the available information

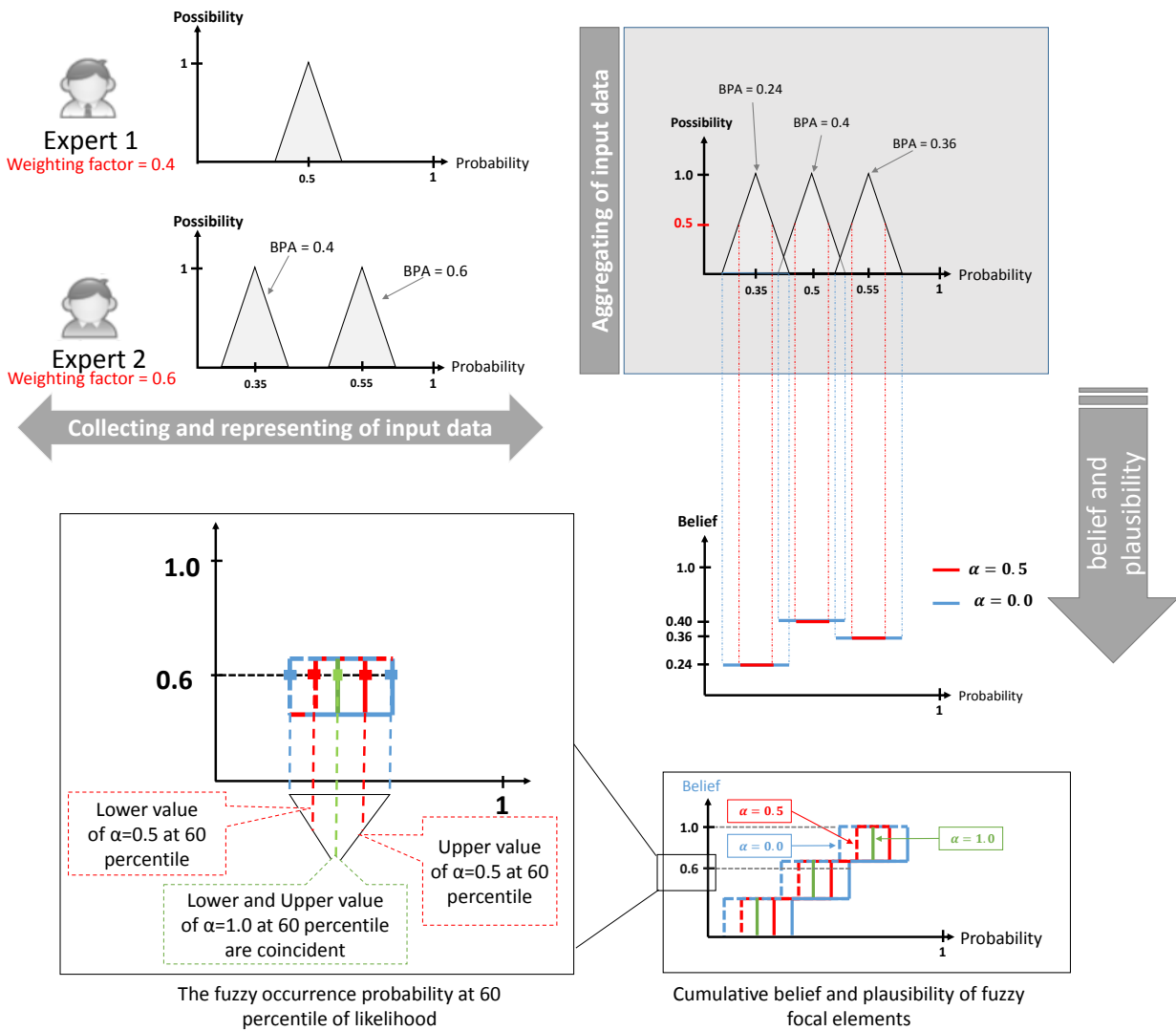


Figure 5.32 – Aggregating multiple sources of data in the same formalism.

Parameter A: The lack of consensus between the different sources of data used are represented in terms of DST with fuzzy focal elements  
 Parameter B: Parameters that known statistically represented by a probability distribution  
 Parameter C: Source term parameter that known with imprecision  
 Parameter D: Fuzzy random variable is used to handle the two type of uncertainty that presented in the same parameter

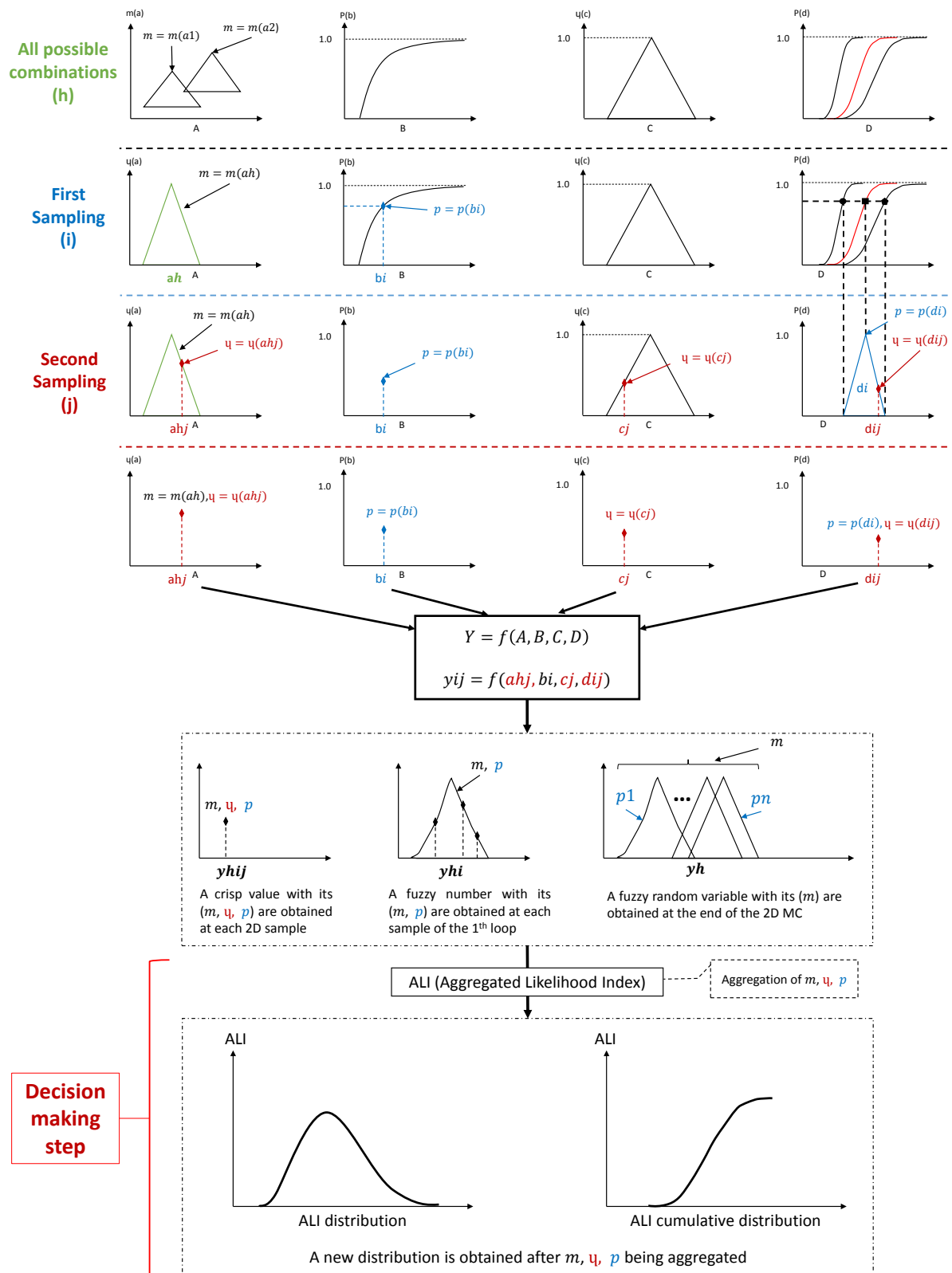
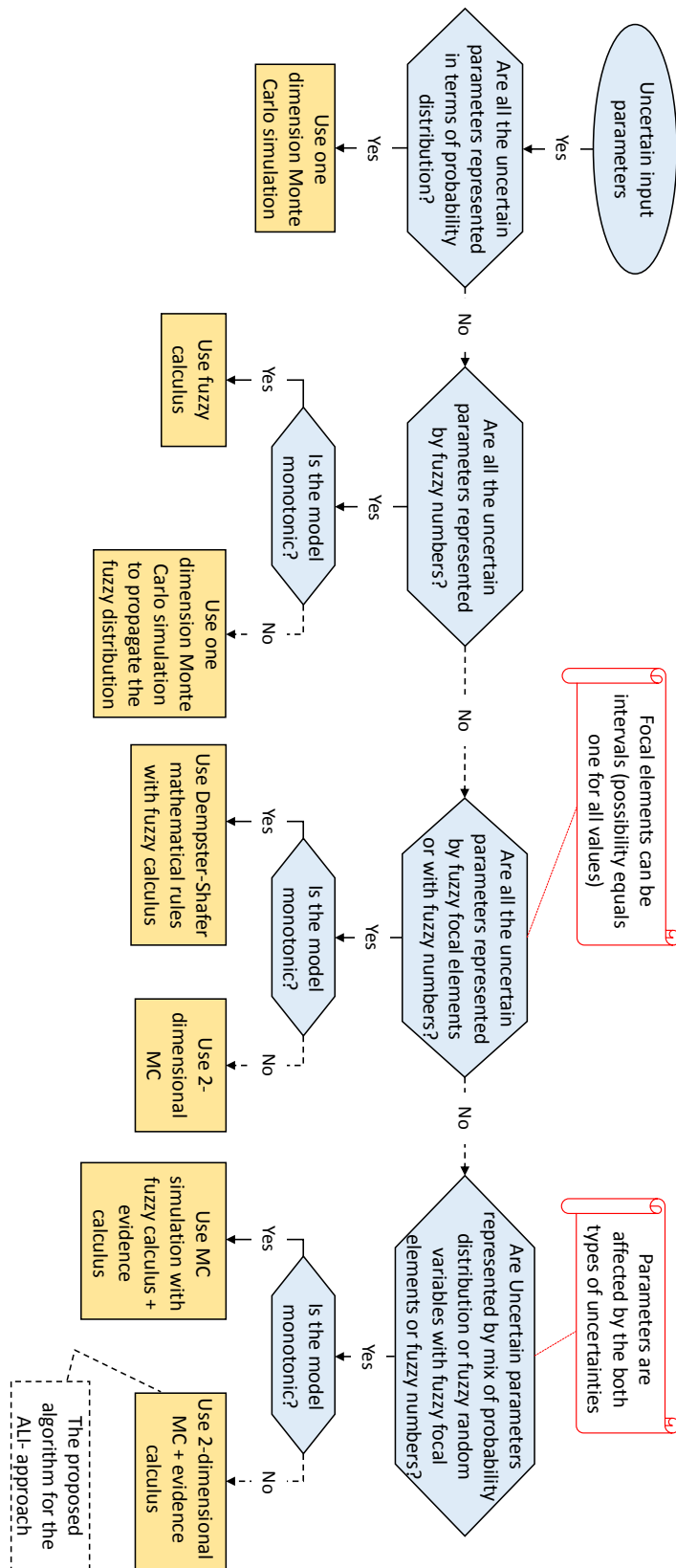


Figure 5.33 – Representing and propagating of uncertainty within the ALI methodology



**Figure 5.34** – Decision tree for selecting the more appropriate uncertainty propagation technique regarding the risk model and uncertain input data

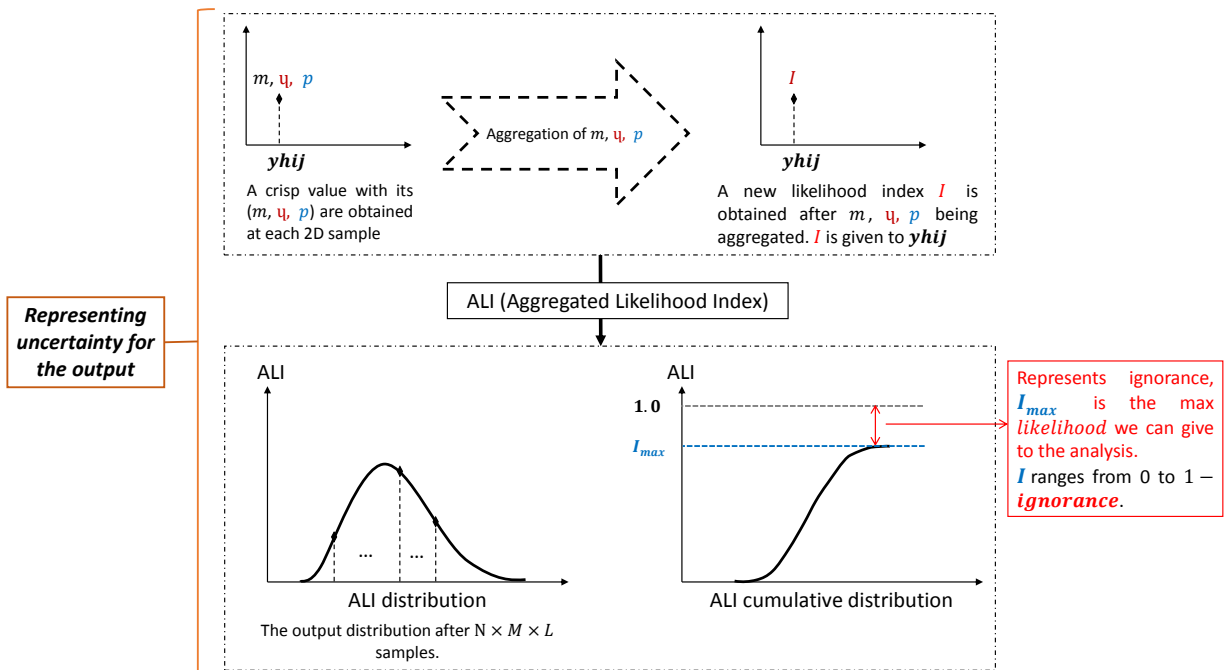


Figure 5.35 – Aggregating the likelihood in a single likelihood index

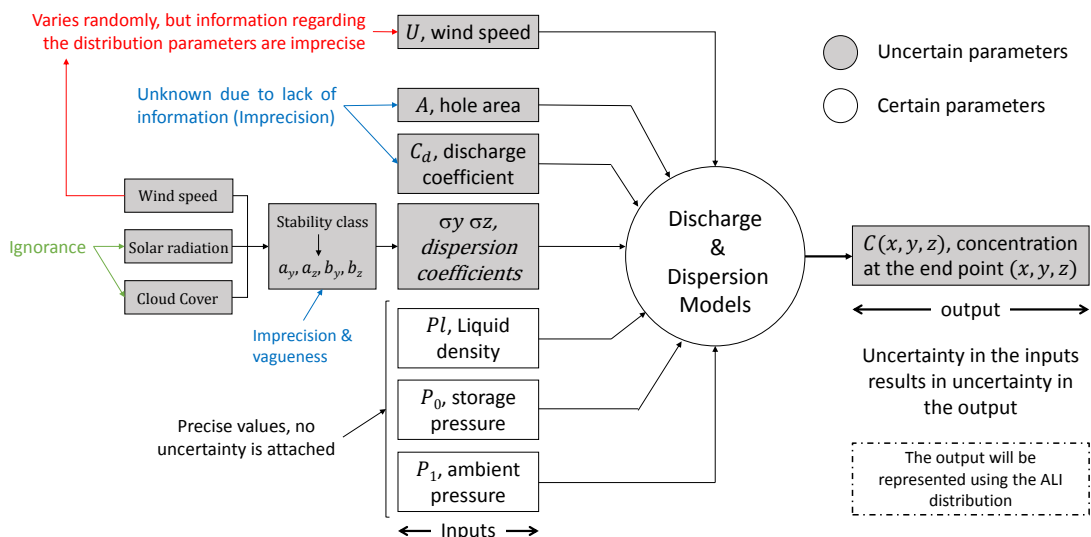


Figure 5.36 – Uncertain and precise input parameters used to calculate the concentration at a specific end-point



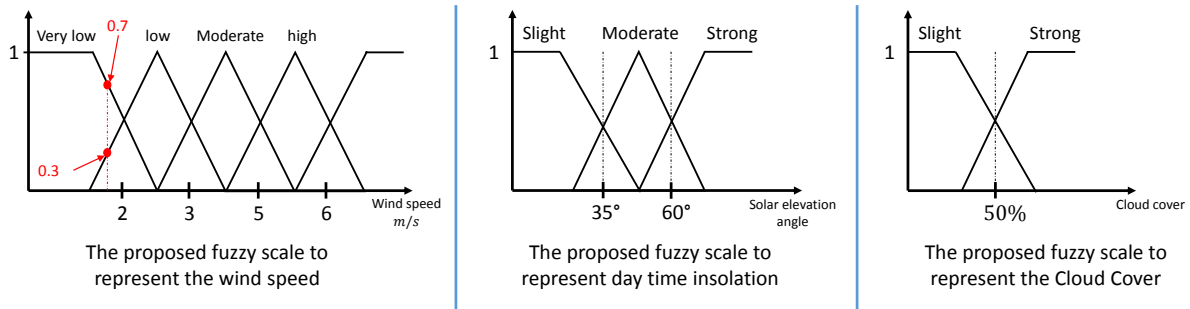


Figure 5.37 – The proposed fuzzy scales for the wind speed and the day time insolation

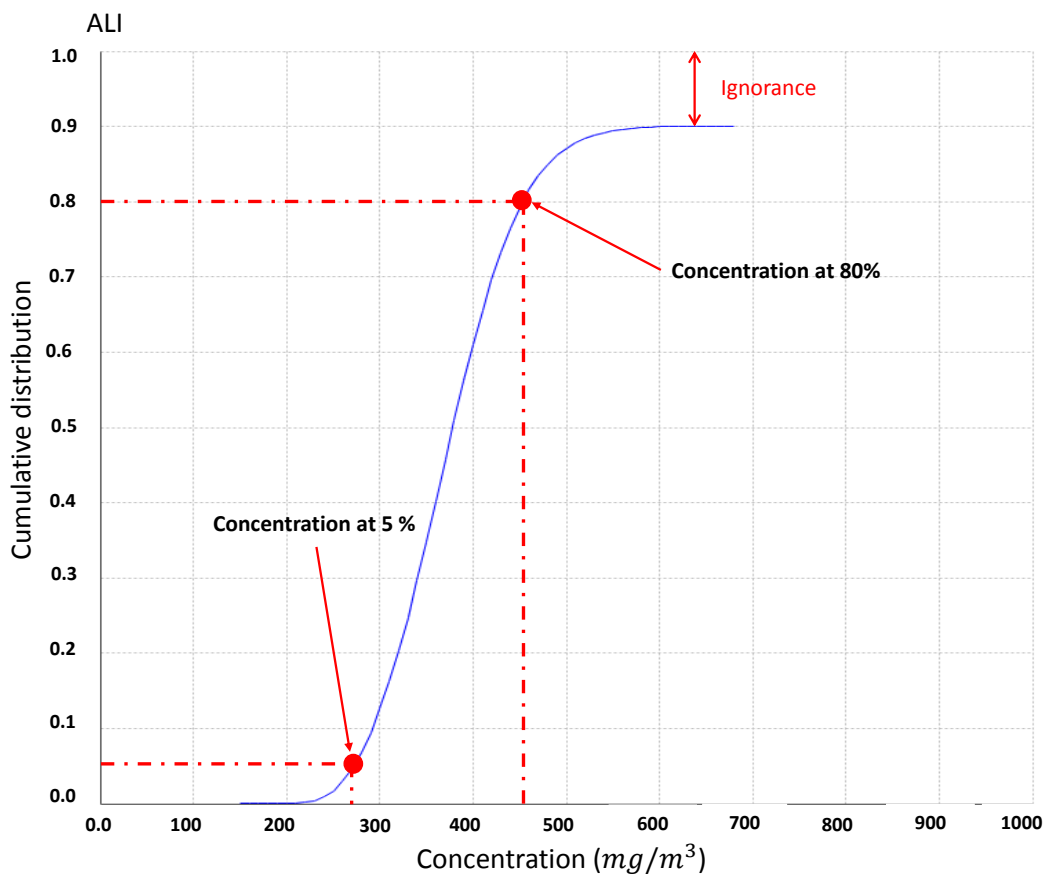
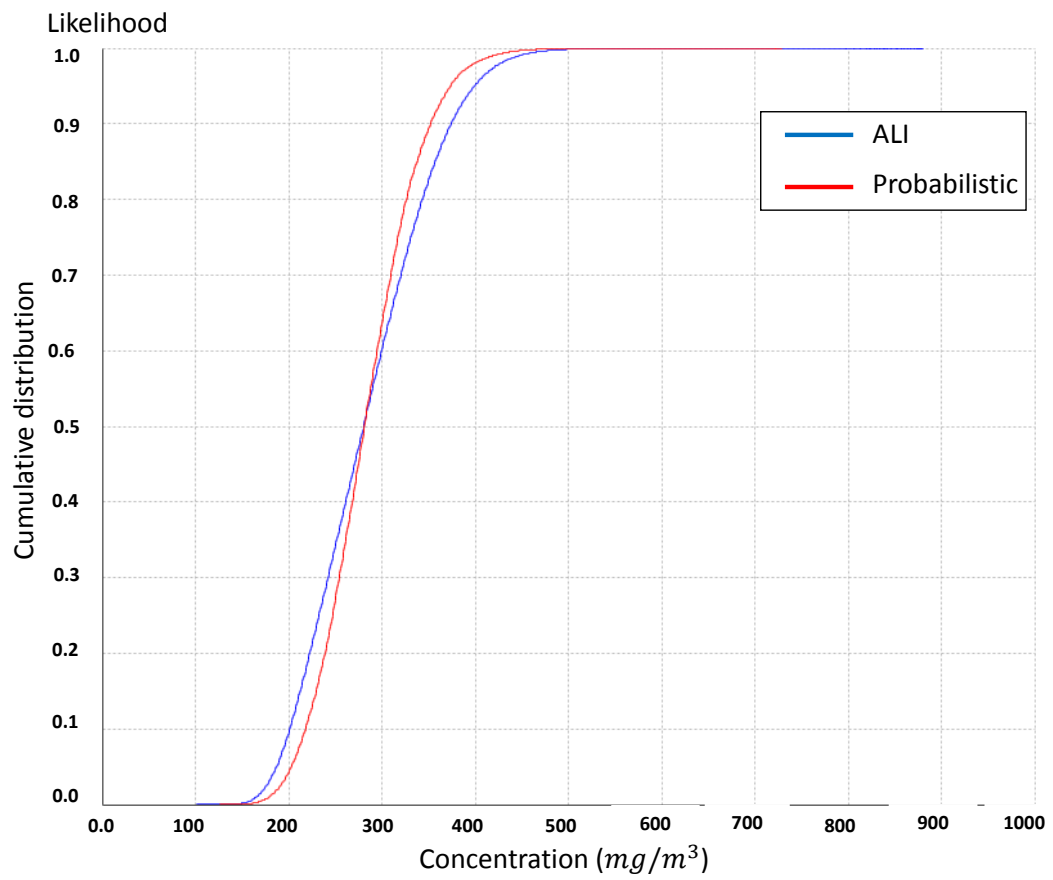


Figure 5.38 – The ALI distribution for the concentration at grid point C(500, 0, 0) using the 2-stages MC simulation



**Figure 5.39** – The cumulative ALI distribution vs the cumulative pure probabilistic distribution for the concentration at grid point  $C(500, 0, 0)$



# 6

## Handling one aspect of completeness uncertainty: introducing cyber-security within industrial risk analysis

**Summary:** In this chapter, a new method that considers safety and security together during industrial risk analysis is proposed. This approach combines bowtie analysis, commonly used for safety analysis, with a new extended version of attack tree analysis, introduced for security analysis of industrial control systems. We then propose an approach for evaluating the risk level based on two-term likelihood parts, one for safety and one for security.

### Summary

---

<b>6.1</b>	<b>Introduction</b>	<b>174</b>
<b>6.2</b>	<b>Global definition of industrial risk</b>	<b>174</b>
6.2.1	Definition of risks related to safety	175
6.2.2	Definition of risks related to security	175
6.2.3	Definition of risks related to safety and security	175
<b>6.3</b>	<b>Methodology for combined safety/security risk analysis</b>	<b>176</b>
6.3.1	Introduction behind the global idea	176
6.3.2	Step-1: representation of a risk scenario	178
6.3.3	Step-2: likelihood evaluation	182
<b>6.4</b>	<b>Case study</b>	<b>188</b>
6.4.1	Description	188

6.4.2	Application . . . . .	189
6.4.3	Discussion and improvement . . . . .	191
<b>6.5</b>	<b>Modeling Stuxnet using the ATBT . . . . .</b>	<b>192</b>
<b>6.6</b>	<b>Limitations and future work . . . . .</b>	<b>195</b>
<b>6.7</b>	<b>Conclusion . . . . .</b>	<b>196</b>

---

---

This work was carried out in collaboration with the INERIS. The results developed in this chapter have been presented in the following articles:



[9] H. ABDO AND J-M. FLAUS AND F. MASSE. *A safety/security risk analysis approach of Industrial Control Systems: A cyber bowtie - combining new version of attack tree with bowtie analysis.* Computers & Security , – (2017).



[6] H. ABDO AND J-M. FLAUS AND F. MASSE. *Towards a better industrial risk analysis: A new approach that combines cyber security within safety.* In *Safety and Reliability Theory and Applications: Proceedings of ESREL*, 1080-1089, Portoroz, Slovenia (2017).

## 6.1 Introduction

This Chapter provides a guidance on addressing one cause of known completeness uncertainty; i.e., the cyber-security related risk contributors that have not been included in the scope of the risk analysis. As discussed in Section 2.6.3, the analysis process should be exhaustive and up-to-date with the new risks to support the decision under consideration. It should be noted that in this study we are interested in cyber-security breaches that can lead to major accident hazards. In other words, accidents that have effects on human life and the environment and not on confidentiality, integrity or availability of information.

The contributions of this Chapter are as follows:

- proposing a new definition of risk that cover safety and security (Section 6.2): combine safety and security definitions of risk in one global definition;
- proposing a risk analysis approach that introduces security within safety (Section 6.3):
  - ✓ developing new modeling technique for identifying and representing safety/security risk scenarios that can lead to major accidents (Section 6.3.2): a new modeling technique is developed by combining the bow-tie analysis with a new extended attack tree;
  - ✓ proposing a likelihood analysis methodology to evaluate the probability of occurrence of safety/security risk scenarios (Section 6.3.3): a qualitative likelihood analysis methodology based on the combined attack tree/bow-tie is proposed. This methodology uses two-term likelihood parts to evaluate the likelihood level of a risk scenario, one for safety and one for security.
- demonstrating the developed approach in real case studies: the application of this approach is demonstrated for safety/security risk analysis using the case study of a risk scenario in a chemical facility (Section 6.4). Section 6.5 demonstrates the applicability of the proposed attack tree in the security domain by modeling the Stuxnet virus.

The chapter finishes by giving hints for future work and drawing some conclusions as presented in Sections 6.6 and 6.7, respectively.

## 6.2 Global definition of industrial risk

In this section, we present the definitions of safety and security related risks (Sections 6.2.1 and 6.2.2, respectively). These two definitions will be used to generate a new global definition of industrial risk that covers safety and cyber-security related risks and suited for

the risk analysis perspectives in each domain. The reasons for proposing a new definition are explained in Section 6.3.

### 6.2.1 Definition of risks related to safety

In general, safety related risk is defined or defined as follows [92]:

$$R_{safety} = \{S_{e_i}, P_{e_i}, X_{e_i}\}; i = 1, 2, \dots, N; \quad (6.1)$$

where

- $R_{safety}$  - safety related risk which is defined as a set of  $\{\}$ ;
- $S_e$  - scenario representation of the undesirable event under study ( $e$ ) by identifying safety causes of  $e$  and its related consequences;
- $P_e$  - likelihood of occurrence of  $S_e$ ;
- $X_e$  - severity of consequences of  $S_e$ ;
- $N$  - is the number of possible scenarios or undesirable events that can cause damages.

### 6.2.2 Definition of risks related to security

In the context of cyber-security, risk is defined in terms of likelihood and effects of a given threat exploiting a potential vulnerability ([150]; [79]):

$$R_{security} = \{(tv)_j, P_{(tv)_j}, X_{(tv)_j}\}; j = 1, 2, \dots, M; \quad (6.2)$$

where

- $R_{security}$  - security related risk which is defined as a set of  $\{\}$ ;
- $tv$  - scenario representation of a security breach: threat or attack ( $t$ ) exploits a vulnerability  $v$ ;
- $P_{tv}$  - likelihood of  $t$  exploits  $v$ ;
- $X_{tv}$  - severity of consequences if  $t$  exploits  $v$ ;
- $M$  - is the number of possible attacks.

### 6.2.3 Definition of risks related to safety and security

This section contributes a new global definition of industrial risk that covers safety and security. In the safety domain, risk is described as a set of undesirable events scenarios  $S_e$  with their related likelihoods and impacts (see Section 6.2.1). In the security domain, risk



is described as a set of scenarios that consist of threats exploiting vulnerabilities with the attached likelihoods and impacts (6.2.2). However, undesirable safety events can occur due to cyber threats after exploiting specific vulnerabilities. Thus, safety/security risk is defined in terms of a triplet as follows:

$$R = (S_{(tv,e)_i}, P(se, sa)_i, X_{(tv,e)_i}); i = 1, 2, \dots, N; \quad (6.3)$$

where

- $S_{(tv,e)}$  - Scenario description of the undesirable event ( $e$ ) that can result from safety incidents (safety causes) or/and security breaches (tv: threats exploit vulnerabilities - see the definition of security risk in Section 6.2.2);
- $P(se, sa)$  - likelihood of occurrence of  $S_{(tv,e)}$ , where  $se$  and  $sa$  are respectively the likelihoods related to security and safety;
- $X_{(tv,e)}$  - Severity of consequences of  $S_{(tv,e)}$ ;
- $N$  - is the number of possible scenarios or undesirable events that can cause damages.

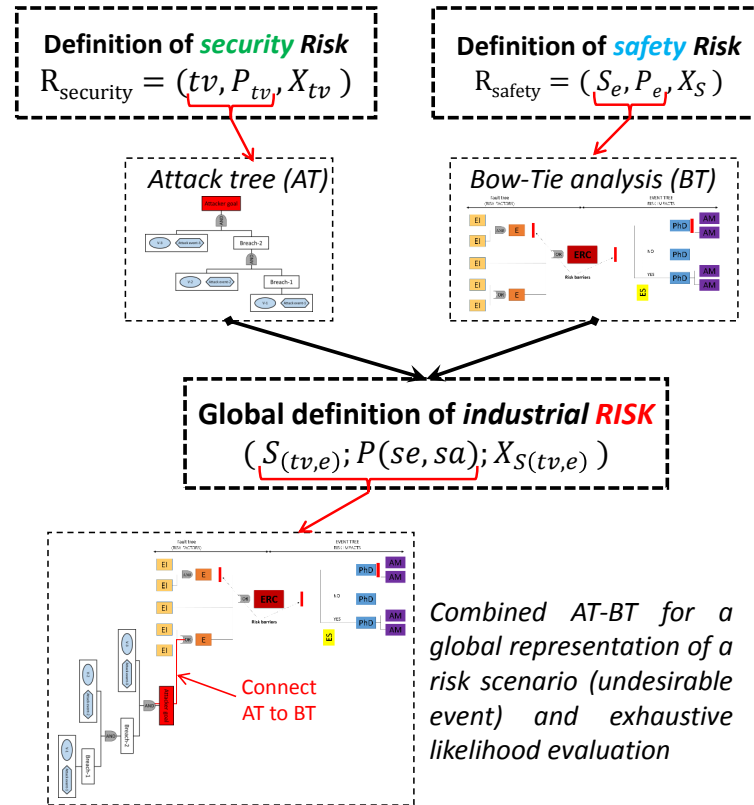
## 6.3 Methodology for combined safety/security risk analysis

### 6.3.1 Introduction behind the global idea

To represent (step 1 in the risk analysis process) and evaluate the likelihood (step 2 in the risk analysis process) of a risk scenario, in the domain of safety, a bowtie analysis is constructed and then we use this bow-tie to calculate the likelihood. In the domain of security, the chain of a security breach is represented by a graph called attack tree. The structures of the trees are close, here we propose a common representation by combining them as presented in Figure 6.1.

In both cases (safety and security), the risk is measured by a pair Likelihood/Severity. However, we realized that it is not possible to use a common qualitative scale for likelihood analysis. Indeed, if we consider an undesired event that can be generated from a component failure (safety) or a cyber attack (security), it is preferable to keep a double measure of likelihood (safety, security) rather than aggregating the two likelihood information into one. This gives a better idea on the importance of the two aspects and eventually, the use of the common model to detect an attack.

Let us take the example of a door lock which is controlled via the Internet. If the IT protection mechanism is moderate reliable, and the lock is difficult to break, then the risk is moderate (it needs a mechanical attack or a cyber attack). In comparison to a purely



**Figure 6.1** – Global definition of risk.

mechanical lock that is a little less resistant, this will present the same level of risk. The two systems appear to be similar whereas only the first one presents a residual cyber risk.

Thus, the proposed methodology to analyze safety/security risks is based on three main steps:

- ✓ identifying risk scenarios: we propose a methodology that combines BT with adjusted AT to identify the safety and security related causes and consequences of the undesirables events being studied. AT is extended to handle the limits of existing ATs as presented in Section 3.6.5. However, combining BT and AT analyses can be effectively used for an integrated safety/security assessment of critical systems. This methodology identifies and considers all safety incidents and security threats that can lead to the same undesirable event generating damages.
- ✓ likelihood evaluation: as BT and AT offer likelihood evaluation for safety and security risk scenarios, respectively, then the combined ATBT offers the same option for a safety/security risk scenario. But, as we said, sources of risk for safety and security are of different nature. Usually the likelihood of cause events related to safety are very low in comparison to the likelihood of security related cause events. For this reason, different likelihood scales, one for safety and another for security are defined to characterize the likelihood of input events. This differentiation helps

in identifying the sequences of events (minimal cut sets) that are purely related to safety, security or to both. The resulting output of different types of cut sets offers richer information for decision making and provides inputs for intrusion detection systems. In the rest of this chapter we are going to prove the importance of considering safety and security together and show that purely security risk sequences should be treated first.

- ✓ severity of consequences evaluation: this step aims to quantify the loss in terms of system assets, human life and environmental damage if the undesirable event has occurred. Here, the severity of an individual scenario is considered to be the same whatever the causes are related to safety or security. This part is not considered in this chapter. A detailed analysis of this step is presented in Chapter 5.

In the rest of this section, we will detail the proposed methodology for a combined safety/security industrial risk analysis. A risk scenario will be a combination of all expected security and safety events that can result in the undesirable event being studied. This modeling will be the first step in our methodology and it is conducted as presented in Section 4.2.2.

Next, we explain the second step that aims to evaluate a risk scenario in terms of likelihood as presented in Section 6.3.3. But, due to the difference in nature between safety and security related events, they will be characterized separately for likelihood analysis.

Figure 6.2 shows the framework to apply the proposed methodology. The proposed methodology will provide a deep, exhaustive analysis on safety/security for industrial risk scenarios in a given facility.

### **6.3.2 Step-1: representation of a risk scenario**

In this section, first we will introduce a new extended AT as depicted in Section 6.3.2.1, respectively. Then, we show how the proposed ATs can be integrated within BT (presented in 2.2.2) for richer and complete safety/security representation of a risk scenario (see Section 6.3.2.2).

#### **6.3.2.1 Security risk analysis using a new extended Attack Tree**

Traditional AT presents some limitations to be used for industrial risk analysis (Section 3.6.5). Showing just the steps that an attacker or a team of attackers follow to achieve a particular goal is not enough to understand the system's weaknesses. On the other hand, traditional AT does not present all the information needed to evaluate the likelihood of a successful attack on the target system. Thus, mapping information on the target system

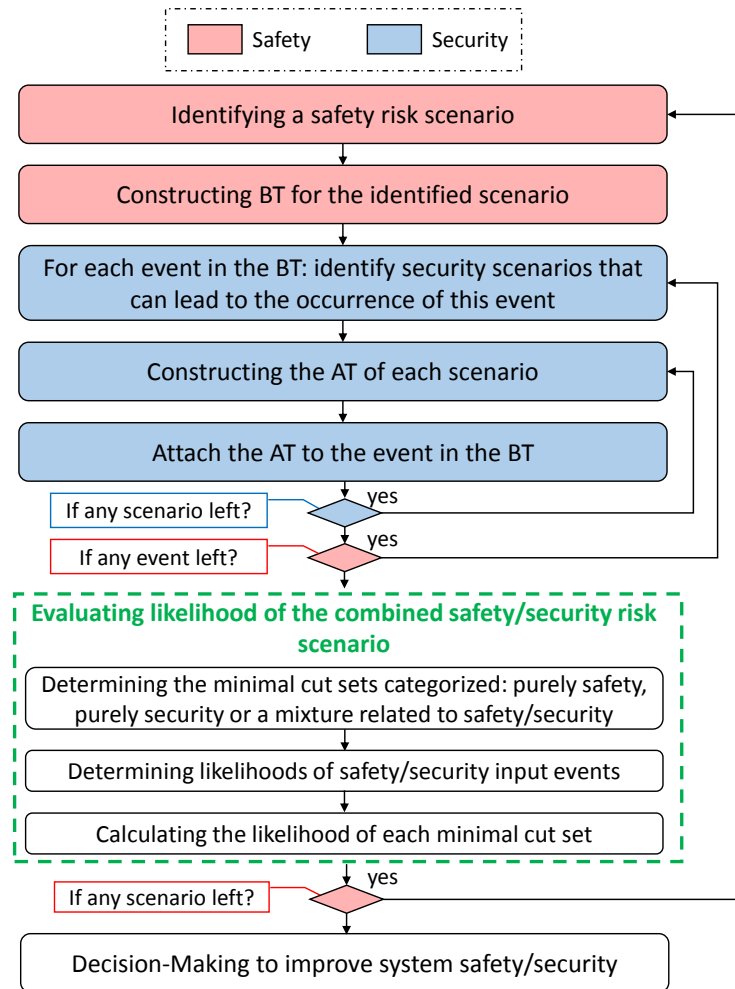


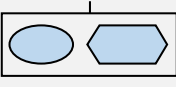
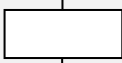
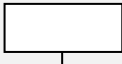


Figure 6.2 – Framework of the proposed approach for safety/security risk analysis.

such as vulnerabilities in addition to attack steps is essential for an effective security risk analysis using AT.

In this section, we will propose an extended version of attack tree with new modeling in order to characterize a security risk scenario. This extended version allows the consideration of significant information such as the target system vulnerabilities to suit the security risk analysis perspective. The AT's leaf nodes (security input events) are represented by a combination of attack events and vulnerabilities. This representation help the decision makers understanding the system's vulnerabilities (or weaknesses) and the different types of attacks that can be contacted in order to provide the right countermeasures.

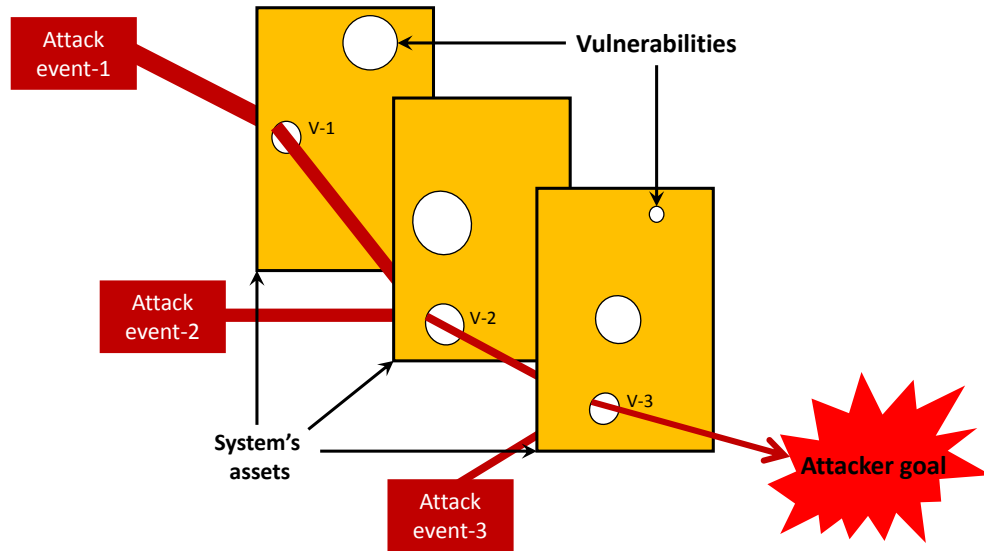
As in BT, the AND/OR gates are used to link the tree's events and define the relationship between them. Table 6.1 presents the term, shape and definition of each event used in the proposed AT.

	shape	Signification	Definition
Input events		Vulnerability	Any step describing a vulnerability required in order to realize the attack
		Attack	The attack process in order to exploit a system vulnerability
		Security basic event	Direct cause of a security breach resulting from exploiting a given vulnerability
		Intermediate	A security breach caused by the occurrence of input events
		Top event	The main goal of an attack generated from one or several security breaches

**Tableau 6.1** – Description of events used for representing an attack scenario.

The goal of this new AT is to model how attackers can exploit system vulnerabilities in order to cause damage. Figure 6.3 shows in a schematic way the reality behind how attackers target a system by exploiting its vulnerabilities. Here, attackers should run three different attacks and exploit three different vulnerabilities in order to achieve their goal. This attack can be modeled by the proposed AT as shown in Figure 6.4. Figure 6.4 shows the breach layers to attain the attack goal. This concept of layers would help propose the right countermeasure in the right place.

It should be noted that different attack events may be needed to exploit a specific vulnerability and vice versa. In these cases, the forms of the basic security events are presented in Figure 6.5. If we take the WannaCry ransomware attack as an example, the attack event is sending an unsolicited email that contains a link to exploit two different



**Figure 6.3** – How attackers exploit system vulnerabilities in order to cause damages.

vulnerabilities: (1) the computer runs Windows operating system that is not updated and (2) the unawareness of the user (if he/she clicks on the link). The security event will be as presented in Figure 6.6.

### 6.3.2.2 Combined ATBT analysis

This step aims to combine AT and BT analyses for a combined safety/security industrial risk analysis. The goal of this combination is to provide a complete representation of risk scenarios by plotting on the same scheme safety and security events that can lead to the same undesirable events. Additionally, integrating ATs within BT analysis can help in understanding how attackers can exploit systems' weaknesses in order to cause damages besides non-deliberate incidents.

This step is conducted as follows:

1. construct BT for the chosen undesirable event being analyzed;
2. for each safety event in BT, identify if there are security incidents that can lead to the occurrence of this event. If yes, construct the AT and attach its goal to the corresponding event (see Figure 6.7). This means that this event can occur due to accidental (safety) or deliberate (security) incidents.

Finally, a cyber-security BT (ATBT) is obtained for the undesirable event being studied.

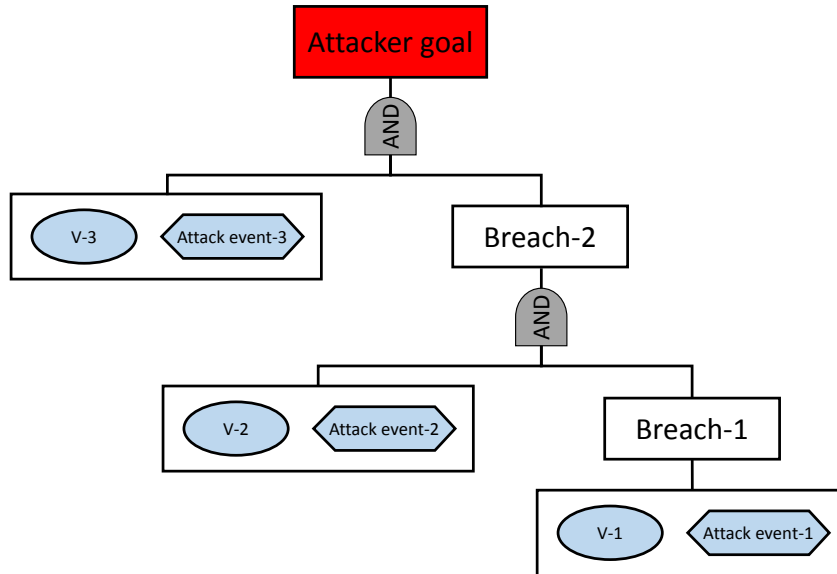


Figure 6.4 – Example of the structure of the proposed attack tree.

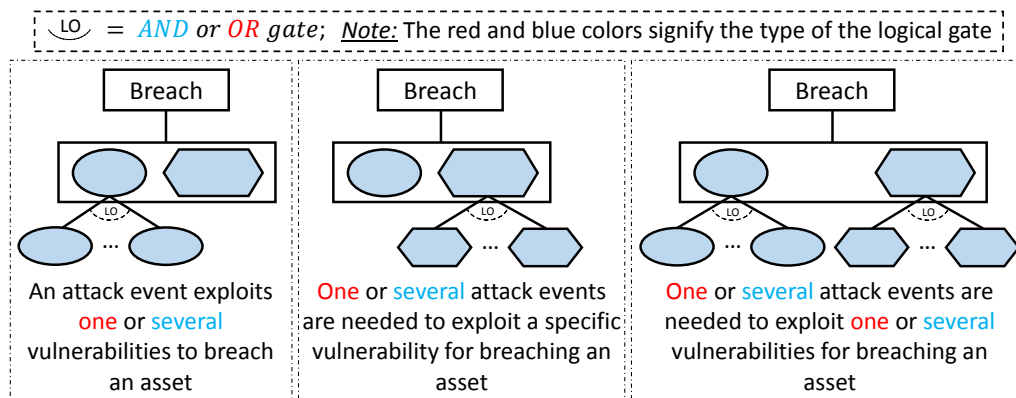
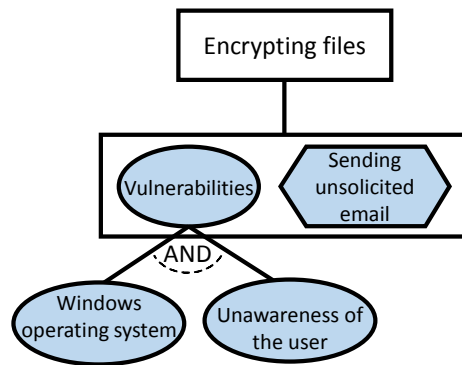


Figure 6.5 – The different form of a security basic event.

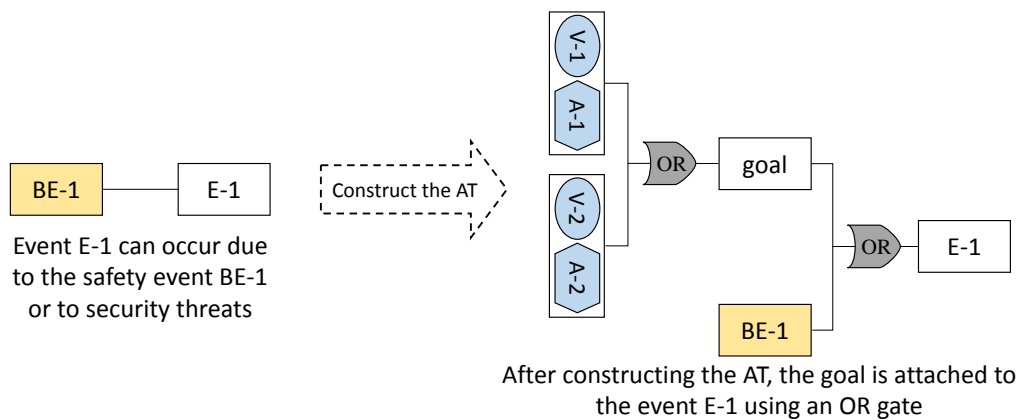
### 6.3.3 Step-2: likelihood evaluation

This section proposes an approach for conducting a qualitative likelihood analysis of a safety/security risk scenario. This likelihood analysis methodology is made up of three main steps: (i) determining the minimal cut sets to understand the structural weaknesses of a system, (ii) characterizing likelihoods of input events using a two-levels representation and (iii) quantify the likelihood of each MC to prioritize the system’s weaknesses (see Sections 6.3.3.1, 6.3.3.2 and 6.3.3.3, respectively).

We should repeat that we are required to characterize likelihood of safety and security events separately because they are intrinsically different and the control in terms of safety or security barriers should be managed independently of the two safety and security aspects.



**Figure 6.6** – Modeling WannaCry ransomware attack using the proposed AT.



**Figure 6.7** – Example of how we attach an AT to its corresponding event in BT.

### 6.3.3.1 Determining minimal cut sets

Finding out the MCs represents the first step of likelihood evaluation in our approach. An MC is the smallest combination of input events which causes the occurrence of the undesirable event. MCs present the different ways in which component failures or events alone or in combination with others make the occurrence of the top event (minimal cut sets with one or several components or events). Here, the MCs are obtained using rules of boolean algebra [163]. Each MC set is a combinations of AND gates containing a set of basic inputs necessary and sufficient to cause the top event (see [72], appendix D for more details).

We separate between three types of minimal cuts:

- purely related to security: all events of the MC are due to deliberate attacks;
- purely related to safety: the MC does not contain any security related event;
- related to a mixture of both security and safety: accidental and deliberate causes exist in the MC.

The importance of this differentiation between types of MCs is to discover the system's



weaknesses where a pure security MC represents a weak point due to the high likelihood of occurrence of security causes. This reasoning will be detailed and demonstrated in the rest of this chapter.

### 6.3.3.2 Characterizing likelihoods of occurrence of input events

In safety, the likelihood of occurrence is the probability (expected frequency) or possibility of something happening. But when we talk about security, the likelihood of occurrence is the probability that a given threat is capable of exploiting a vulnerability (or set of vulnerabilities).

Likelihood analysis can be qualitative or quantitative depending on the type of available data. This data is either quantitative derived from historical incident or qualitative based on experts' elicitations. Because of the difficulties in estimating quantitative likelihood of occurrence of an attack or an accidental cause, a qualitative scale is used. The advantage of the qualitative methodology is its simplicity of applying and understanding by the relevant personnel.

As we presented in the beginning of this section, there are different concepts to define likelihood related to safety and security. Due to the deviation in the likelihood translation, high likelihood in safety is different than high likelihood in security regarding the number of observed safety and security incidents (we see cyber attacks on critical facilities every day, while safety incidents are rare). Two different scales  $L_s$  : *security* and  $L_f$  : *safety* of respectively five and six levels are proposed. The first level of each scale represents an undefined value (likelihood equals zero) in order to specify if an event is purely related to safety or security. Thus, each event is characterized by couples  $(L_s, L_f)$ .

Based on this likelihood representation in terms of couples, we can differ between three different types of events presented as follows:

- events that are purely related to safety with likelihood  $(N/A, L_f)$  for each event;
- events that are purely related to cyber-security with likelihood  $(L_s, N/A)$  for each event. If the event is a security cause (basic event in terms of two parts),  $L_s$  will depend on the vulnerability level and the technical difficulties of conducting the attack as we will detail in Section 6.3.3.2.2;
- events (intermediate events) related to both safety and security with likelihood  $(L_s, L_f)$  for each event.

#### 6.3.3.2.1 Characterizing likelihood for safety risk events

Likelihood characterization here aims to determine the likelihood of occurrences of input events (BEs and Es in BT) and the likelihood of failures of risk barriers according

Qualitative scale	Safety Level	Designation	Quantitative meaning
<b>Likelihood</b>	N/A	<b>Not Applicable:</b> event is purely related to security, not safety	
	E	<b>Very unlikely:</b> event that is practically impossible, very low chance of happening	0
	D	<b>Unlikely:</b> Low chance of occurrence even if we consider several systems of the same type, but has to be considered as a possibility	$10^{-5}$
	C	<b>Moderate:</b> may occur during total operational life if considering several systems of the same type	$10^{-4}$
	B	<b>Likely event:</b> may occur during total operational life of a system	$10^{-3}$
	A	<b>Very likely event:</b> can frequently occur (several times) during operational life	$10^{-2}$

**Tableau 6.2** – Qualitative scale to characterize the frequency of input safety events.

to a specific scale. The same scale presented in Section 2.2.2 for safety analysis is used in this study. Table 6.2 presents the safety scale with the added first level (N/A).

#### 6.3.3.2.2 Characterizing likelihood for security risk events

In the context of a security risk analysis, the likelihood of occurrence depends on the capability that a given threat (or set of threats) exploiting a potential vulnerability (or set of vulnerabilities). Thus, the likelihood is a function of the difficulty of performing a needed attack to exploit a vulnerability, and the level of vulnerability depending on the existing counter measures. In this article, two different criteria are considered to determine the likelihood of a security initial event presented as follow:

- vulnerability level: given to a vulnerability in the ATBT to represent how easy or hard exploiting this vulnerability depending on the existing countermeasures. Figure 6.8(a) shows the three different levels proposed to evaluate this criterion. Level 1 (E) means that the vulnerability is easy to be exploited (for example, a password that should be a number of four digit represents an easy vulnerability). Vulnerabilities of level 2 (M) or 3 (H) are harder to be exploited due to the presence of security measures. If we take the same example, a password that should be a number of eight digit would be of level 2. While an eight digit password that should contains lower and upper case letters in addition to numbers would be of level 3;
- technical difficulty of conducting an attack: given to an attack event to show the needed level of expertise or difficulty to conduct the attack. Figure 6.8(b) presents the levels of difficulty of an attack inspired from [28]. Four levels  $\{T, M, D, VD\}$  are used to describe the difficulty of executing an attack. (T) is the easiest to conduct where normal computer skills are required (for example, running a Denial

Qualitative scale	Security Level	Designation
<b>Likelihood</b>	N/A	<b>Not Applicable:</b> Event is purely related to safety, not security
	1	<b>Low:</b> High unlikely to occur, attack is hard to perform, existence of effective security measures
	2	<b>Moderate:</b> Possibility to occur, but existed security measures reduce the likelihood of occurrence
	3	<b>High:</b> Likely to occur, limited countermeasures are presented
	4	<b>Strong:</b> Is almost certain to occur, system is an easy target

**Tableau 6.3** – Qualitative scale to characterize the likelihood of input security events.

Of Service Attack). (M) demands some programming and security skills (running an SQL injection). (D) needs a hacking expert (man in the middle attack). (VD) is the hardest where a team of competent hackers are needed to conduct the attack (implementing a sophisticated warm).

These two criteria should then be combined in order to provide a likelihood for the security initial (or basic) events. The difficulty of the attack is combined with the vulnerability levels as presented in Figure 6.8(c). Four different security likelihood levels in addition to the N/A level are proposed to represent the combination. The definition of each security likelihood level is presented in Table 6.3. From Table 6.3, we can note that likelihood levels of security events are different from those of safety events (Table 6.2).

### 6.3.3.3 Calculating the likelihoods of MCs

This step aims to prioritize the system weaknesses by calculating the likelihood of each MC in order to help decision makers propose the right countermeasure where MCs with highest likelihood should be treated first.

Calculating the likelihood of an MC only needs the AND gate to be solved. AND gate signifies that the output event occurs if all its input events have occurred. Since qualitative scales are used for safety and security likelihood characterization, the min rule is used to solve the AND gate. Suppose an AND gate with  $n$  input events  $EV_i, i = 1, \dots, n$ , the output likelihood is calculated as presented in Eq 6.4 [86].

$$\begin{aligned}
 L(AND_{out}) &= \min[L(EV_i)] = (\min[L_{security}(EV_i)], \min[L_{safety}(EV_i)]) \\
 &= (\min[L_{security}(EV_1), \dots, L_{security}(EV_n)], \\
 &\quad \min[L_{safety}(EV_1), \dots, L_{safety}(EV_n)])
 \end{aligned}
 \tag{6.4}$$

Likelihood levels		Likelihood of safety events					
		E	D	C	B	A	N/A
Likelihood of security events	N/A	VL	L	M	H	VH	NS
	4	VL	L	M	H	VH	VH
	3	VL	L	M	H	H	H
	2	VL	L	M	M	M	H
	1	VL	L	L	L	L	M

NS: Not Significant   
 VL: Very Low   
 L: Low   
 M: Moderate   
 H: High   
 VH: Very high

Tableau 6.4 – Analysis scale - Overall likelihood.

where  $L(EV_1), \dots, L(EV_n)$  are the likelihoods of occurrence attached to  $EV_1, \dots, EV_n$ , respectively.

Finally, for each MC, the two determined likelihoods for safety and security should be taken together to provide one meaningful likelihood to be used for prioritizing MCs and for risk evaluation using the likelihood-consequence risk matrix (which is not discussed in this chapter). Table 6.4 presents the overall scale regarding the proposed safety and security scales. This overall scale defines five different qualitative expressions from low (L) to very high (VH).

It should be noted that this overall-likelihood can not replace the double part likelihoods ( $L_{security}, L_{safety}$ ) which is important for decision-making and in choosing the right countermeasure, because decision makers should know if the high likelihood is related to safety, security or to both.

Figure 6.9 presents an example on how to calculate the likelihood of an MC. The MC in Figure 6.9 presents four basic events, two are related to safety ( $BE - 1$  and  $BE - 2$ ) and the other two are security related ( $SBE - 1$  and  $SBE - 2$ ). Based on the proposed approach, experts are asked to characterize the likelihood of safety basic events, and (i) difficulty of attacks and (ii) exploitability of vulnerabilities for security basic events. From (i) and (ii), the likelihood security basic events are determined based on Figure 6.8 (For example,  $SBE - 1$  is of level 4 since the vulnerability level is E and the needed attacker skills are T). The dashed rectangle beside each event in the figure presents its likelihood. These likelihood are then propagated through the MC. The likelihood of events SE-1 and E-1 are calculated based on Eq 4.5.  $L(E - 1) = \min(L(BE-1), L(BE-2)) = (\min[L_{security}(BE-1), L_{security}(BE-2)], \min[L_{safety}(BE-1), L_{safety}(BE-2)]) = (\min[N/A, N/A], \min[A, C]) = (N/A, C)$ , where  $L(E - 2) = \min(L(SBE-1), L(SBE-2)) = (\min[4, 3], \min[N/A, N/A]) = (3, N/A)$ . The likelihood

of the top event is equal to  $\min(L(E-1), L(E-2)) = (\min[N/a; 3], \min[C; N/A]) = (3, C)$  which is of level Moderate (M) based on Table 6.4.

This approach will be illustrated in the next section and applied to an overheating scenario in a chemical reactor.

## 6.4 Case study

### 6.4.1 Description

This case study illustrates the implementation of the proposed approach, which can be applied in any industrial context. The case study concerns an industrial site of a propylene oxide polymerisation reactor [2]. The same reactor presented in Chapter 5 is used for this case study. The structure of the reactor is more detailed in this chapter. The reactor runs a high exothermic chemical reaction at high pressure. It is located in a manufacturing site located south of a small town. Risks associated with the operation of the reactor are of high consequences.

In a systematic representation of the reactor, a production system, a cooling system and a power supply are interacting in order to perform the operation under normal conditions (regulated temperature and pressure). Components of these systems (valves, pumps, etc.) are controlled by Siemens PLCs and supervised by a SCADA system. The information collected by the SCADA system is accessible by all the site managers from their offices using wireless remote control. The manager of the utility can control the facility using its tablet or smart phone via Internet. Controlling the process via Internet would allow the manager to handle the situation from where he/she is before it is too late, rather than waking up at midnight racing to the plant to handle the situation. Figure 6.10 shows the architecture of the system under study. The architecture of the system is taken from the “Risk analysis: socio-technical and industrial systems” book [63].

The reactor is used in batch mode to run a chemical reaction in order to produce a product C from two reactives A and B. The temperature of the reaction is regulated with industrial water. At the end of the reaction, after the mixture A, B is completely transformed. The output C is transferred toward another unit in the facility by opening the valve XV33021. This process is controlled by PLC1.

The cooling system E33040 receives cold industrial water as input which is used to cool down the content of reactor R33030 using a double jacket. The temperature of the cooling system and the water flow rate are measured by the sensor TI33061 and TI33062, respectively. The data collected by these two sensors is sent to PLC2 which regulates the water flow rate by controlling P1, P2, CV33063 and XY33027. Under

normal conditions, the pressure in the reactor is less than six bars when the temperature is controlled under 120 °C. An automated safety valve PSV33009 opens in the case of over-pressure to limit the pressure to 10 bars. This safety valve is connected to host computers. After PSV33009 opened, the exhausted gases are cleaned by scrubber.

## 6.4.2 Application

In this case study, the most likely undesirable scenario with the highest consequences due to overheating/overpressure is considered for risk analysis. This scenario can be generated after the occurrence of deliberate attacks or accidental errors. Overheating occurs if the temperature and pressure exceed the threshold.

The two first steps for risk analysis (risk identification and likelihood evaluation) using the proposed methodology are applied on the overheating scenario as depicted in Sections 6.4.2.1 and 6.4.2.2, respectively.

### 6.4.2.1 Step-1: Constructing ATBT for safety/security analysis

This step contains two sub-steps as presented in the proposed methodology:

1. constructing the BT for safety analysis: Figure 6.11 presents the BT for the undesirable event under study. The undesirable main event is an overheating and increase in pressure inside the reactor. This event occurs after an abnormal increasing of the temperature and pressure which is due to:
  - (a) an error in the cooling system: this event can be generated from accidental failures if the valve XYSV33027 breaks down (BE-1), pumps P1 or P2 breakdown (BE-2 or BE-3), the valve XYSV33063 breaks down (BE-4), or failure in the power supply (BE-5). It can also be initiated by deliberate attacks on the control system (as detailed in the next paragraph: constructing ATs);
  - (b) over loading: an excessive loading of the reactor due to a human error (BE-6);
  - (c) agitation system breakdown: if the power supply or the motor of the system breaks down (BE-7 or BE-8).

However, this rise in pressure is limited by an automated safety valve. If this does not accomplish its purpose due to mechanical failure (BE-9) or cyber-attack (see next paragraph), it would result in the explosion of the reactor. Thus, nine safety related basic events are investigated as causes of the overheating in the reactor.

2. constructing ATs for security analysis: two events in the BT of Figure 6.11 can occur due to security breaches. The first event is the failure of the automated safety valve due to an attack on the hardware (SBE-6 in Figure 6.11: exploiting the

control surveillance vulnerability by running a Doorknob rattling attack). To run a doorknob attack, the attacker attempts a few common username and password combinations on several computers resulting in very few failed login attempts. This attack can go undetected unless the data related to login failures from all the hosts are collected and aggregated to check for doorknob-rattling from any remote destination ([166]; [18]). The second is by sabotaging the cooling system after gaining unauthorized access to the SCADA system. SCADA system can be exploited by attacking the computer software or the communication network as shown in Figure 6.12 and detailed below:

- attacking the communication network: this can be achieved by sending a malicious email to steal access information from an authorized target (employees inside the facility) to exploit the no existence of email surveillance (exploit confidentiality), see SBE-5 in Figure 6.12. Or, exploiting the weakness of the encryption algorithm (integrity of information) used for communication between the SCADA and the control level. Different attacks can be performed to exploit integrity: message spoofing, replay attack or man in the middle attack (SBE-4). The communication network can also be hacked by running a Denial Of Service attack where the system is vulnerable and reached from a big sized network (SBE-3).
- attacking the computer software: variety of applications software are implemented to complete the functionality of the control system. Furthermore, there are large databases that save confidential information and data about the process. SCADA applications software are susceptible to be hacked by sophisticated threats. Most of these applications are written in C programming language which make them vulnerable to the Buffer Overflow attack (SBE-2). This attack aims to insert lines of assembly codes such that can result in corrupting the memory. A successful Buffer Overflow attack can corrupt data, crash the program, or cause the execution of malicious codes. SQL injection also represents a threat to the computer software (SBE-1). SQL injection is one of the top Web attacks that affects the security of SCADA systems. It occurs when an adversary is able to manipulate a malicious SQL query into a Web application that fails to properly sanitize the query.

#### 6.4.2.2 Step-2: Likelihood evaluation

- (a) determining minimal cut sets: The ATBT shown in Figure 6.11 yields to 21 MCs. All minimal combinations of basic events that result in the occurrence of the main event are identified. Figure 6.13 shows an example of the MC number

19. The MC in Figure 6.13 is related to mixture safety/security because safety and security basic events (respectively BE-9 and SBE-4 in the Figure) should occur together to cause the occurrence of the undesirable event. These MCs are divided into 7 that are purely related to security, 7 that are purely related to safety and 7 that are related to mixture safety/security.
- (b) characterizing likelihood of occurrence of input events: Experts in the field are asked to characterize likelihoods of safety and security basic events. The characterized likelihoods in terms of couple  $(L_{security}, L_{safety})$  are drawn beside the basic events in the ATBT (see figures).
- (c) calculating likelihood of MCs: safety/security likelihood of each MC is calculated using Eq 6.4 as shown in Table 6.5. Then the overall likelihood of each MC is determined based on Table 6.4. As an example, Figure 6.13 presents calculating the likelihood of MC number 19. MC number 19 contains two basic events BE-9 and SBE-4 with likelihoods equal to  $(N/A, D)$  and  $(4, N/A)$  (derived based on Figure 6.8(c)), respectively. After propagating these likelihoods, the likelihood of the explosion event is equal to  $(4, D)$  which is of level  $L$ .

### 6.4.3 Discussion and improvement

As shown in Table 6.5, the MCs ranked high (H) and (VH) are purely due to cybersecurity. This reveals the importance of considering security risks during safety risk analysis. However, the presence of a safety event in an MC will lead to less likelihood of occurrence. We can clearly see that between MC-5 and MC-19 where their attached likelihoods are equal to VH and L respectively, MC-19 is of less likelihood because it contains the accidental event BE-9.

To show the importance of analyzing safety and security together, a burst disk is added which represents a mechanical component (no security breaches are related) as improvement for the process. The re-determination of MCs shows that there is no MC that is related to pure security. Table 6.6 shows the re-determined MCs with their re-estimated likelihoods. The introduced improvement diminishes the likelihoods into the lowest level. Thus, the presence of a mechanical failure (safety event) in a cut set will insure the prevention of malicious attacks and vice versa. For these reason, safety and security being treated together will lead to a better risk analysis and effective decision making.



	MCS	Likelihood	Level		MCS	Likelihood	Level
1	SBE-1; SBE-6	(2, N/A)	M	12	BE-6, BE-9	(N/A, D)	L
2	SBE-2; SBE-6	(2, N/A)	M	13	BE-7, BE-9	(N/A, D)	L
3	SBE-3; SBE-6	(4, N/A)	VH	14	BE-8, BE-9	(N/A, D)	L
4	SBE-4(V-4, A-4.1) ; SBE-6	(3, N/A)	H	15	SBE-1; BE-9	(2, D)	L
5	SBE-4(V-4, A-4.2) ; SBE-6	(4, N/A)	VH	16	SBE-2; BE-9	(2, D)	L
6	SBE-4(V-4, A-4.3) ; SBE-6	(3, N/A)	H	17	SBE-3; BE-9	(4, D)	L
7	SBE-5; SBE-6	(4, N/A)	VH	18	SBE-4(V-4, A-4.1) ; BE-9	(3, D)	L
8	BE-1, BE-9	(N/A, D)	L	19	SBE-4(V-4, A-4.2) ; BE-9	(4, D)	L
9	BE-2, BE-3, BE-9	(N/A, D)	L	20	SBE-4(V-4, A-4.3) ; BE-9	(3, D)	L
10	BE-4, BE-9	(N/A, D)	L	21	SBE-5; BE-9	(4, D)	L
11	BE-5, BE-9	(N/A, D)	L				

Purely security related MC
  Mix related MC
  Purely safety related MC

**Tableau 6.5** – The identified MCs for the scenario under study.

## 6.5 Modeling Stuxnet using the ATBT

Stuxnet is one of the most sophisticated and advanced computer worms that can affect ICS. For this purpose and in order to present the utility of the proposed AT in the security domain and not just to combine safety/security, we modeled the Stuxnet worm to examine the impact of computer worms on industrial control systems. In the rest of this section, we detail how Stuxnet operates to sabotage the control system.

The different operations (attacks) and vulnerabilities Stuxnet exploits are modeled in Figures 6.14, 6.15, 6.16.

Stuxnet is one of the most sophisticated worm that was designed to target a specific Siemens PLC ([51]; [114]). The main goal of Stuxnet is to gain unauthorized access to this PLC in order to attack and sabotage the industrial system [94]. To do so, Stuxnet shall install itself after being injected into the facility, spread via the network to find the PLC and lastly run the attack as respectively presented in Figures 6.16, 6.15 and 6.14. From Figure 6.16, Stuxnet can infect a computer inside using different paths. Injecting an infected removable drives and open it. The virus

	MCS	Likelihood	Level		MCS	Likelihood	Level
1	SBE-1; SBE-6; BE-10	(2, E)	VL	12	BE-6, BE-9; BE-10	(N/A, E)	VL
2	SBE-2; SBE-6; BE-10	(2, E)	VL	13	BE-7, BE-9; BE-10	(N/A, E)	VL
3	SBE-3; SBE-6; BE-10	(4, E)	VL	14	BE-8, BE-9; BE-10	(N/A, E)	VL
4	SBE-4(V-4, A-4.1) ; SBE-6; BE-10	(3, E)	VL	15	SBE-1; BE-9; BE-10	(2, E)	VL
5	SBE-4(V-4, A-4.2) ; SBE-6; BE-10	(4, E)	VL	16	SBE-2; BE-9; BE-10	(2, E)	VL
6	SBE-4(V-4, A-4.3) ; SBE-6; BE-10	(3, E)	VL	17	SBE-3; BE-9; BE-10	(4, E)	VL
7	SBE-5; SBE-6; BE-10	(4, E)	VL	18	SBE-4(V-4, A-4.1) ; BE-9; BE-10	(3, E)	VL
8	BE-1, BE-9; BE-10	(N/A, E)	VL	19	SBE-4(V-4, A-4.2) ; BE-9; BE-10	(4, E)	VL
9	BE-2, BE-3, BE-9; BE-10	(N/A, E)	VL	20	SBE-4(V-4, A-4.3) ; BE-9; BE-10	(3, E)	VL
10	BE-4, BE-9; BE-10	(N/A, E)	VL	21	SBE-5; BE-9; BE-10	(4, E)	VL
11	BE-5, BE-9; BE-10	(N/A, E)	VL				

Purely security related MC
  Mix related MC
  Purely safety related MC

**Tableau 6.6** – The re-identified MCs after the added improvement.

will spread to the computer by exploiting the Auto-run or the LNK vulnerabilities. Or via internet by sending a malicious email as modeled by SBE-3. After infecting a computer inside the facility, Stuxnet installs itself by stealing a digital certificate (exploiting the Realtek vulnerability) and loading a dropper (.ddl) file (exploiting the Windows vulnerabilities) as represented by SBE-4 and SBE-5, respectively.

The second step of Stuxnet is to spread inside the facility searching for its target (Siemens PLC). Stuxnet can spread using different ways as shown in Figure 6.15 and presented below ([118]):

- spread via WinCC vulnerability: Stuxnet searches for computers running the SCADA interface Simatic WinCC and connects into WinCC using a password hard-coded (SBE-6). Then attacks using SQL injection (SBE-7). If these two have been done successfully, Stuxnet uploads and copies itself on the WinCC computer.
- spread via Network shares (SBE-8): Stuxnet can exploit the existing of shared folders to spread throughout the local network. It places a Trojan.dropper file to install the virus on the target computers that share the same folders.
- spread via the MS10-061 print spooler 0-day vulnerability (SBE-9): Stuxnet uploads copies of itself on remote computers by exploiting this vulnerability. By executing these copies, Stuxnet infects the remote machines.
- spread via the MS08-067 SMB vulnerability (SBE-10): if a remote computer has this vulnerability, Stuxnet can send a malformed path over SMB (a protocol for sharing files and other resources between computers) to execute arbitrary code on the remote machine, thereby propagating itself to it.

The last step for Stuxnet represents the attack phase to compromise and sabotage the SCADA system (see Figure 6.14). After spreading and finding its target, Stuxnet checks the connection to PLC as well as other specific configuration (PLC model, Profibus configuration, speed regulators number). Stuxnet sends this information to its senders in order to get updated (SBE-11). Once it is updated, Stuxnet looks for WinCC/Step7 software on the control PC used to configure the PLC in order to proceed infecting and modifying PLC function blocks. If found, it installs a rootkit: it loads a library file (s7otbxdx.dll) used for the communication between the control PC and the PLC, renames it (s7otbxsx.dll) and inserts malicious codes into the new file (SBE-12). Beside this step, Stuxnet conceals the attack activities (SBE-13): it collects data for a period of 13 to 90 days before conducting the attack and sending modified control data. Thus, the malware operates without being detected. As result, Stuxnet sends wrong control data and displays to the operator that the system is under normal conditions.

The MCs of Stuxnet scenario with their likelihoods are determined as presented in Table 6.7. Table 6.7 presents 20 different paths Stuxnet can follow to sabotage the system. The given input data for SBEs is subjective and can be changed depending on the structure of the system under study. The results reflect the potential attack sequences and gave a likelihood of occurrence of each one. These results would help preventing or reducing the likelihoods of high likelihood attack sequences by providing the right security measures in the right place.

	MCS	Likelihood	Level		MCS	Likelihood	Level
1	SBE-1(V1, A-1.1); SBE-2(V-2.1, A-2); SBE-4; SBE-5; SBE-6; SBE-7; SBE-11; SBE-12; SBE-13	(1, N/A)	L	11	SBE-1(V1, A-1.1); SBE-2(V-2.2, A-2); SBE-4; SBE-5; SBE-9; SBE-11; SBE-12; SBE-13	(3, N/A)	H
2	SBE-1(V1, A-1.2); SBE-2(V-2.1, A-2); SBE-4; SBE-5; SBE-6; SBE-7; SBE-11; SBE-12; SBE-13	(1, N/A)	L	12	SBE-1(V1, A-1.2); SBE-2(V-2.2, A-2); SBE-4; SBE-5; SBE-9; SBE-11; SBE-12; SBE-13	(3, N/A)	H
3	SBE-1(V1, A-1.1); SBE-2(V-2.2, A-2); SBE-4; SBE-5; SBE-6; SBE-7; SBE-11; SBE-12; SBE-13	(1, N/A)	L	13	SBE-1(V1, A-1.1); SBE-2(V-2.1, A-2); SBE-4; SBE-5; SBE-10; SBE-11; SBE-12; SBE-13	(2, N/A)	M
4	SBE-1(V1, A-1.2); SBE-2(V-2.2, A-2); SBE-4; SBE-5; SBE-6; SBE-7; SBE-11; SBE-12; SBE-13	(1, N/A)	L	14	SBE-1(V1, A-1.2); SBE-2(V-2.1, A-2); SBE-4; SBE-5; SBE-10; SBE-11; SBE-12; SBE-13	(2, N/A)	M
5	SBE-1(V1, A-1.1); SBE-2(V-2.1, A-2); SBE-4; SBE-5; SBE-8; SBE-11; SBE-12; SBE-13	(3, N/A)	H	15	SBE-1(V1, A-1.1); SBE-2(V-2.2, A-2); SBE-4; SBE-5; SBE-10; SBE-11; SBE-12; SBE-13	(2, N/A)	M
6	SBE-1(V1, A-1.2); SBE-2(V-2.1, A-2); SBE-4; SBE-5; SBE-8; SBE-11; SBE-12; SBE-13	(3, N/A)	H	16	SBE-1(V1, A-1.2); SBE-2(V-2.2, A-2); SBE-4; SBE-5; SBE-10; SBE-11; SBE-12; SBE-13	(2, N/A)	M
7	SBE-1(V1, A-1.1); SBE-2(V-2.2, A-2); SBE-4; SBE-5; SBE-8; SBE-11; SBE-12; SBE-13	(3, N/A)	H	17	SBE-3; SBE-4; SBE-5; SBE-6; SBE-7; SBE-11; SBE-12; SBE-13	(1, N/A)	L
8	SBE-1(V1, A-1.2); SBE-2(V-2.2, A-2); SBE-4; SBE-5; SBE-8; SBE-11; SBE-12; SBE-13	(3, N/A)	H	18	SBE-3; SBE-4; SBE-5; SBE-8; SBE-11; SBE-12; SBE-13	(2, N/A)	M
9	SBE-1(V1, A-1.1); SBE-2(V-2.1, A-2); SBE-4; SBE-5; SBE-9; SBE-11; SBE-12; SBE-13	(3, N/A)	H	19	SBE-3; SBE-4; SBE-5; SBE-9; SBE-11; SBE-12; SBE-13	(2, N/A)	M
10	SBE-1(V1, A-1.2); SBE-2(V-2.1, A-2); SBE-4; SBE-5; SBE-9; SBE-11; SBE-12; SBE-13	(3, N/A)	H	20	SBE-3; SBE-4; SBE-5; SBE-10; SBE-11; SBE-12; SBE-13	(2, N/A)	M

Tableau 6.7 – The identified MCs for the Stuxnet attack scenario.

## 6.6 Limitations and future work

The methodology in this chapter focuses on introducing cyber-security related threats within industrial risk analysis. Beside the advantages this methodology presents, it hides some limits:

- the step of characterizing likelihood for security related events (Section 6.3.3.2.2) could be improved. Many other criteria could be taken into account to evaluate the likelihood of security related events: connectivity of systems, control of internal and external stakeholders, technology and communication protocols used, organization set up to monitor and patch vulnerabilities, etc.;

- the approach presented for likelihood analysis is qualitative. As we said, the advantage of this qualitative approach is its simplicity of applying and understanding by the relevant personnel. But, this qualitative approach is subjective and heavily dependent on the experience of experts who are performing the analysis and it can be influenced by personal idiosyncrasies. Consequently, it can lead to inaccurate and imprecise risk predictions. Moreover, statistical data, if available, is lost by using a qualitative likelihood analysis. Therefore, moving towards a semi-quantitative approach that uses statistical data if available or qualitative data if not is a focus of interest;
- in general, each dangerous phenomenon is partly analyzed in terms of likelihood/severity. Nevertheless, in case of security, it can be imagined that several scenarios can occur at the same time. An attacker could try to cause several phenomena at the same time to maximize the effects. For example, if there are several tanks of dangerous materials, in traditional risk analysis, the fire of each tank is assessed separately. But, if there are several simultaneous fires, the severity would be greater and the the intervention to reduce accidents would be more difficult.

These limits will enhance future research about introducing cyber-security with safety risk analysis.

## 6.7 Conclusion

The use of technology in critical facilities exposes systems' safety to security related threats. These threats are due the use of Internet, standardized protocols and electronic components for connectivity and remote controls.

Nowadays, most of the existing approaches for industrial risk analysis ignore cyber-security. On the other hand, cyber-security analysis ignore the benefits of non-digital systems in decreasing the likelihood of security scenarios. In light of security threats, there is an urgent need for complete and effective safety risk analysis. That is why this chapter proposes an approach that integrates ATs with BT analysis for a combined safety and security industrial risk analysis. Bowtie analysis is used for analyzing safety accidents. A new concept of Attack Tree is introduced to consider potential malicious attacks that can affect the system's safety. The steps of combining AT within BT is presented and the process for likelihood evaluation is explained.

There is complexity in quantifying likelihoods of attacks and a lack of consistency in the likelihood of occurrence between deliberate and accidental causes of risk. For these reasons, two different qualitative likelihood scales one for safety and another for security are proposed for representing the likelihood of basic events related to safety and security.

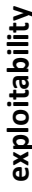
The different likelihood scales lead to three different types of events sequences (MCs). A qualitative mathematical rule is used to calculate the likelihoods of MCs.

The outputs of the approach show important results in terms of representation of risk scenarios as well as in likelihood quantification. MCs due to purely safety, security or both can be separately extracted. This separation between MCs helps understand the origins of risk and provide the right control measures.

The application of the proposed approach on an undesirable scenario in a chemical reactor shows that the highly likelihood MCs are purely related to security. The added improvement diminishes the unacceptable likelihood to an acceptable level. The results show that the moves from purely security MCs to mix safety/security MCs is the safest risk treatment.

The applicability of the proposed AT in the security domain is also approved by modeling the Stuxnet virus. A graphical representation of the different steps and event combinations of Stuxnet is provided. The minimal attack sequences of Stuxnet are extracted and their likelihoods are calculated.

At the end of this chapter, we outlined the limits presented in our proposed approach. These limits revolve around the use of a qualitative approach for likelihood analysis. We invite researchers to address the list of the limits, and to continue the work in considering safety and security jointly. Shared understanding of the challenges facing the domain will facilitate its rapid maturing.

Qualitative scale	Vulnerability Level	Designation
<b>exploitability</b>  Level of difficulty to exploit a given vulnerability	1	<b>Easy (E)</b> : No countermeasures are presented
	2	<b>Medium (M)</b> : Countermeasures are presented
	3	<b>Hard (H)</b> : Countermeasures existed with continuous review and improvements.

(a) Qualitative scale to characterize the vulnerability levels.

Qualitative scale	Difficulty Level	Designation
<b>Technical difficulty of an attack</b>	1	<b>Trivial (T)</b> : Little technical skill required
	2	<b>Moderate (M)</b> : Average cyber hacking skills required
	3	<b>Difficult (D)</b> : Demands a high degree of technical expertise
	4	<b>Very difficult (VD)</b> : Beyond the known capability of today's best hackers

(b) Qualitative scale to characterize the difficulty of conducting an attack.

Likelihood levels		Technical difficulty of an attack			
		T	M	D	VD
Exploitability	E	4	4	3	2
	M	4	3	2	1
	H	2	2	1	1

(c) Combining attack difficulty levels with the vulnerability levels to determine the likelihood of security input events.

**Figure 6.8** – Characterizing the likelihood of security basic events.

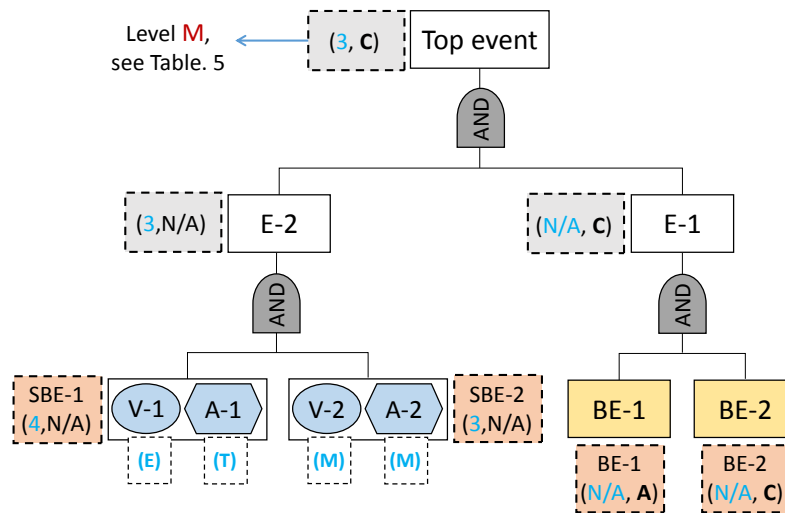


Figure 6.9 – Example of how calculating the likelihood of an MC.

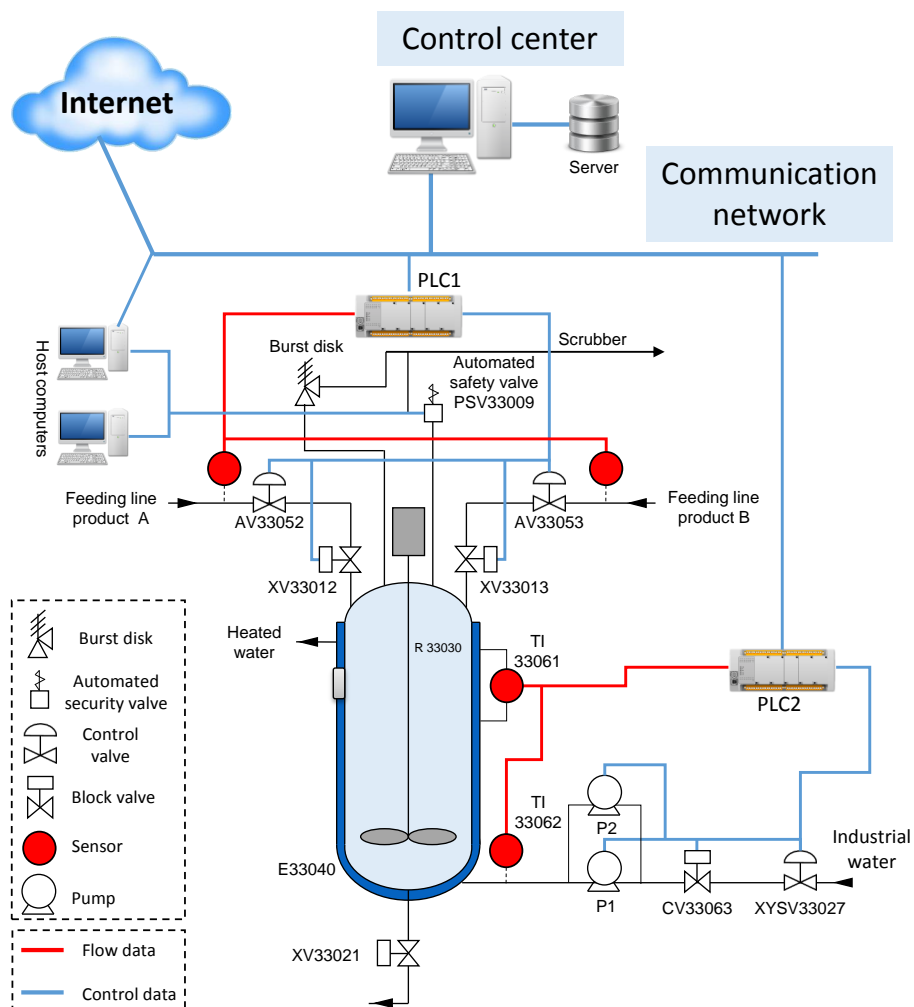


Figure 6.10 – The chemical reactor with its control system structure.



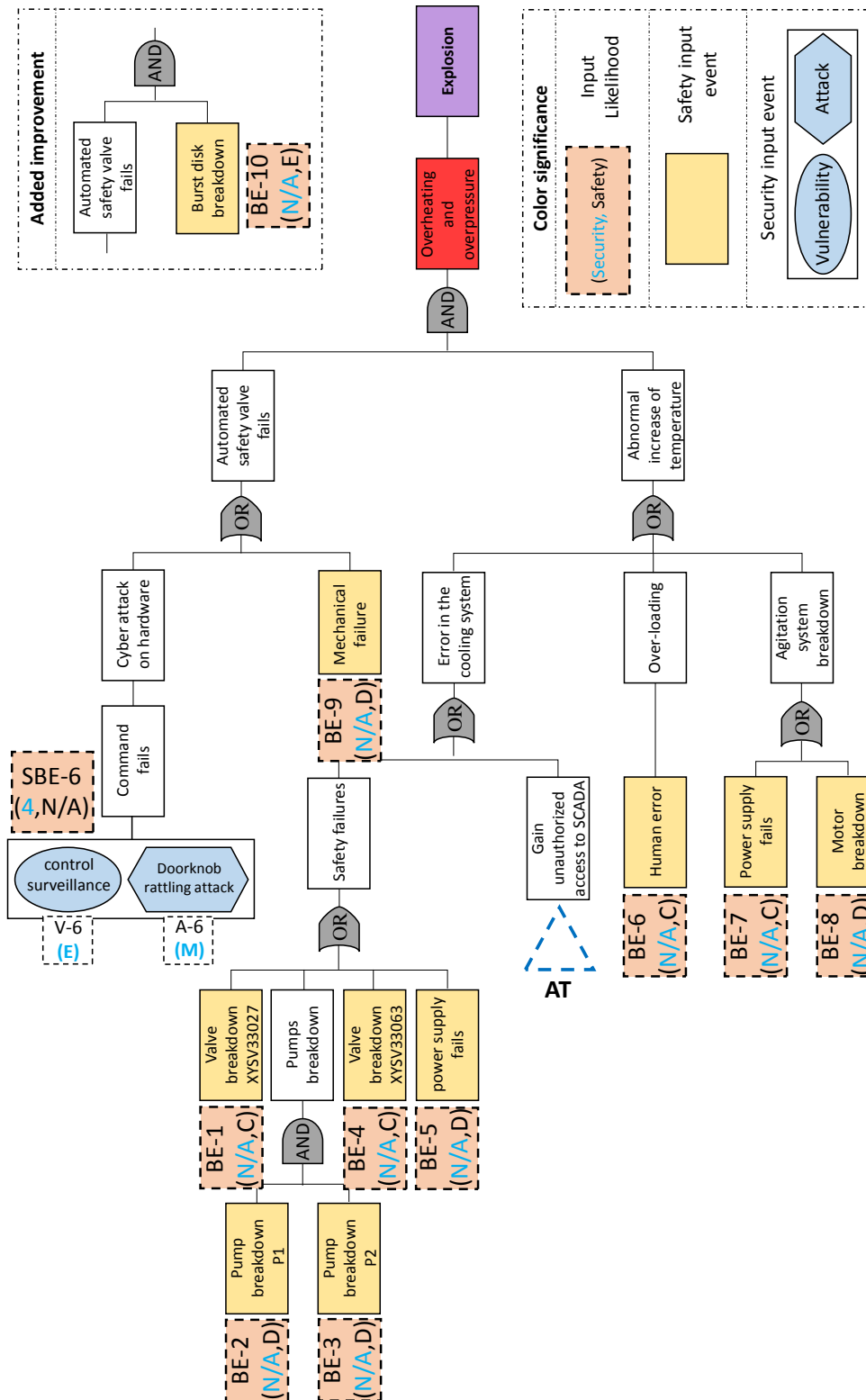


Figure 6.11 – Combined ATBT of the scenario under study.

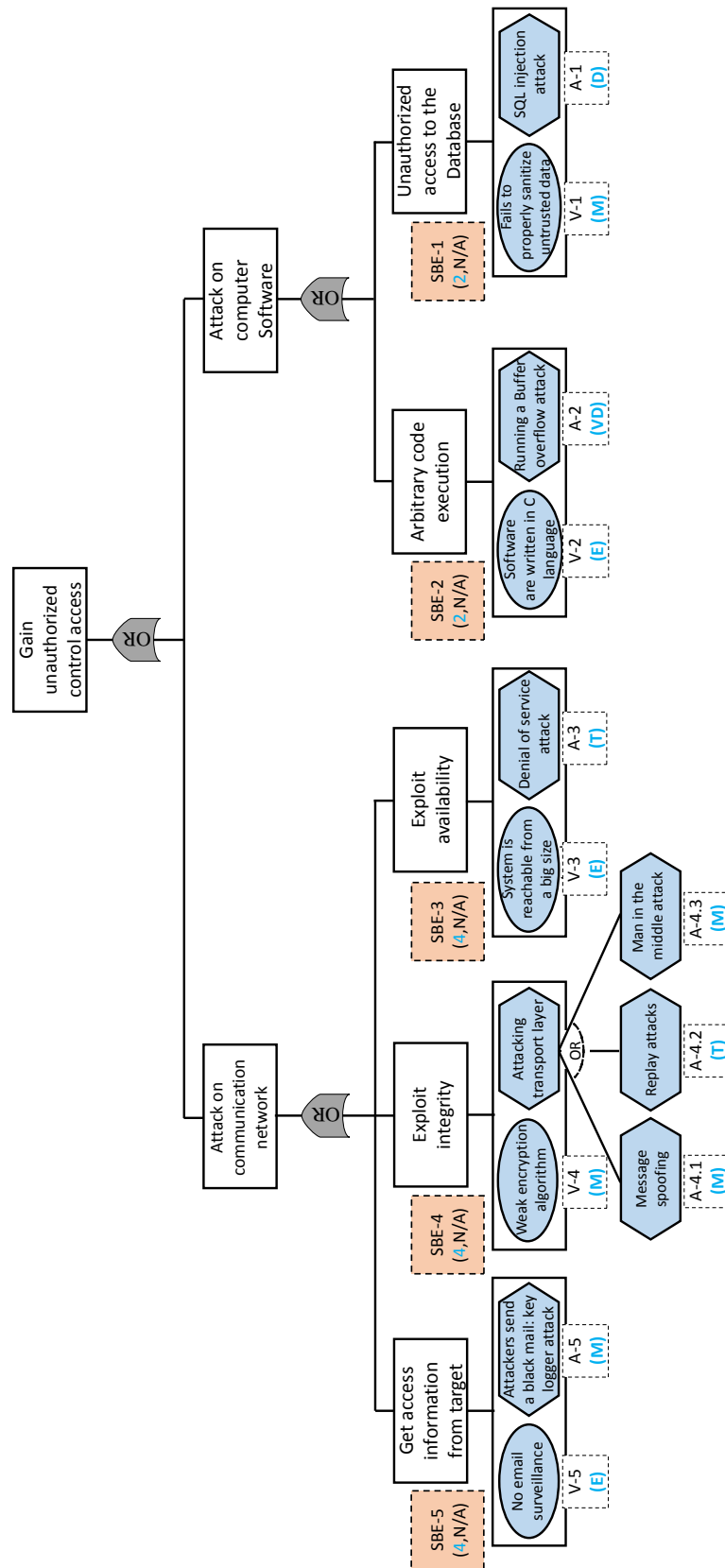


Figure 6.12 – AT for the goal: gain unauthorized access to SCADA.

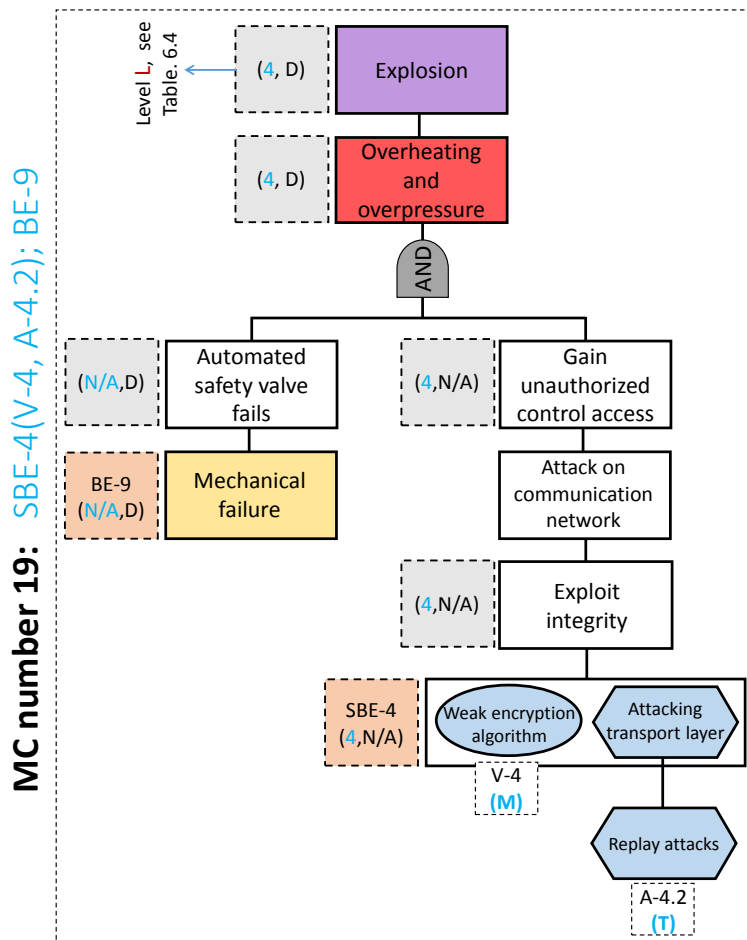


Figure 6.13 – Calculating the likelihood of MC number 19.

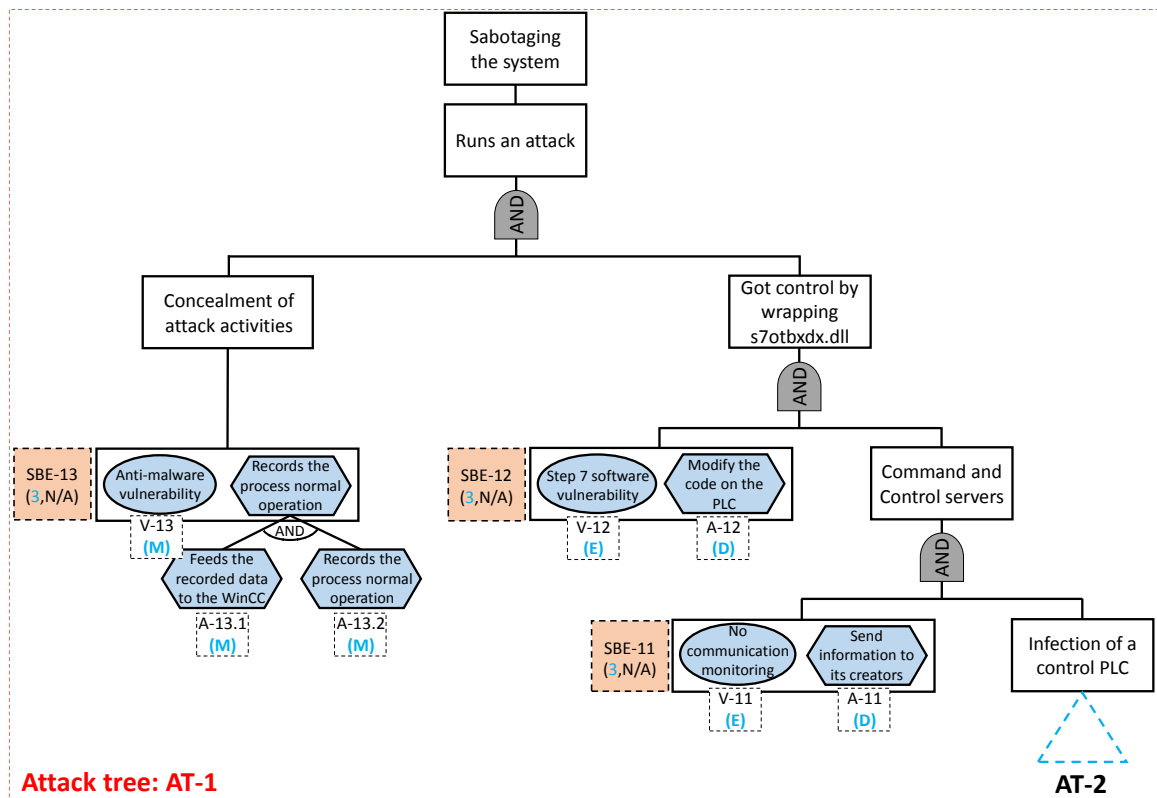


Figure 6.14 – The top objective of Stuxnet.

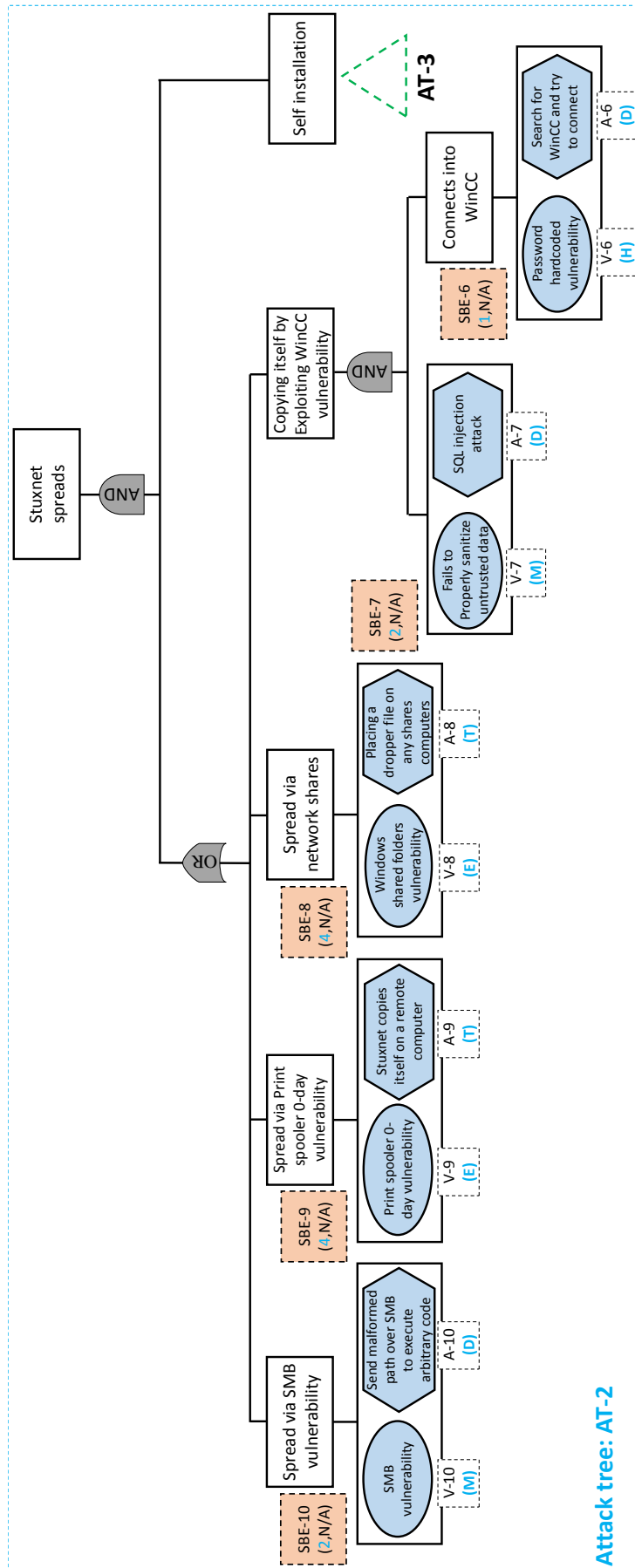


Figure 6.15 – Attack tree of the “spreading of Stuxnet”.

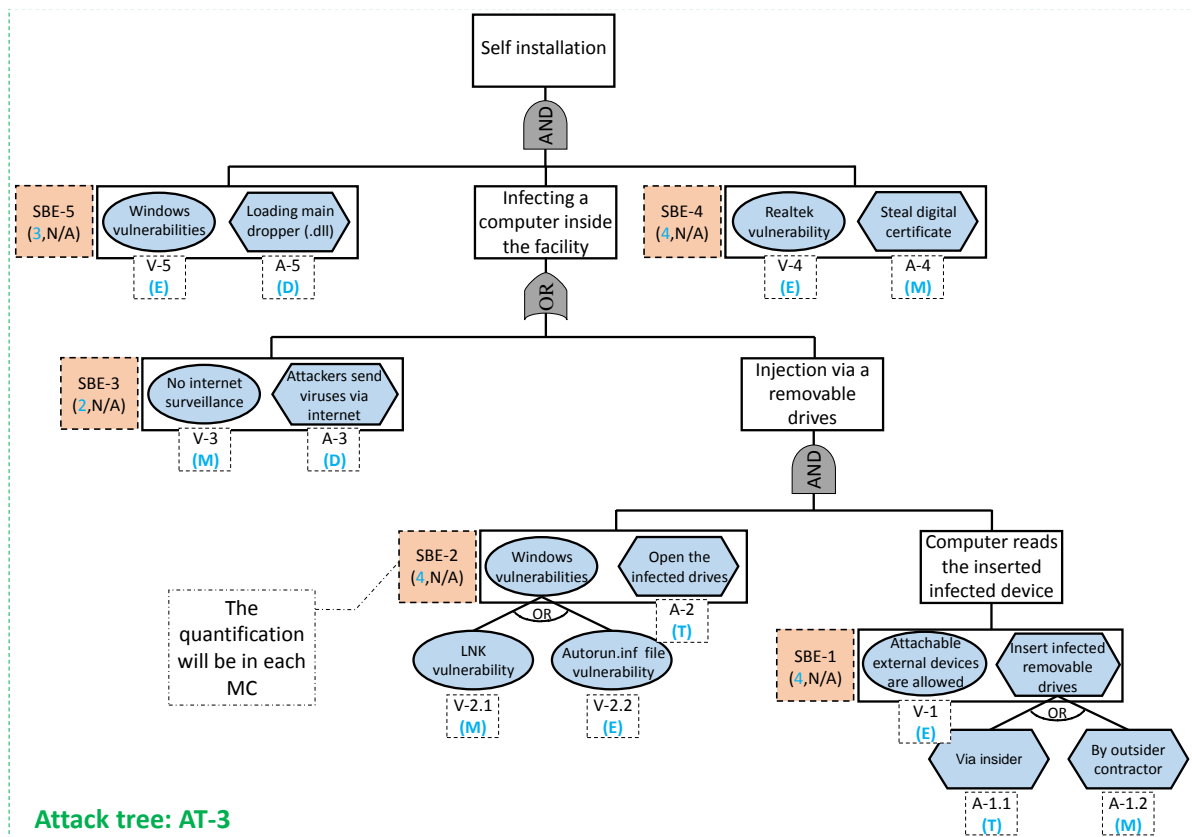
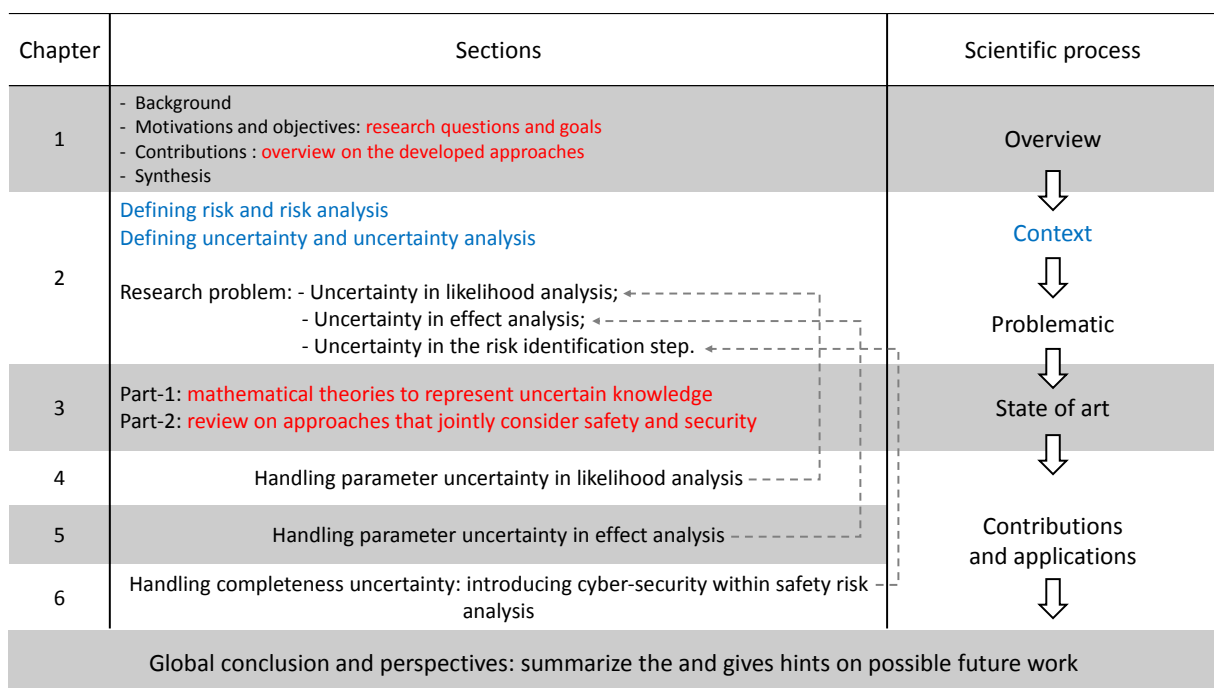


Figure 6.16 – Attack tree of “Stuxnet self installation”.



# Global conclusion and perspectives

Risk analysis is a critical part for regulatory decision-making related to high-risk risk industries. A systematic risk analysis process is made up of three steps: risk identification of undesirable risk scenarios, likelihood analysis and effect analysis. Identifying a risk scenario aims to explore how an undesirable hazard can be developed starting from causes and ending with the consequences. Likelihood analysis aims to estimate the likelihood of occurrence of risk scenarios. Effect or consequence analysis aims to calculate the effects of risk scenarios on human life and the environment. The INERIS uses the bow-tie analysis for identifying risk scenarios and analyzing their likelihoods. Complex mathematical models are used for effect analysis.



**Figure 6.17** – A global vision of the thesis structure.

However, most risk analysis studies struggle with uncertainty analysis and yet uncertainty with respect to its sources, types and causes is not well treated. Quantifying uncertainty during risk analysis has become an important part of effective decision-making



and health risk analysis. This thesis is motivated by the need of quantifying uncertainty in order to perform a sound risk analysis. Figure 6.17 presents how this document is structured in order to present the research problems and detail the proposed solutions.

Chapter 1 provides an overview on the motivations and objectives for conducting this work. Chapter 2 presents the different steps of a risk analysis process, the definition of uncertainty, the different sources, types and causes of uncertainty, why is it important to address uncertainty and how to address each source of uncertainty. The rest of chapter 2 highlights the research problem by identifying what sources of uncertainty affect the INERIS risk analysis process. Parameter and completeness uncertainties affect the analysis process as follows:

1. the likelihood analysis step suffers from parameter uncertainty: the interval semi-quantitative methodology used by the INERIS to estimate the probabilities of risk scenarios might lead to unreliable estimation due to uncertainty. Interval analysis is not the right mathematical theory to treat uncertainty during likelihood analysis. The use of interval analysis may lead to probability underestimation in specific cases. However, replacing interval analysis with the right mathematical theory to deal with uncertainty is the solution.
2. the effect analysis step suffers from parameter uncertainty: input parameters for these effect models are fraught with uncertainties. These uncertainties might be of different types (aleatoric, epistemic or mix of both types) depending on the types of these input parameters and the sources of available data. Epistemic uncertainty can be of different causes: imprecision, subjectivity, ignorance and lack of consensus. However, yet parameter uncertainty regarding its causes is not well treated. Different mathematical theories are needed to accurately represent each cause of parameter uncertainty. For this reason, a global new approach that use the right mathematical theory to treat each cause of parameter uncertainty is needed;
3. the identification process is incomplete: one of the most important incompleteness source is the introduction of connected systems and digital technology in process industries. This introduction creates new cyber-security vulnerabilities that can be exploited by sophisticated threats and lead to undesirable safety accidents. Thus, identifying these vulnerabilities during risk analysis becomes an important part for effective industrial risk evaluation. However, nowadays, safety and security are analyzed separately when they should not be. This is because a security threat can lead to the same dangerous phenomenon as a safety incident. However, these security related causes are not considered during risk analysis. Thus, a new risk analysis methodology that introduces security within safety risk analysis is an important need.

After presenting the context of the study and subtracting the research problems, a review of the literature on the subject was presented in Chapter 3. We have focused particularly on the different mathematical theories to represent uncertain knowledge and the existing approaches that jointly consider safety and security during risk analysis. We compare these approaches and discuss their limits in dealing with the problem we are facing.

Chapters 4, 5 and 6 present our contributions to deal with the above highlighted problems. These contributions are summarize in the next section. Section 6.7 gives hints on future work related to each contribution.

## Contributions of this work

After investigation of the different sources of uncertainty that affect the INERIS risk analysis process, three main contributions are developed in this work.

The first contribution is the development of a fuzzy semi-quantitative approach to treat parameter uncertainty in the likelihood analysis step (Chapter 4). As we said previously, the interval semi-quantitative approach used by the INERIS for probability analysis may result in probability underestimation in some cases due to uncertainty. For this reason, Chapter 4 proposes a fuzzy semi-quantitative bow-tie analysis to address data uncertainty and overcome the limits of the interval-based approach. A fuzzy-based approach is used for handling subjectivity and imprecision in the input parameters. By introducing fuzzy logic, we solved the problem of probability underestimation of interval-based approach. The discreteness and vagueness between frequency classes are removed. A measurement error in the input data will not affect the decision making anymore. Furthermore, the fuzzy semi-quantitative approach shows a simple more conservative approach than the interval semi-quantitative, the pure quantitative and the qualitative approaches;

The second contribution is the development of a global approach to deal with parameter uncertainty in effect analysis (Chapter 5). Here, we rode out from the fact that different methods are needed to represent variability, imprecision, ignorance and lack of consensus. Chapter 5 first proposes a mixed fuzzy-probabilistic approach to separately treat variability and imprecision. This approach is then compared with the existing approaches to treat parameter uncertainty in effect analysis. As result, a guideline on how the best to treat parameter uncertainty is provided. We concluded that probability distributions, fuzzy numbers and evidence theory are the most suitable theories to deal with variability, imprecision and vagueness. Based on this conclusion, a new approach that treat each cause of parameter uncertainty with the right theory is developed. This approach combines probability, possibility and belief to separately represent and treat

variability, imprecision and ignorance. A new propagation algorithm is develop that uses 2-stages MC simulation and evidence calculus to propagate the represented uncertain parameters through the effect models. The approach is named the ALI-approach because after propagating uncertainty we aggregate the different likelihood indexes (probability, possibility and belief) in a single one. The output of the ALI approach after aggregating the likelihood indexes is a single distribution. This approach provides the most precise treatment of parameter uncertainty and simpler output for decision makers by aggregating likelihood indexes;

The third contribution is the proposition of a new approach for the treatment of incompleteness uncertainty by introducing cyber-security beside safety within industrial risk analysis (Chapter 6). This approach combines bow-tie analysis developed by the INERIS, with a new extended version of attack tree analysis, introduced for security analysis of industrial control systems. The proposed new version of attack tree depict more information about the target system in order to provide a more meaningful modeling of security risk scenarios. The combined use of bow-tie and attack tree provides an exhaustive representation of risk scenarios in terms of safety and security. We then propose an approach for evaluating the risk level based on two-term likelihood parts, one for safety and one for security. Two-term likelihood parts are proposed to declare the difference between safety and security related risk causes.

These contributions are demonstrated on real case studies. This work has resulted in publications in international journals ([4], [7] and [9]). We also published our findings at international conferences ([2], [3], [5] and [6]) and national ([50]).

## Suggestions for Future Work

As general perspectives for this work, we mention here the most important:

- the fuzzy semi-quantitative approach can be extended by using multiple sources of data in probability analysis. Different data bases or experts may provide different probabilities regarding the same parameter. This result in lack of consensus between the different sources of data. Thus, rating and aggregating the data from different sources will lead to a more robust probability quantification approach for bow-tie analysis. Rating the sources of data can be based on how detailed the data bases are, and the level of expertise of experts in the filed of the analysis. The questions here are: how to combine all sources of data in the same formalism, and how to propagate this information through the bow-tie model. By using all information from the all available sources of data, the decision making will be more effective and conservative;

- in this document, we discuss decision-making under uncertainty in a simplified way. A decision making framework under uncertainty can be developed to facilitate making decisions for the concerned authorities. Providing guidelines for decision makers is an important research subject. The guidelines would help in choosing the best degree of confidence for the decision regarding specific criteria such as the types of related hazards, the surrounding environment etc,. Based on this, a reasonable decision would be made rather than being non-conservative or too conservative;
- the qualitative approach for a combined safety/security risk analysis can be extended by proposing a more robust likelihood evaluation technique based on the developed cyber bow-tie. The fuzzy semi-quantitative approach developed in Chapter 6 for likelihood analysis can be adjusted and applied for safety/security likelihood analysis. The idea is to propose fuzzy scales to characterize the criteria used to evaluate the likelihood of security related events (attack difficulties and vulnerability levels). The likelihood of security related events will be represented in terms of frequency classes with their membership degrees. Moving toward a fuzzy semi-quantitative approach based on the proposed cyber bow-tie analysis (the ATBT developed in Chapter 6) will provide more precise likelihood evaluation of safety/security risks. Section 6.4.3 of Chapter 6 details more possible improvements and future work;
- omitting or ignoring risk events can affect both the accuracy and precision of the analysis results. So, it might be desirable to address the unknown completeness uncertainty and study its impact on the analysis. Providing a level of confidence on how complete the analysis is will provide a helpful information for risk assessors and decision makers. Knowing if the analysis is complete or not would help its updating once more information is available in the future.



# Bibliography

- [1] H. ABDO et J-M FLAUS. A mixed fuzzy probabilistic approach for risk assessment of dynamic systems. *IFAC-PapersOnLine*, 48(3):960 – 965, 2015. 15th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2015.
- [2] H. ABDO et J-M. FLAUS. A mixed fuzzy probabilistic approach for risk assessment of dynamic systems. *IFAC-PapersOnLine*, 48(3):960–965, 2015.
- [3] H. ABDO et J-M. FLAUS. Uncertainty quantification in bow-tie analysis: A mixed approach of fuzzy theory with dempster-shafer theory of evidence. *In Risk, Reliability and Safety: Innovating Theory and Practice: Proceedings of ESREL (Glasgow, Scotland)*, pages 2743–2750. Taylor & Francis, 2016.
- [4] H. ABDO et J-M. FLAUS. Uncertainty quantification in dynamic system risk assessment: a new approach with randomness and fuzzy theory. *International Journal of Production Research*, pages 1–24, 2016.
- [5] H. ABDO, J-M. FLAUS et F. MASSE. Fuzzy semi-quantitative approach for probability evaluation using bow-tie analysis. pages 376–376, 2017.
- [6] H. ABDO, J-M. FLAUS et F. MASSE. Towards a better industrial risk analysis: A new approach that combines cyber security within safety. pages 179–187, 2017.
- [7] H. ABDO, J-M. FLAUS et F. MASSE. Uncertainty quantification in risk assessment - representation, propagation and treatment approaches: Application to atmospheric dispersion modeling. *Journal of Loss Prevention in the Process Industries*, 2017.
- [8] H. ABDO, M. KAOUK, J.-M. FLAUS et F. MASSE. Fuzzy semi-quantitative approach for probability evaluation using bow-tie analysis. *reliability engineering & system safety*, 2017.
- [9] H. ABDO, M. KAOUK, J.-M. FLAUS et F. MASSE. A safety/security risk analysis approach of industrial control systems: A cyber bowtie – combining new version of attack tree with bowtie analysis. *Computers & Security*, 72(Supplement C):175 – 195, 2017.
- [10] Gotz ALEFELD et Jurgen HERZBERGER. *Introduction to interval computation*. Academic press, 1983.
- [11] Götz ALEFELD et Günther MAYER. Interval analysis: theory and applications. *Journal of Computational and Applied Mathematics*, 121(1–2):421 – 464, 2000.

- [12] Abdulmohsen ALMALAWI, Xinghuo YU, Zahir TARI, Adil FAHAD et Ibrahim KHALIL. An unsupervised anomaly-based detection approach for integrity attacks on scada systems. *Computers & Security*, 46:94 – 110, 2014.
- [13] G APOSTOLAKIS. The concept of probability in safety assessments of technological systems. *Science*, 250(4986):1359–1364, 1990.
- [14] N.S. ARUNRAJ, Saptarshi MANDAL et J. MAITI. Modeling uncertainty in risk assessment: An integrated approach with fuzzy set theory and monte carlo simulation. *Accident Analysis and Prevention*, 55(0):242 – 255, 2013.
- [15] Robert B ASH et Catherine DOLEANS-DADE. *Probability and measure theory*. Academic Press, 2000.
- [16] Terje AVEN. On the need for restricting the probabilistic analysis in risk assessments to variability. *Risk analysis*, 30(3):354–360, 2010.
- [17] Terje AVEN. Some reflections on uncertainty analysis and management. *Reliability Engineering & System Safety*, 95(3):195–201, 2010.
- [18] Luis AYALA. *Cybersecurity Lexicon*. Springer, 2016.
- [19] Bilal M AYYUB et George J KLIR. *Uncertainty modeling and analysis in engineering and the sciences*. CRC Press, 2006.
- [20] Jean BACCOU, Eric CHOJNACKI, Catherine MERCAT-ROMMENS et Cédric BAUDRIT. Extending monte carlo simulations to represent and propagate uncertainties in presence of incomplete knowledge: application to the transfer of a radionuclide in the environment. *Journal of environmental engineering*, 134(5):362–368, 2008.
- [21] Nadeau S. Gbodossou A. BADRI, A.. Proposal of a risk-factor-based analytical approach for integrating occupational health and safety into project risk evaluation. 2012.
- [22] Morton L BARAD. Project prairie grass, a field program in diffusion. volume ii. Rapport technique, DTIC Document, 1958.
- [23] C BAUDRIT, S DESTERCKE et PH WUILLEMIN. Unifying parameter learning and modelling complex systems with epistemic uncertainty using probability interval. *Information Sciences*, 2016.
- [24] Cédric BAUDRIT, Inés COUSO et Didier DUBOIS. Joint propagation of probability and possibility in risk analysis: Towards a formal framework. *International Journal of Approximate Reasoning*, 45(1):82–105, 2007.
- [25] Pierre BIEBER, Jean-Paul BLANQUART, Gilles DESCARGUES, Michael DULUCQ, Yannick FOURASTIER, Eric HAZANE, Mathias JULIEN, Laurent LÉONARDON et Gabrielle SAROUILLE. Security and safety assurance for aerospace embedded systems. In *Proceedings of the 6th International Conference on Embedded Real Time Software and Systems, Toulouse, France*, pages 1–10, 2012.

- 
- [26] Patrick BILLINGSLEY. *Probability and measure*. John Wiley & Sons, 2008.
- [27] GA BRIGGS. Lift off of buoyant gas initially on the ground. *ADTL Contribution File*, (87), 1973.
- [28] Eric J BYRES, Matthew FRANZ et Darrin MILLER. The use of attack trees in assessing vulnerabilities in scada systems. *In Proceedings of the international infrastructure survivability workshop*, 2004.
- [29] Netherlands National Institute of Public Health Center for EXTERNAL SAFETY et the ENVIRONMENT. *Reference manual BEVI risk assessments*, volume v3.1. 2009.
- [30] Sanjay S CHAUHAN et David S BOWLES. Dam safety risk assessment with uncertainty analysis. *Ancold Bulletin*, pages 73–88, 2004.
- [31] Yulia CHERDANTSEVA, Pete BURNAP, Andrew BLYTH, Peter EDEN, Kevin JONES, Hugh SOULSBY et Kristan STODDART. A review of cyber security risk assessment methods for scada systems. *Computers & Security*, 56:1 – 27, 2016.
- [32] Yulia CHERDANTSEVA, Pete BURNAP, Andrew BLYTH, Peter EDEN, Kevin JONES, Hugh SOULSBY et Kristan STODDART. A review of cyber security risk assessment methods for scada systems. *computers & security*, 56:1–27, 2016.
- [33] Bongsik CHU, Sangick LEE et Daejun CHANG. Determination of design accidental fire load for offshore installations based on quantitative risk assessment with treatment of parametric uncertainty. *Journal of Loss Prevention in the Process Industries*, 45:160 – 172, 2017.
- [34] Rituparna CHUTIA, Supahi MAHANTA et D DATTA. Uncertainty modelling of atmospheric dispersion by stochastic response surface method under aleatory and epistemic uncertainties. *Sadhana*, 39(2):467–485, 2014.
- [35] RobertT. CLEMEN et RobertL. WINKLER. Combining probability distributions from experts in risk analysis. *Risk Analysis*, 19(2):187–203, 1999.
- [36] International Electrotechnical COMMISSION *et al.*. Functional safety-safety instrumented systems for the process industry sector". *IEC61511*, 1:9–5, 2003.
- [37] JR González DAN, A GUIX, V MARTÍ, Josep ARNALDOS et RM DARBRA. Monte carlo simulation as a tool to show the influence of the human factor into the quantitative risk assessment. *Process Safety and Environmental Protection*, 102:441–449, 2016.
- [38] VJ DAOO, NS PANCHAL, Faby SUNNY et V Venkat RAJ. Scintillometric measurements of daytime atmospheric turbulent heat and momentum fluxes and their application to atmospheric stability evaluation. *Experimental thermal and fluid science*, 28(4):337–345, 2004.
- [39] Ryks J. Fazil A. DAVIDSON, V.J.. Fuzzy risk assessment tool for microbial hazards in food systems. 2006.



- [40] Valérie DE DIANOUS, Céline DEUST, Charlotte BOUISSOU, Régis FARRET et Sylvain CHAUMETTE. Prise en compte de la probabilité dans les études de dangers. *Préventique Sécurité*, (95):32–37, 2014.
- [41] Bruno DE FINETTI. Theory of probability. a critical introductory treatment. 1979.
- [42] Inc. DELL. Dell Security Annual Threat Report. Rapport technique, 01 2015.
- [43] Yong DENG, Rehan SADIQ, Wen JIANG et Solomon TESFAMARIAM. Risk analysis in a linguistic environment: a fuzzy evidential reasoning-based approach. *Expert Systems with Applications*, 38(12):15438–15446, 2011.
- [44] Thierry DENOEU. Maximum likelihood estimation from uncertain data in the belief function framework. *Knowledge and Data Engineering, IEEE Transactions on*, 25(1):119–130, 2013.
- [45] U.S. department of ENERGY. MACCS2 Computer Code Application Guidance for Documented Safety Analysis, Final Report. June 2004.
- [46] Armen DER KIUREGHIAN et Ove DITLEVSEN. Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105–112, 2009.
- [47] M DROUIN. *Guidance on the Treatment of Uncertainties Associated with PRAs in Risk-informed Decision Making: Main Report*. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Office of Nuclear Reactor Regulation, 2009.
- [48] Didier DUBOIS et Henri PRADE. *Possibility theory*. Wiley Online Library, 1988.
- [49] U.S. Environmental Protection Agency (EPA). Process for conducting probabilistic risk assessment, rags volume 3 part a. 2001.
- [50] H. Abdo F. MASSE et J-M. FLAUS. Vers une approche intégrant les exigences de cybersécurité à la maîtrise des risques d’accidents majeurs pour les icpe). In *12ème Congrès International Pluridisciplinaire en Qualité, Sûreté de fonctionnement et Développement durable (Bourges, France)*, 2017.
- [51] Nicolas FALLIERE, Liam O MURCHU et Eric CHIEN. W32. stuxnet dossier. *White paper, Symantec Corp., Security Response*, 5(6), 2011.
- [52] Willliam FELLER. *An introduction to probability theory and its applications*, volume 2. John Wiley & Sons, 2008.
- [53] Refaul FERDOUS, Faisal KHAN, Rehan SADIQ, Paul AMYOTTE et Brian VEITCH. Handling data uncertainties in event tree analysis. *Process safety and environmental protection*, 87(5):283–292, 2009.
- [54] Refaul FERDOUS, Faisal KHAN, Rehan SADIQ, Paul AMYOTTE et Brian VEITCH. Fault and event tree analyses for process systems risk analysis: uncertainty handling formulations. *Risk Analysis*, 31(1):86–107, 2011.

- 
- [55] Refaul FERDOUS, Faisal KHAN, Rehan SADIQ, Paul AMYOTTE et Brian VEITCH. Handling and updating uncertain information in bow-tie analysis. *Journal of Loss Prevention in the Process Industries*, 25(1):8–19, 2012.
- [56] Refaul FERDOUS, Faisal KHAN, Rehan SADIQ, Paul AMYOTTE et Brian VEITCH. Analyzing system safety and risks under uncertainty using a bow-tie diagram: An innovative approach. *Process Safety and Environmental Protection*, 91(1):1–18, 2013.
- [57] Refaul FERDOUS, Faisal KHAN, Brian VEITCH et Paul R AMYOTTE. Methodology for computer aided fuzzy fault tree analysis. *Process safety and environmental protection*, 87(4):217–226, 2009.
- [58] S FERSON. Using fuzzy arithmetic in monte carlo simulation of fishery populations. *Management strategies for exploited fish populations. Alaska Sea Grant College Program, Fairbanks*, pages 595–608, 1993.
- [59] Scott FERSON. *RAMAS Risk Calc 4.0 software: risk assessment with uncertain numbers*. CRC press, 2002.
- [60] Scott FERSON et Lev R GINZBURG. Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, 54(2):133–144, 1996.
- [61] Scott FERSON, William L. OBERKAMPF et Lev GINZBURG. Validation of imprecise probability models, 2009.
- [62] Donald G FIRESMITH. Common concepts underlying safety security and survivability engineering. Rapport technique, Software Engineering Institute, December 2003.
- [63] Jean-Marie FLAUS. *Risk analysis: socio-technical and industrial systems*. John Wiley & Sons, 2013.
- [64] French National Institute for INDUSTRIAL ENVIRONMENT et Risks (INERIS). *Guide de maitrise des risques technologiques dans les depots de liquides inflammables*, volume 8. 10. 2008.
- [65] French National Institute for INDUSTRIAL ENVIRONMENT et Risks (INERIS). *Proposition d’une mthode semi-quantitative d’valuation des probabilités d’inflammation*. 2015.
- [66] John A FOSTER. Psychophysical causal relations. *American Philosophical Quarterly*, pages 64–70, 1968.
- [67] Igor Nai FOVINO et Marcelo MASERA. Through the description of attacks: A multidimensional view. *In International Conference on Computer Safety, Reliability, and Security*, pages 15–28. Springer, 2006.
- [68] Igor Nai FOVINO, Marcelo MASERA et Alessio De CIAN. Integrating cyber attacks within fault trees. *Reliability Engineering & System Safety*, 94(9):1394 – 1402, 2009. ESREL 2007, the 18th European Safety and Reliability Conference.

- [69] Raymond FREEMAN, Angela SUMMERS *et al.*. Evaluation of uncertainty in safety integrity level calculations. *Process Safety Progress*, 2016.
- [70] Rubin E.S. FREY, H.C.. Evaluate uncertainties in advanced process technologies. 1992.
- [71] L GOOIJER, N CORNIL et CL LENOBLE. An international comparison of four quantitative risk assessment approaches - a benchmark study based on a fictitious lpg plant. *Process safety and environmental protection*, 90(2):101–107, 2012.
- [72] Stanley S GROSSEL. Guidelines for chemical process quantitative risk analysis: ; by center for chemical process safety; american institute of chemical engineers, new york, ny, 2000, pp. 750., 2001.
- [73] Jiwen W GUAN et David A BELL. *Evidence theory and its applications*. Elsevier Science Inc., 1991.
- [74] Dominique GUYONNET, Bernard BOURGINE, Didier DUBOIS, Hélène FARGIER, Bernard CÔME et Jean-Paul CHILÈS. Hybrid approach for addressing uncertainty in risk assessments. *Journal of environmental engineering*, 129(1):68–78, 2003.
- [75] J. HALLIKAS, V.-M. VIROLAINEN et M. TUOMINEN. Understanding risk and uncertainty in supplier networks—a transaction cost approach. *International Journal of Production Research*, 40(15):3519–3531, 2002.
- [76] Steven R HANNA, Joseph C CHANG et Mark E FERNAU. Monte carlo estimates of uncertainties in predictions by a photochemical grid model (uam-iv) due to uncertainties in input variables. *Atmospheric Environment*, 32(21):3619–3628, 1998.
- [77] DUANE A HAUGEN. Project prairie grass. a field program in diffusion. volume 3. Rapport technique, DTIC Document, 1959.
- [78] J.C. HELTON, J.D. JOHNSON et W.L. OBERKAMPF. An exploration of alternative approaches to the representation of uncertainty in model predictions. *Reliability Engineering and System Safety*, 85(1–3):39 – 71, 2004. Alternative Representations of Epistemic Uncertainty.
- [79] Morgan HENRIE. Cyber security risk management in the scada critical infrastructure environment. *Engineering Management Journal*, 25(2):38–45, 2013.
- [80] Matthew H HENRY, Ryan M LAYER, Kevin Z SNOW et David R ZARET. Evaluating the risk of cyber attacks on scada systems via petri net analysis with application to hazardous liquid loading operations. In *Technologies for Homeland Security, 2009. HST'09. IEEE Conference on*, pages 607–614. IEEE, 2009.
- [81] A.J. HOBDAY, A.D.M. SMITH, I.C. STOBUTZKI, C. BULMAN, R. DALEY, J.M. DAMBACHER, R.A. DENG, J. DOWDNEY, M. FULLER, D. FURLANI, S.P. GRIFFITHS, D. JOHNSON, R. KENYON, I.A. KNUCKEY, S.D. LING, R. PITCHER, K.J. SAINSBURY, M. SPORCIC, T. SMITH, C. TURNBULL, T.I. WALKER, S.E. WAYTE, H. WEBB, A. WILLIAMS, B.S. WISE et S. ZHOU. Ecological risk assessment for the effects of fishing. *Fisheries Research*, 108(2–3):372 – 384, 2011.

- 
- [82] Diane HONEYCUTT, Northrop GRUMMAN *et al.*. Developing a framework to improve critical infrastructure cybersecurity. 2013.
- [83] David HOURTOLOU et Olivier SALVI. Aramis project: development of an integrated accidental risk assessment methodology for industries in the framework of seveso ii directive. *In International Conference on Safety and Reliability (ESREL 2003)*, pages 829–836, 2003.
- [84] David HOURTOLOU et Olivier SALVI. Aramis project: Achievement of the integrated methodology and discussion about its usability from the case studies carried out on real test seveso ii sites. *In 11. International Symposium on Loss Prevention and Safety Promotion in the Process Industry*, pages 1133–1143. PetroChemEng. Praha, 2004.
- [85] IEC61508 IEC. 61508 functional safety of electrical/electronic/programmable electronic safety-related systems. *International electrotechnical commission*, 1998.
- [86] INERIS. Agrégation semi-quantitative des probabilités dans les études de dangers des installations classées - omega probabilités. 2015.
- [87] ANSI ISA. 84.00. 01-2004 part 3 (iec 61511 modified) functional safety: Safety instrumented systems for the process industry sector. *Research Triangle Park, NC*, 2004.
- [88] Afshin JAMSHIDI, Samira Abbasgholizadeh RAHIMI, Daoud AIT-KADI et Angel RUIZ. A comprehensive fuzzy risk-based maintenance framework for prioritization of medical devices. *Applied Soft Computing*, 32:322 – 334, 2015.
- [89] CW JOHNSON. Cybersafety: on the interactions between cybersecurity and the software engineering of safety-critical systems. *Achieving System Safety*, pages 85–96, 2012.
- [90] C JOLY, S DESCOURRIERE, R FARRET et B DEBRAY. L'étude de dangers d'une installation classée, 2006.
- [91] Anssi KÄKI, Juuso LIESIÖ, Ahti SALO et Srinivas TALLURI. Newsvendor decisions under supply uncertainty. *International Journal of Production Research*, 53(5): 1544–1560, 2015.
- [92] Stanley KAPLAN et B John GARRICK. On the quantitative definition of risk. *Risk analysis*, 1(1):11–27, 1981.
- [93] Bilge KARABACAK et Ibrahim SOGUKPINAR. Isram: information security risk analysis method. *Computers & Security*, 24(2):147 – 159, 2005.
- [94] Stamatis KARNOUSKOS. Stuxnet worm impact on industrial cyber-physical system security. *In IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society*, pages 4490–4494. IEEE, 2011.
- [95] E KENTEL et MM ARAL. Probabilistic-fuzzy health risk modeling. *Stochastic Environmental Research and Risk Assessment*, 18(5):324–338, 2004.

- [96] Arvind KEPRATE et RM Chandima RATNAYAKE. Enhancing offshore process safety by selecting fatigue critical piping locations for inspection using fuzzy-ahp based approach. *Process Safety and Environmental Protection*, 102:71–84, 2016.
- [97] Joris KOORNNEEF, Mark SPRUIJT, Menso MOLAG, Andrea RAMIREZ, Wim TURKENBURG et André FAAIJ. Quantitative risk assessment of co2 transport by pipelines—a review of uncertainties and their impacts. *Journal of Hazardous Materials*, 177(1–3):12 – 27, 2010.
- [98] Andrew J KORNECKI, Nary SUBRAMANIAN et Janusz ZALEWSKI. Studying inter-relationships of safety and security for software assurance in cyber-physical systems: Approach based on bayesian belief networks. *In Computer Science and Information Systems (FedCSIS), 2013 Federated Conference on*, pages 1393–1399. IEEE, 2013.
- [99] Andrew J KORNECKI et Janusz ZALEWSKI. Safety and security in industrial control. *In Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research*, page 77. ACM, 2010.
- [100] Siwar KRIAA, Marc BOUISSOU et Ludovic PIÈTRE-CAMBACÉDÈS. Modeling the stuxnet attack with bdmp: Towards more formal risk assessments. *In Risk and Security of Internet and Systems (CRiSIS), 2012 7th International Conference on*, pages 1–8. IEEE, 2012.
- [101] Siwar KRIAA, Ludovic PIETRE-CAMBACEDES, Marc BOUISSOU et Yoran HALGAND. A survey of approaches combining safety and security for industrial control systems. *Reliability Engineering & System Safety*, 139:156–178, 2015.
- [102] Siwar KRIAA, Ludovic PIETRE-CAMBACEDES, Marc BOUISSOU et Yoran HALGAND. A survey of approaches combining safety and security for industrial control systems. *Reliability Engineering & System Safety*, 139:156–178, 2015.
- [103] Goldy KUMAR et J. MAITI. Modeling risk based maintenance using fuzzy analytic network process. *Expert Systems with Applications*, 39(11):9946 – 9954, 2012.
- [104] Matthew LEITCH. Iso 31000: 2009—the new international standard on risk management. *Risk Analysis*, 30(6):887–892, 2010.
- [105] CJA LENOBLE, ES KOOI et FN ANTOINE. Benchmark study for a flammable liquid depot: comparison of two risk assessments. *RIVM report 620001002, INERIS DRA-09-102989-08638A*, 2012.
- [106] Weize LI, Lun XIE, Zulan DENG et Zhiliang WANG. False sequential logic attack on scada system and its physical impact analysis. *Computers & Security*, 58:149 – 159, 2016.
- [107] Ching-Torng LIN et Mao-Jiun J WANG. Hybrid fault tree analysis using fuzzy sets. *Reliability Engineering & System Safety*, 58(3):205–213, 1997.
- [108] Yuan LIU et Hong MAN. Network vulnerability assessment using bayesian networks. *In Proc. SPIE*, volume 5812, pages 61–71, 2005.

- 
- [109] Kim E LOWELL et Kurt K BENKE. Uncertainty and risk analysis in hydrological models for land-use management. *In Accuracy 2006: International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, pages 740–749. Citeseer, 2006.
- [110] JD MARKOVIC-PETROVIC et MD STOJANOVIC. An improved risk assessment method for scada information security. *Elektronika ir Elektrotehnika*, 20(7):69–72, 2014.
- [111] Adam Stanistaw MARKOWSKI et Dorota SIUTA. Fuzzy logic approach to calculation of thermal hazard distances in process industries. *Process Safety and Environmental Protection*, 92(4):338 – 345, 2014. Loss Prevention 2013.
- [112] W MARQUARDT. Chemical engineering dynamics. von j. ingham, i. dunn, e. heinzle und e. prenusil. vch verlagsgesellschaft mbh, weinheim 1994. 701 s. mit 430 abb., geb., dm 276,-. *Chemie Ingenieur Technik*, 67(4):500–501, 1995.
- [113] M.S. Pamela MCCAULEY-BELL et Adedeji B. BADIRU. A fuzzy linguistics model for job related injury risk assessment. *Computers & Industrial Engineering*, 23(1):209 – 212, 1992.
- [114] Robert MCMILLAN. Siemens: Stuxnet worm hit industrial systems. *Computerworld*, 14, 2010.
- [115] MERGE. Merge safety & security project–multi-concerns interactions system engineering.
- [116] Ramon E MOORE. *Interval analysis*, volume 4. Prentice-Hall Englewood Cliffs, 1966.
- [117] Ramon E MOORE, R Baker KEARFOTT et Michael J CLOUD. *Introduction to interval analysis*. SIAM, 2009.
- [118] Paul MUELLER et Babak YADEGARI. The stuxnet worm. *Département des sciences de l’informatique, Université de l’Arizona*, 2012.
- [119] C NEGOITA, L ZADEH et H ZIMMERMANN. Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*, 1:3–28, 1978.
- [120] A. NICHOLSON, S. WEBBER, S. DYER, T. PATEL et H. JANICKE. Scada security in the light of cyber-warfare. *Computers & Security*, 31(4):418 – 436, 2012.
- [121] Emanuel PARZEN. *Modern probability theory and its applications*. John Wiley & Sons, Incorporated, 1960.
- [122] Frank PASQUILL et Paul MICHAEL. Atmospheric diffusion. *Physics Today*, 30:55, 1977.
- [123] S PATEL, R TANTALEAN, P RALSTON et J GRAHAM. Supervisory control and data acquisition remote terminal unit testbed. *Intelligent Systems Research Laboratory technical report TR-ISRL-05-01, Department of Computer Engineering and Computer Science. Louisville, Kentucky: University of Louisville*, 2005.

- [124] Nicola PEDRONI, Enrico ZIO, Elisa FERRARIO, Alberto PASANISI et Mathieu COU-  
PLET. Hierarchical propagation of probabilistic and non-probabilistic uncertainty  
in the parameters of a risk model. *Computers & Structures*, 126:199–213, 2013.
- [125] Ludovic PIÈTRE-CAMBACÉDÈS et Marc BOUISSOU. Modeling safety and security  
interdependencies with bdmp (boolean logic driven markov processes). *In Systems  
Man and Cybernetics (SMC), 2010 IEEE International Conference on*, pages 2852–  
2861. IEEE, 2010.
- [126] Ludovic PIÈTRE-CAMBACÉDÈS, Yann DEFLESSELLE et Marc BOUISSOU. Security  
modeling with bdmp: from theory to implementation. *In Network and Information  
Systems Security (SAR-SSI), 2011 Conference on*, pages 1–8. IEEE, 2011.
- [127] Ludovic PIETRE-CAMBACEDES, Edward L QUINN et Leroy HARDIN. Cyber security  
of nuclear instrumentation & control systems: overview of the iec standardization  
activities. *IFAC Proceedings Volumes*, 46(9):2156–2160, 2013.
- [128] Abel PINTO, Rita A RIBEIRO et Isabel L NUNES. Fuzzy approach for reducing  
subjectivity in estimating occupational accident severity. *Accident Analysis & Pre-  
vention*, 45:281–290, 2012.
- [129] David POLLARD. *A user’s guide to measure theoretic probability*, volume 8. Cam-  
bridge University Press, 2002.
- [130] Cameron-I.T. QUELCH, J.. Uncertainty representation and propagation in quantifie  
risk assessment using fuzzy sets. 1994.
- [131] Samira Abbasgholizadeh RAHIMI, Afshin JAMSHIDI, Daoud AIT-KADI et An-  
gel Ruiz BARTOLOME. Risk-based decision making framework for prioritizing pa-  
tients’ access to healthcare services by considering uncertainties. *In Industrial Engi-  
neering and Systems Management (IESM), 2015 International Conference on*, pages  
291–297. IEEE, 2015.
- [132] Samira Abbasgholizadeh RAHIMI, Afshin JAMSHIDI, Angel RUIZ et Daoud AIT-  
KADI. A new dynamic integrated framework for surgical patients’ prioritization  
considering risks and uncertainties. *Decision Support Systems*, 2016.
- [133] Christian RASPOTNIG, Peter KARPATI et Vikash KATTA. A combined process for  
elicitation and analysis of safety and security requirements. *In Enterprise, business-  
process and information systems modeling*, pages 347–361. Springer, 2012.
- [134] Michael ROTH et Peter LIGGESMEYER. Modeling and analysis of safety-critical cy-  
ber physical systems using state/event fault trees. *In SAFECOMP 2013-Workshop  
DECS (ERCIM/EWICS Workshop on Dependable Embedded and Cyber-physical  
Systems) of the 32nd International Conference on Computer Safety, Reliability and  
Security*, page NA, 2013.
- [135] Arpan ROY, Dong Seong KIM et Kishor S TRIVEDI. Attack countermeasure trees  
(act): towards unifying the constructs of attack and defense trees. *Security and  
Communication Networks*, 5(8):929–943, 2012.

- 
- [136] Rehan SADIQ, Elise SAINT-MARTIN et Yehuda KLEINER. Predicting risk of water quality failures in distribution networks under uncertainties using fault-tree analysis. *Urban Water Journal*, 5(4):287–304, 2008.
- [137] Avinash SAMVEDI, Vipul JAIN et Felix T.S. CHAN. Quantifying risks in a supply chain through integration of fuzzy ahp and fuzzy topsis. *International Journal of Production Research*, 51(8):2433–2442, 2013.
- [138] Todd SAX et Vlad ISAKOV. A case study for assessing uncertainty in local-scale regulatory air quality modeling applications. *Atmospheric Environment*, 37(25):3481–3489, 2003.
- [139] Telemetry & Remote SCADA Solutions SCHNEIDER ELECTRIC. Scada systems. Rapport technique, Schneider Electric, Ontario K2K 2A9, Canada, 2012.
- [140] Bruce SCHNEIER. Modeling security threats. *In Dr. Dobb's Journal*, 1998.
- [141] Timothy L SELNOW et Matthew SEEGER. Exploring the boundaries of crisis communication: The case of the 1997 red river valley flood. *Communication Studies*, 52(2):153–167, 2001.
- [142] Scott J SHACKELFORD, Andrew A PROIA, Brenton MARTELL et Amanda N CRAIG. Toward a global cybersecurity standard of care: Exploring the implications of the 2014 nist cybersecurity framework on shaping reasonable national and international cybersecurity practices. *Tex. Int'l LJ*, 50:305, 2015.
- [143] Glenn SHAFER *et al.*. *A mathematical theory of evidence*, volume 1. Princeton university press Princeton, 1976.
- [144] Arnold F SHAPIRO. Fuzzy random variables. *Insurance: Mathematics and Economics*, 44(2):307–314, 2009.
- [145] Jinsoo SHIN, Hanseong SON et Gyunyoung HEO. Cyber security risk evaluation of a nuclear i&c system using bayesian networks and event trees. *Nuclear Engineering and Technology*, 2016.
- [146] David H SLADE. Meteorology and atomic energy, 1968. Rapport technique, Environmental Science Services Administration, Silver Spring, Md. Air Resources Labs., 1968.
- [147] Philippe SMETS et Robert KENNES. The transferable belief model. *Artificial intelligence*, 66(2):191–234, 1994.
- [148] Jason Edwin STAMP et Philip LaRoche CAMPBELL. A classification scheme for risk assessment methods. Rapport technique, Sandia National Laboratories, 2004.
- [149] Stanley Smith STEVENS. Mathematics, measurement, and psychophysics. 1951.
- [150] Keith STOUFFER, Joe FALCO et Karen SCARFONE. Guide to industrial control systems (ics) security. *NIST special publication*, 800(82):16–16, 2011.



- [151] Teruo SUNAGA. Theory of an interval algebra and its application to numerical analysis. *Japan Journal of Industrial and Applied Mathematics*, 26(2-3):125–143, 1958.
- [152] J TADMOR et Y GUR. Analytical expressions for the vertical and lateral dispersion coefficients in atmospheric diffusion. *Atmospheric Environment (1967)*, 3(6):688–689, 1969.
- [153] D Bruce TURNER. *Workbook of atmospheric dispersion estimates: an introduction to dispersion modeling*. CRC press, 1994.
- [154] A.M VADEBY. Modeling of relative collision safety including driver characteristics. *Accident Analysis and Prevention*, 36(5):909 – 917, 2004.
- [155] Jan-Willem GM van der PAS, Vincent AWJ MARCHAU, Warren E WALKER, GP VAN WEE et SH VLASSENROOT. Isa implementation and uncertainty: A literature review and expert elicitation study. *Accident Analysis & Prevention*, 48:83–96, 2012.
- [156] Ying-Ming WANG et Taha M.S. ELHAG. A fuzzy group decision making approach for bridge risk assessment. *Computers and Industrial Engineering*, 53(1):137 – 148, 2007.
- [157] Joe WEISS. Industrial control system cyber security and the critical infrastructures. *INSIGHT*, 19(4):33–36, 2016.
- [158] Robb C WILCOX et Bilal M AYYUB. Uncertainty modeling of data and uncertainty propagation for risk studies. *In null*, page 184. IEEE, 2003.
- [159] Henk WM WITLOX. Overview of consequence modelling in the hazard assessment package phast. Rapport technique, American Meteorological Society, 2010.
- [160] Ronald YAGER, Mario FEDRIZZI et Janus KACPRZYK. Advances in the dempster-shafer theory of evidence. 1994.
- [161] B YUAN. Fuzzy sets and fuzzy logic: theory and applications, 1995.
- [162] Yang Chyuan YUAN, SY CHEN, DJ LEPOIRE et R ROTHMAN. *RISKIND-A Computer Program for Calculating Radiological Consequences and Health Risks from Transportation of Spent Nuclear Fuel*. Argonne National Laboratory, 1995.
- [163] Wang YUANHUI. Safety system engineering. *Tianjin: Tianjin University Publishing House*, 1999.
- [164] L.A. ZADEH. Fuzzy sets. *Information and Control*, 8(3):338 – 353, 1965.
- [165] Edmundas Kazimieras ZAVADSKAS, Zenonas TURSKIS et Jolanta TAMOŠAITIENE. Risk assessment of construction projects. *Journal of civil engineering and management*, 16(1):33–46, 2010.

- [166] Bonnie ZHU, Anthony JOSEPH et Shankar SASTRY. A taxonomy of cyber attacks on scada systems. *In Internet of things (iThings/CPSCoM), 2011 international conference on and 4th international conference on cyber, physical and social computing*, pages 380–388. IEEE, 2011.
- [167] H-J ZIMMERMANN. An application-oriented view of modeling uncertainty. *European Journal of operational research*, 122(2):190–198, 2000.
- [168] Enrico ZIO. Reliability engineering: Old problems and new challenges. *Reliability Engineering & System Safety*, 94(2):125–141, 2009.
- [169] Enrico ZIO et GE APOSTOLAKIS. Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories. *Reliability Engineering & System Safety*, 54(2-3):225–241, 1996.
- [170] Enrico ZIO et Terje AVEN. Industrial disasters: Extreme events, extremely rare. some reflections on the treatment of uncertainties in the assessment of the associated risks. *Process Safety and Environmental Protection*, 91(1):31–45, 2013.
- [171] Enrico ZIO et Nicola PEDRONI. Methods for representing uncertainty.
- [172] Enrico ZIO et Nicola PEDRONI. *Uncertainty characterization in risk analysis for decision-making practice*. FonCSI, 2012.
- [173] Miremadi S.G. ZONOUZ, S.A.. A fuzzy monte carlo simulation approach for fault tree analysis. 2006.



# Annexe A: Publications list

## H Journal articles



[4] ABDO, HOUSSEIN AND FLAUS, JEAN-MARIE. *Uncertainty quantification in dynamic system risk assessment: a new approach with randomness and fuzzy theory*. International Journal of Production Research **54**, 5862-5885 (2016).



[7] H. ABDO, J-M. FLAUS, AND F. MASSE. *Uncertainty quantification in risk assessment - Representation, propagation and treatment approaches: Application to atmospheric dispersion modeling*. Journal of Loss Prevention in the Process Industries.



[9] H. ABDO AND J-M. FLAUS AND F. MASSE. *A safety/security risk analysis approach of Industrial Control Systems: A cyber bowtie - combining new version of attack tree with bowtie analysis*. Computers & Security , - (2017).

## I International conferences



[5] H. ABDO, J-M. FLAUS, F.MASSE. *Fuzzy semi-quantitative approach for probability evaluation using Bow-Tie analysis*. Safety and Reliability Theory and Applications (2017).



[2] H. ABDO, AND J-M. FLAUS. *A mixed fuzzy probabilistic approach for risk assessment of dynamic systems*. In *15th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2015*, Ottawa, Canada (2015).



[3] H. ABDO, AND J-M. FLAUS. *Uncertainty quantification in bow-tie analysis: A mixed approach of fuzzy theory with Dempster-Shafer theory of evidence*. In *Risk, Reliability and Safety: Innovating Theory and Practice: Proceedings of ESREL*, Glasgow, Scotland (2016).



[6] H. ABDO, J-M. FLAUS AND F. MASSE. *Towards a better industrial risk analysis: A new approach that combines cyber security within safety*. In *Safety and Reliability Theory and Applications: Proceedings of ESREL*, 1080-1089, Portoroz, Slovenia (2017).

## J National conferences



[50] F. MASSE, H. ABDO AND J-M. FLAUS. *Vers une approche intégrant les exigences de cybersécurité à la maîtrise des risques d'accidents majeurs pour les ICPE*). In *12ème Congrès International Pluridisciplinaire en Qualité, Sûreté de fonctionnement et Développement durable*, Bourges, France (2017).

## K Under submission



[8] H. ABDO, J-M. FLAUS, F.MASSE. *Fuzzy semi-quantitative approach for probability evaluation using Bow-Tie analysis*. reliability engineering & system safety. under submission.

---