Uncertainty analysis and decision-aid: methodological, technical and managerial contributions to engineering and R&D studies

Alberto Pasanisi

MÉMOIRE

Présenté en vue de l'obtention de la

HABILITATION À DIRIGER DES RECHERCHES

par

Alberto Pasanisi

Docteur de l'École Nationale du Génie Rural, des Eaux et des Forêts

Analyse d'incertitudes et aide à la décision : contributions méthodologiques, techniques et managériales aux études d'ingénierie et de R&D

Uncertainty analysis and decision-aid: methodological, technical and managerial contributions to engineering and R&D studies

Soutenu le 22 mai 2014
Devant le jury composé par :

Prof. Jean-Michel Poggi	Université Paris Descartes	Président
Prof. Frédéric Y. Bois	Université de Technologie de Compiègne	Rapporteur
Prof. Anne Gégout-Petit	Université de Lorraine	Rapporteur
Prof. Olivier Roustant	École des Mines de St. Étienne	Rapporteur
Prof. Antoine Grall	Université de Technologie de Troyes	Examinateur
Prof. Éric Parent	AgroParisTech	Examinateur
Prof. Enrico Zio	École Centrale Paris	Examinateur

Acknowledgments

A alma do Mundo é alimentada pela felicidade das pessoas. Ou pela infelicidade, inveja, ciúme. Cumprir sua Lenda Pessoal é a única obrigação dos homens. Tudo é uma coisa só.
"E quando você quer alguma coisa, todo o Universo conspira para que você realize seu desejo".
Paulo Coelho, O Alquimista (1988)

The Soul of the World is nourished by people's happiness. And also by unhappiness, envy, and jealousy. To realize one's destiny is a person's only real obligation. All things are one.
"And, when you want something, all the universe conspires in helping you to achieve it".
Paulo Coelho, The Alchemist (translation by Alan R. Clarke, 1992)

A little story about this manuscript, to start with. Academic titles are not just lines on a curriculum vitae (as naive people can think) but always the ends of adventures, and I take the opportunity here to briefly remind the main steps of this exciting project.

This manuscript is the achievement of a work that actually took one year and a half. Although the idea germinated in my mind several years ago, this project actually started during a dinner at my home for celebrating my fortieth birthday in September 2012 ... 40 years: an ideal time for taking stock of the work done so far and, especially, for looking for new challenges.

I started that night discussing with Éric Parent who warmly encouraged me to undertake this work and gave me the first advices. Lucien Duckstein was also there. I remembered him the very last question he asked me at the end of my PhD defence in February 2004. "Alberto, when will you write another PhD thesis?" Well, a HDR thesis is not a PhD thesis... but nevertheless a thesis; so that, I was finally able to answer his question, more than eight year later. *Meglio tardi che mai*!

Following Éric's advices, I began collecting the first ideas and listing my papers, communications, reports and several contributions to be possibly put into evidence in a HDR.

Here comes the second major contributor to this project: Jean-Michel Poggi. I will always remember our fruitful discussion in a cold November morning during which I presented my first ideas and he gave me back a number of useful advices. His role in this adventure is simply crucial as he helped me to refine the project and to properly present it, together with a tailored academic curriculum vitae.

The third person with a special role in this work is Antoine Grall, who listened with interest to my ideas, starting to

take form, and definitely oriented me towards the *Université de Technologie de Compiègne* (UTC). I established the first contact with UTC at the beginning of 2013 and I completed the application file in Summer 2013. The application included a summary document, the writing of which was an extremely useful experience and greatly helped my in collecting the ideas for the writing of this HDR thesis. The Scientific Council of UTC examined and accepted my application in September 2013 ... and (some hundred hours of work after) here this manuscript is!

I would like to thank, first of all, Éric Parent, Jean-Michel Poggi and Antoine Grall (recalled here in the appearing order in the story of this adventure) for their help and their advices and for the time they dedicated me in spite of their busy schedules. *Merci, encore et encore*.

I warmly thank Anne Gégout-Petit, Frédéric Bois and Olivier Roustant for the honour they have done me by reviewing this manuscript as well as Enrico Zio for being part of the jury, together with the aforementioned persons.

I also take the opportunity to thank here the UTC for welcoming me and especially Julie Jarek (*École Doctorale de l'UTC*) who have been an extremely helpful and effective contact, as well as AgroParisTech for hosting the HDR defence.

I am particularly grateful to the colleagues that read this document before the submission and helped me to improve it with respect to its form and content: Merlin Keller, Bertrand Iooss, Nicolas Bousquet, Éric Parent (again).

Special thanks are addressed to Jacques Bernier, an example and a source of inspiration. I took the liberty to resume in this manuscript some of his papers that deserve more light.

Another story that I will not tell. I want here to especially thank all people who helped me to get rid of recent difficult moments and to let me understand a number of teachings I will always bring with me. I cannot remember here all people that helped me: many discussions with colleagues but also phone-calls with my brother Francesco and my parents, conversations with Fabienne, football games and drinks with Alessandro, Italian and Greek sunny days with my wife are also part (the good one) of this history. It is not worth writing anything else here

Co-authors. This manuscript puts into evidence a number of works carried in the last 15 years. Of course, and luckily, I did not work alone during this period. In this sense, this manuscript, is a sort of collective work. I want to make here a graphical tribute (cf. Figure 1) to my co-authors by representing their names into a so-called *word-cloud* (cf. also Chapter VIII, page 148) in which the size of the name is a function of the number of papers I wrote with. Besides the "statistical" curiosity, I take the opportunity to gratefully thank each one of them.

EDF colleagues. I had the chance to work in an extremely stimulating environment, at the Dept. *Management des Risques Industriels* (MRI) of EDF R&D. I take the opportunity here to make a tribute especially to some of the numerous colleagues who work or worked with me within the *Incertitudes* project: Anne Dutfoy (OpenTURNS leader and passionate researcher on probabilistic dependence), Bertrand Iooss (the *señor researcher* who every one wishes to have in his project team), Mathieu Couplet (amazing how many good ideas he can have simultaneously), the "Bayesian fellows" Merlin Keller, Nicolas Bousquet and Sophie Ancelet (currently at IRSN), Anne-Laure Popelin (who did and will continue doing a great job, I am sure of it), Michaël Baudin (whose arrival in late 2011 completed the team's skills), the probabilistic mechanics experts Marc Berveiller and Géraud Blatman, the PhD students Guillaume Damblin (I will find the time to follow your PhD, until its end), Shuai Fu and Richard Fischer as well as the OpenTURNS crucial contributors Ivan Dutka-Malen (until 2012) and Renaud Barate. I wish you all the best for the future project (particularly to Anne-Laure). It was simply great to work with you!

Many thanks also to Emmanuel Remy (ok, you are not Bayesian, but it does not matter) and Frédéric Hostyn (endless source of engineering common sense and great football expert), as well as Françoise Talbot, Françoise Massot and Isabelle

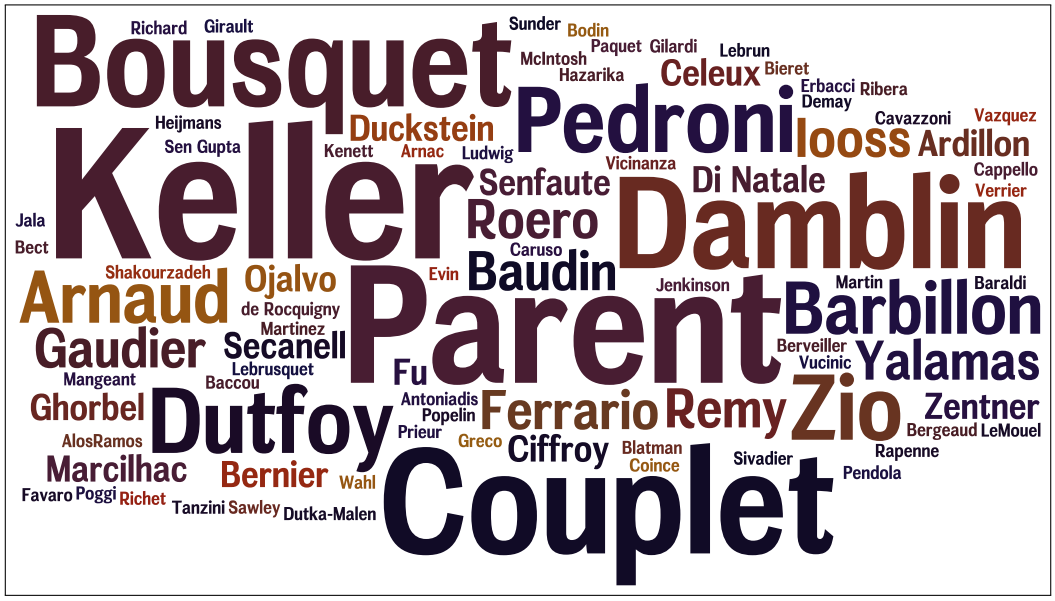


Figure 1 – Word-cloud of the list of the coauthors of the papers of mine cited in this document, built thanks to the Wordle™ tool (www.wordle.net). The greater the size of the name, the higher the number of papers I coauthored with.

Périlhous, for their effective support. I also remember here the interesting and pleasing coffee talks in Chatou and Clamart with Roman Sueur, Jean Anglès, Marie Gallois, Mohamed Hibti, Kateryna Dorogan, Mathieu Anderhalt, Hélène Bickert, David Albrecht, Jean-Baptiste Desmottes, Dominique Marion, François Beaudouin, Fanny Douard, Denis Lachêne ... I certainly forget somebody.

I take the opportunity to thank the MRI managers who encouraged me in this adventure, especially François Billy, Stéphane Fortier , Laurent Billet and Sylvie Jahan (to whom I am particularly grateful), as well as Stéphane Andrieux and Claire Waast-Richard (members of the Direction Board of EDF R&D) and Ange Caruso (responsible of the Information Technology Program).

It is worth reminding here the activity of the *Réseau Statistique et Probabilité* of EDF R&D in which I have been honoured to be member of the Board: many thanks in particular to Philippe Chabault, Silvia Fagnoni and Christian Derquenne for the wonderful job we did in these last years.

I wish to address here many thanks to other colleagues, from other Departments of EDF R&D and Engineering, with whom I had the opportunity to cooperate and in particular Joseph Ojalvo, Mathieu Schumann, Vincent Lefèbvre, Nicole Goutal, Pietro Bernardara, Philippe Ciffroy, Gloria Senfaute, Christian Chauliac.

Scientific societies. The *Société Française de Statistique* (SFdS) greatly accompanied my activities during these last years. I warmly thank in particular the past-presidents Avner Bar-Hen (who gave me the opportunity to create in 2009 the thematic group *Fiabilité et Incertitudes*) and Jean-Michel Poggi and the current president Anne Gégout-Petit for the missions they give me within this wonderful Society, in which I proudly serve as member of the Council. And many special thanks to Servane Bianciardi for her effective help in the organisation of events.

It is worth reminding the fruitful cooperation established between SFdS and the *European Network For Industry and Business Statistics* (ENBIS) for which I am grateful in particular to the recent ENBIS presidents and past-presidents as well as to Ron Kenett with whom (and Jean-Michel Poggi) I had the pleasure and the honour to share the organization of

the joint 2014 SFdS-ENBIS Spring Meeting and to co-edit a related special issue of an international journal.

The *Institut pour la Maîtrise des Risques* has also been a powerful vector for spreading the works and the activities of my project. I have been honoured to co-chair (with my comrade Nicolas Fischer) the thematic group *Incertitudes et Industrie*. I take the opportunity to thank, in particular, André Lannoy for the interesting discussions and his kindness.

Colleagues outside EDF. Most of the activities I carried during these last years were made in cooperation with academic and industrial partners. In particular these works took benefit from a long-standing fruitful cooperation with EADS (nowadays, *Airbus Group*) and CEA, for which I particularly thank Fabien Mangeant and Vincent Bergeaud.

The OPUS project (I had the honour to lead) was a fantastic technical and human experience. I thank again the colleagues with whom I shared that great adventure and in particular Anestis Antoniadis (what an honour for me to co-edit a special issue of a scientific journal with him!), Jayant Sen Gupta, Jean-Marc Martinez, Emmanuel Vazquez, Laurent Lebrusquet, Julien Bect, Gilles Fleury, Josselin Garnier, Michaël Baudin (again), Christophe Prud'homme, Érik Herbin. The journey continues with the CHORUS project, with some new team-mates. *Bon vent !*

The *OpenTURNS* consortium is a pillar of the activity of EDF R&D about computer experiments. It was really a pleasure to work with Fabien Mangeant (again), Pierre Benjamin, Régis Lebrun, Thierry Yalamas, Julien Schueller ... and welcome to our new partners of IMACS!

It is worth noting the fruitful cooperation with the *GdR MASCOT NUM* with whom I had the opportunity to co-organize some scientific events these last years. I remember here the fruitful and inspiring discussions with Fabrice Gamboa and Luc Pronzato, as well as the special issue on sensitivity analysis of the *Journal de la SFdS* with Clémentine Prieur and François Wahl.

Many thanks to the academic partners with whom I had the opportunity to cooperate: Éric Parent (again), great teacher and friend and and attentive PhD supervisor, Gilles Celeux (who gave me the great honour to serve in the Program Committee of the *Journées de Statistique*) in 2013, Enrico Zio (who always succeeds to find some time for his colleagues, in spite of his huge number of missions) and his team of the Politecnico di Milano and École Centrale Paris / SUPÉLEC (in particular, Nicola Pedroni and Piero Baraldi), Mitra Fouladirad for the fruitful cooperation in the framework of the course "Industrial feedback and safety" at UTT.

Family. My final thoughts go to my family who encouraged and supported me: my parents without whom nothing of all that would have been possible, my grandparents, the one here on this earth and the other ones left away but still close to me, who compassionately look at what I do, even if I am far, my brother (who also helped me for the writing of the first chapter and provided me some wonderful pictures of our joint Italian and Brazilian early works).

And, dulcis in fundo, my warmest thanks go to my wife, who patiently freed up time (and space) for the week-ends I spent writing this manuscript and my daughters (my two little *principesse*) for their many lovely attentions and for putting at my disposal their wonderful room, where I wrote part of this manuscript.

Sandrine, Chiara, Diane, I do not know if one day you will read it ... but this work is dedicated to you!

Paris, XII Arrondissement, March 2014

Contents

I Foreword **1**

1 About engineering . . . 1
2 Engineers, data and statistical modelling . . . 2
3 Engineers and computer models . . . 3
4 Recommending Decisions . . . 4
5 Structure of the document . . . 5
6 Biographical summary . . . 7

II On the practical use of two discrete lifetime models **13**

1 Rationale and basic definitions . . . 14
2 Inverse Pólya model . . . 16
3 Modelling ageing by means of IPD: a major limitation . . . 17
4 The Weibull-1 model . . . 19
5 Modelling ageing by means of the Weibull-1 distribution . . . 20
6 Ageing: two numerical examples . . . 21
7 Approximating Weibull-1 by continuous Weibull model . . . 22
7.1 Empirical study . . . 23
7.2 Inference from actual feedback data . . . 25
8 Discussion . . . 26

III Markov chain modelling of industrial systems deterioration **29**

1 Introduction . . . 30
1.1 Different data structures . . . 31
2 Bayesian estimation of transition probabilities . . . 32
2.1 Complete sequence problem . . . 33
2.2 Incomplete sequence problem, ignorable DCM . . . 33
2.3 Incomplete sequence problem, non-ignorable DCM . . . 35
2.4 Aggregate data problem . . . 36

2.5 A four-dimensional simulated case study . . . 38
3 Accelerating the MH algorithm using adaptive approaches . . . 38
4 Numerical experiments . . . 42
5 Discussion . . . 44
5.1 Main ideas and results . . . 44
5.2 Directions of further research . . . 45
6 An application to turbine cracks propagation . . . 47
7 Modelling water meters deterioration . . . 49
7.1 Rationale and modelling assumptions . . . 49
7.2 Different sources of information . . . 51
7.3 Variants of the main model . . . 52
7.4 Search for explanatory variables . . . 52

IV Uncertainties in numerical simulation **59**
1 Computer simulation: opportunities and issues . . . 61
2 Advanced numerical simulation at EDF R&D . . . 64
3 Different kinds of problems . . . 66
4 The pioneering works of Jacques Bernier . . . 70
5 The need for specific methods . . . 74
6 The need for specific tools . . . 78
7 A very active research community . . . 80
7.1 Working groups and research consortia . . . 80
7.2 The ANR OPUS project . . . 81
7.2.1 Context and background . . . 81
7.2.2 Project Structure . . . 82
7.2.3 Some achievements . . . 84
7.3 Other collaborative funded projects . . . 87

V The extra-probabilistic temptation **89**
1 A very early study: fuzzy rule-based metamodel of river pollution . . . 90
1.1 The problem . . . 91
1.2 The fuzzy model . . . 93
1.2.1 Learning . . . 93
1.2.2 Fuzzy rule-based computing . . . 94
1.2.3 Numerical example . . . 95
1.2.4 Discussion . . . 95
2 Propagation of "hybrid" uncertainties . . . 95
2.1 Motivations . . . 95
2.2 Hybrid possibilistic-probabilistic framework . . . 98
2.3 Hybrid possibilistic-probabilistic "hierarchical" framework . . . 100
2.4 Further works and comments . . . 103

VI Bayesian point estimation of probabilistic risk criteria **105**
1 Coming back to the "quantity of interest" . . . 106
2 Decision theory and Bayesian estimation of the quantities of interest . . . 108
2.1 Cost functions and Bayesian estimation . . . 108
2.2 Reinterpreting the common predictive approach . . . 111
3 A simple numerical example . . . 112
3.1 Probabilistic model . . . 113
3.2 Synthetic dataset . . . 113
4 Comments . . . 117
5 Further considerations . . . 120
5.1 A more general result . . . 120
5.2 Descriptive approaches . . . 122
6 Extension to a functional criterion: the fragility curve . . . 122
6.1 Accounting for estimation uncertainties . . . 123
6.2 Bayesian estimation of fragility curves . . . 125
6.3 On-going work on seismic hazard assessment . . . 127

VII An application of multi-criteria decision-aid **131**
1 Decision making: a difficult process . . . 132
1.1 Methods based on a single synthesising criterion . . . 133
1.2 Methods based on the outranking . . . 135
2 Energy retrofitting of existing building: a specific and complex problem . . . 139
3 REFLEX Overview . . . 141
3.1 Retrofitting solutions generation . . . 141
3.2 Retrofitting solutions evaluation . . . 143
3.3 Retrofitting solutions' filtering and ranking . . . 144
4 Discussion and perspectives . . . 146

VIII Afterword **147**
1 Concluding remarks . . . 147
2 On-going works and perspectives . . . 149

Appendix A: Proofs and additional material **153**

Appendix B: Curriculum vitae **161**

References **173**

Chapter I

Foreword

Al primo parlar che si fece di peste, don Ferrante fu uno de' più risoluti a negarla. [...] "In rerum natura," diceva, "non ci son che due generi di cose: sostanze e accidenti; e se io provo che il contagio non può esser nè l'uno nè l'altro, avrò provato che non esiste, che è una chimera." [...]
La scienza è scienza; solo bisogna saperla adoprare." [...]
Su questi bei fondamenti, non prese nessuna precauzione contro la peste; gli s'attaccò; andò a letto, a morire, come un eroe di Metastasio, prendendosela con le stelle.
Alessandro Manzoni, I promessi sposi (1842)

On the very first whisper of pestilence, Don Ferrante was one of the most resolute, and ever afterwards one of the most persevering, in denying it. [...] "In rerum natura" he used to say, "there are but two species of things, substances and accidents; and if I prove that the contagion cannot be either one or the other, I shall have proved that it does not exist, that it is a mere chimera. [...]
Science is science; only we must know how to employ it." [...]
On these grounds, he used no precautions against the pestilence; took it, went to bed, and went to die, like one of Metastasio's heroes, quarrelling with the stars.
Alessandro Manzoni, The betrothed (translated by Burns Publishers, 1844)

1 About engineering

A (not so original) way for introducing some reflections about a specific topic is to start from its usual definition.

According to the on-line Collins Dictionary (cf. http://www.collinsdictionary.com), engineering is "*the profession of applying scientific principles to the design, construction, and maintenance of engines, cars, machines, etc. (mechanical engineering), buildings, bridges, roads, etc. (civil engineering), electrical machines and communication systems (electrical engineering), chemical plant and machinery (chemical engineering), or aircraft (aeronautical engineering)*".

This definition sounds perfectly good: it is essentially focused on the technical objects the engineers cope with. On the other hand, although the first part of the definition, evoking some "*scientific principles*" may seems vague, it also reminds that methodological tools used by engineers are extremely varied.

Here comes another (and more challenging) definition of engineers. Some years ago, few months before obtaining my high school diploma, I took part in a day-long orientation workshop intended to young students for guiding the choice of their university studies. Among all conferences, I particularly remember the passage of a professor of the Faculty of Engineering of the University of Napoli Federico II (unfortunately I do not remember his name), in which he defined an engineer as someone who "*has learnt to learn*". This phrase, maybe not fully original, and surely not restricted to engineers

only, impressed me (and actually played a role in my personal choice to be an engineer) but I did not understand, in that present moment, how it was well suited to me.

Engineering studies are hard and long and provide a very wide set of mathematical and methodological tools, mostly based on the application of physics to technical systems.

In spite of the quality of the tools provided by classical engineering education, due to the extremely wide variety of problems engineers must cope with in practice, in most cases, they need to complete their own toolbox with other methods, traditionally coming from the domain of applied mathematics.

Actually, one of the main activity of engineers is to provide forecasts of the behaviour of systems (which can be very complex) by means of appropriate predictive models. Most of the models used by engineers are deterministic, but often the problem to be solved contains several sources of uncertainties, the analyst must cope with:

- because the behaviour of the system under investigation is intrinsically stochastic and the question posed is to make predictions under the base of feedback data and/or expertise,
- because the system under investigation is fairly well described by a deterministic model (a computer code), the inputs of which are tainted with uncertainties.

In the first case, the engineer must use a statistical model, and most often build his/her own model, suited to the problem to be solved. In the second case, he/she must couple a probabilistic and a deterministic models. In both situations, he/she must cope with uncertainties and take them into account in the provided results.

Another issue in modern engineering is that providing forecasts is not enough. What is asked to engineers is essentially *solving problems*: analysing the results and making decisions (or more modestly providing recommendations) are a fundamental part of engineering work.

Throughout this document, taking as motivations some technical questions I had to answer during my career, I present some methods, non conventional in the classical engineering background, which ideally complete it.

In particular, I insist on the issues sketched in the next three sections.

2 Engineers, data and statistical modelling

Reliability assessment of industrial components is a key issue in engineering. Assessing the lifetime of an equipment is a capital input of safety and assets management studies as well as of the definition of maintenance policies and spare parts stock.

Even though, in theory, physical phenomena leading to the loss of operability can always be imagined and possibly modelled, it is common to admit that failures are random events. The problem posed here is thus to imagine a (more or less) complex statistical model, infer its parameters, based on available information (data and/or expertise), and finally establish predictions.

In other problems, when a system passes during its lifetime across several states corresponding to more or less deteriorated operating conditions, one can be interested also in modelling the deterioration dynamics: what is the probability for the system to be in a given state, given its age? What is the period the system spends in a given state or in a given set of states?

It is to notice that in many other technical fields, engineers cope with random variables: measurement errors, natural phenomena (rain and snow precipitations, floods, earthquakes) etc.

The main resource for solving these problems is constituted by data. In the case of industrial reliability, data come from feedback: for instance, if one must assess the probability distribution function of the random variable "lifetime" of a system, it is intuitive to imagine fitting it on a random sample of lifetimes, which is easy to obtain if one knows exactly all the commissioning and the failure dates.

When assessing deterioration models, ideally, data should be repeated observations of states (or given features of interest, related to states) for the same systems at different dates. But unfortunately, data the engineers cope with are often much less informative.

In real life, data can be scarce, censored, incomplete and, on the one hand, a preliminary analysis taking into account the technical information available about both the studied system and the data collection procedure is necessary, on the other hand, the statistical model chosen must be adapted to data. Actually, models must fit to data and not the opposite! That adds an additional difficulty: reality is much more complex than common statistical models and here arises the need for more advanced mathematical methods for modelling and inference.

Although data are almost always far from being ideal, normally another source of information, often neglected or not fully exploited, is available: expertise. Engineers, technicians, operators may have a deep and valuable knowledge of the behaviour of a system; for instance, based on his/her own experience, one can give an interval for the lifetime of a component or its mean. Taking into account expertise is traditionally the main motivation for the use of Bayesian methods in industry (of course, there are many others and some of them will be highlighted in the remainder of this manuscript!).

3 Engineers and computer models

Nowadays, computer modelling is probably the most powerful tool in engineering. Most of the work of engineers (in particular young engineers), consists, very roughly speaking, in running numerical simulations. The quite recent availability of more and more computational power gave to analysts the possibility to model more and more accurately, more and more complex systems. I started working in the 90's and I belong to the first generation of engineers who have always had one (or more) "personal" computers on their desk. Yet, as undergraduate student, I have had classes of technical drawing using T-squares, pencils and technical pens (and actually I learnt AutoCad® by myself a couple of years later). I have also have been taught (more or less learned) to use tables or graphical methods (Figure I.1), like the funicular polygon, the Mohr circles, or the Bergeron-Schnyder method for water hammer calculation. One can argue that it is just a matter of tools for solving a problem. The added value is in the equations to be solved and not in the tools to be used, exactly like using a LaTeX editor instead of another ... but the question is much deeper than that.

When calculations were long and tedious, engineers were naturally pushed to be parsimonious. A very long phase of planning and careful choice of the input parameters always preceded the calculation phase.

Modern computer simulation may have a dangerous side. In some situations, the actual physical knowledge of the system under investigation could not be essential to run an experiment, once the model has been implemented: hence, the engineer risks becoming a simple operator just modifying input files and storing output files. Of course, good engineers plan calculation today too and the image of qualified workers using computer codes as machine-guns is somehow caricatural, but the risk exists. I will get back to some issues concerning computer simulation in the beginning of Chapter IV.

In addition, the model itself and/or its inputs could be tainted with uncertainties of different natures, which must necessarily be taken into account. Hence, the analyst has to assess the uncertainty of output results. This task is quite tricky, especially by a computing viewpoint, as it is easy to imagine that, for doing that, a great number of runs of a possibly costly computer code is necessary. For capably carrying this works, analysts need (i) a proper framework for

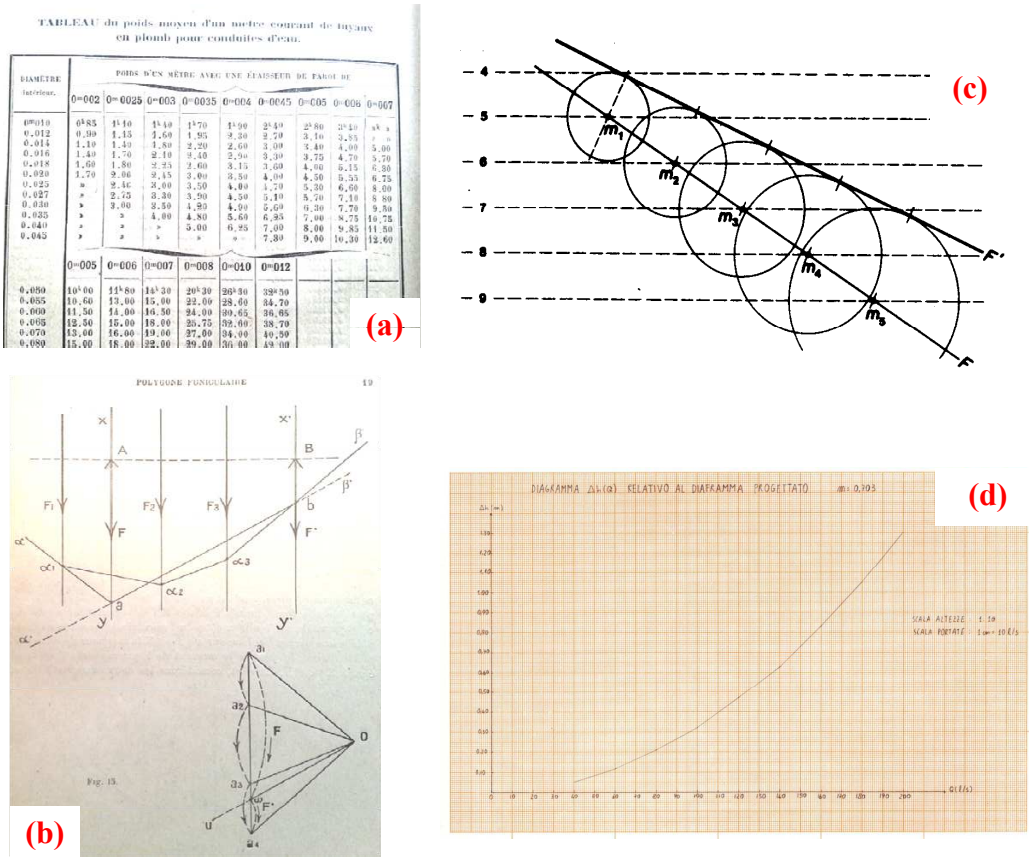


Figure I.1 – Some examples of older ways for effectively performing computations: (a) table for evaluating the weight of 1 meter of lead-pipe as a function of its internal diameter and thickness (from [Duffau 1903]), (b) resolution of a problem of statics by mean of the funicular polygon technique (from [Bayle 1946]), (c) determining fronts of sea waves in shallow water under refraction, by means of Huyghens' circles technique (from [Benassai 1973]), (d) head losses vs. discharge diagram for a flow-rate restrictor (Hydraulics practical of the author, academic year 1993-1994).

posing the problem and planning calculations, (ii) proper and adapted mathematical methods to reduce the number of calculations to be run, (iii) proper software tools for effectively coupling deterministic and probabilistic models and (iv) high performance computers for reducing the overall computing time.

Another additional and actually important question in computer simulation is the need for effective tools for interpreting, resuming and visualising (possibly uncertain) results. That is actually a wide-open field of investigation at the frontier between mathematics, computer science and visual perception.

4 Recommending Decisions

After predictive models (deterministic or probabilistic) have been proposed, assessed and run, an even more complicated question may arise: how using these predictions in order to decide, or (more often) to provide arguments for a rational decision? In engineering (and actually in everyday life) deciding is a difficult problem, basically because of two reasons.

The first reason is the presence of uncertainties. We live in an uncertain world and, even if one can imagine to reduce some kinds of uncertainties, the best he/she can do for providing useful inputs to a decision problem is to properly quantify

all sources of uncertainties tainting the outcomes of each decision. Decision theory under uncertainty is a wide topic; in this manuscript, I mostly focus on statistical decision theory (and make use of), which is indeed also a fundamental ingredient of the Bayesian setting. In particular, I deal with the following problem rooted in the domain of industrial safety: how to build point estimators of risk indicators (like quantiles or failure probabilities) in a decisional context?

The second reason is that the decision must be taken in presence of multiple and antagonistic criteria. For instance, a decision maker often faces situations where alternatives having high initial costs have also, in return, low running costs. The problem becomes even more complicated when non-monetary criteria come into play: how to arbitrate between costs, CO_2 emissions and user's comfort? Roughly speaking, two different paths for solving this problem can be taken.

The first one is to reduce the multi-criteria problem to a single-criterion one, by building a scalar function taking as input arguments the different criteria and returning the overall utility (or cost) of every action.

The second one is to use methods based on the aggregation of outranking relations between couples of alternatives. An alternative outranks another if it is *significantly* better than its opponent with respect to *enough* criteria (notice that behind these words different mathematical methods for performing comparisons are defined). Then, outranking relations are used to build a final ranking, which can possibly present situations of equivalence or incomparability between alternatives. That is the well known family of the ELECTRE methods.

In both situations, it must be clear that the goal of the engineer is twofold: it is not only providing the *best* solutions to a problem but also (and this is much more important) clearly present, generally to decision-makers not familiar with underlying mathematical tools, the reasons why they are the best. And here comes again the issue of presenting methods and results. Understandability and clearly stated hypotheses are key factors if one wants decision-makers actually follow engineers recommendations.

5 Structure of the document

Besides this introductory chapter and the final one, summarizing some conclusions and perspectives, the main body of the document is structured into six chapters (numbered from II to VII).

- Chapter II deals with statistical lifetime analysis of industrial components and in particular with problems concerning systems that do not operate continuously but rather "on demand", for which, in theory, the use of discrete probabilistic models is recommended.

- Chapter III is concerned with the statistical modelling, by means of Markov chains, of the deterioration of pieces of equipment when, as it is often the case in the industrial practice, available data are scarce and incomplete.

- Chapter IV relates to my main activity of these recent years, namely the technical management of works and projects on the quantitative assessment of uncertainties in computer simulation. It is also a sort of introduction for the two following chapters which are rather concerned with more technical and scientific activities I carried in this domain.

- Chapter V sketches some works (of different natures) carried in the framework of computer simulation by means of "extra-probabilistic" mathematical tools, i.e. based on uncertainty analysis frameworks alternative to the (classical) probability theory. In particular, the main tools used are fuzzy sets and possibility distributions.

- Chapter VI describes a methodology, firmly rooted in the Bayesian theory, for building point estimators of risk criteria (e.g. failure probability) by means of computer experiments, in presence of uncertainties tainting the inputs

of the model as well as their probability distributions. This methodology is compared to methods (popular within the engineering community) based on the so-called *predictive* distribution.

- Chapter VII, finally, gives some details about works applying multi-criteria decision-aid methods and tools to the domain of the energy efficiency of residential buildings, and in particular to the problem of energy retrofitting.

Each chapter is introduced by a short introductory section, entitled "Reading notes", which gives details on (i) the technical context of the works (why, when and where these activities have been carried), (ii) the contributions I provided to the technical problem and (iii) the already published material from which the chapter (or parts of it) is inspired or excerpted. Even in the case where most of the chapter consists in excerpted text, I have preferred incorporating the text in the main body of this manuscript rather than "appending" the paper "as is" in its original format, for sake of clarity and readability (avoiding different formats of texts and references and incoherent notations). That gave me also the opportunity to make little changes to the original texts when needed. Of course, the original papers and their co-authors are properly mentioned in the "Reading notes" sections (subsections entitled "Structure of the chapter and credits").

The order of chapters is not chronological. On the one hand, because of my "double life" of engineer and researcher, some works have been carried (sometimes discontinuously) over quite long periods during which other and very different works have been also carried and finalized. As an example, more than eight years passed between the very first formalization of the algorithm for the estimation of Markov transition matrices described in Chapter III [Pasanisi 2004a] and its publication in a scientific journal [Pasanisi *et al.* 2012a].

On the other hand, the choice of the order of the presented topics allows gradually introducing some concepts which will be used in the following chapters (e.g. the concepts of reliability and quantity of interest). The proposed order, hopefully, makes the reading easier. Works and contributions are not always presented with the same level of detail, for sake of brevity.

Two appendices complete the document: one is concerned with some additional material and proofs I decided to put at the end to avoid burdening the text, the other one is my curriculum vitae.

The document has numerous references. A list of communications and publications I am author or co-author of is also reproduced in Appendix, at the end of my Curriculum vitae (page 167).

Main message. The main message delivered by this document, which is also the connecting thread of my technical and scientific activity is that advanced mathematical methods and tools, particularly from the domain of probability and statistics, are necessary to solve engineering problems. These methods, most of engineers are not fully familiar with, are becoming more and more essential in the industrial practice.

In spite of the complexity of some methods described in the remainder, I wrote this document following an engineer's viewpoint, insisting on the technical problems and on the motivations for the use of sophisticated methods and tools.

But the goal of this document is also to summarize and highlight my own works and my contributions to these different technical fields (naturally humble, I definitely enjoyed very much to write a document focused, in some way, on myself). For every class of problems, my personal contributions are put into evidence. These contributions are of different natures:

- methodological: adapting, enhancing, or criticizing the use of particular methods in particular contexts,
- technical: solving particular engineering problems,
- managerial: organizing and driving research activities and projects.

Nevertheless, in spite of the diversity of methods, tools, fields of application and personal contributions given, all of these works stay coherently in the same mainline: enhancing engineering studies by means of advanced mathematical tools for coping with uncertainties and recommending decisions.

About introductory quotes. The texts of chapters are preceded by quotes, coming from books and songs which I particularly appreciated. I am an avid reader and many books (novels, stories or essays), which are seemingly not concerned at all with my technical and research domain, have definitely been a source of inspiration. I decided to reproduce the quoted text in its original language as a form of tribute to the authors. When needed, an English translation is provided.

6 Biographical summary

It is interesting, at the end of this introductory chapter, to give some biographical elements summarizing the evolution of my studies and my professional career. Two features are particular relevant in my opinion. First, I have always worked in the context of advanced engineering studies, where the borderline between engineering, R&D and academic research is often thin. Second, my studies and my professional activities are very varied from more than one perspective: geographical places, cultural environments, technical and scientific disciplines.

Cf. Curriculum vitae (page 161) for a more classical (and complete) presentation of my career.

Early period (before 1998). A seemingly surprising element concerning my education is that my high school studies have mainly been literary. It is a peculiarity of the Italian school system: a significant number of university students in scientific disciplines, come actually from literary high school studies. As a young student, I had the possibility to discover and deeply study subjects as Latin, ancient Greek, philosophy, Italian and classical literature. That gave me a great curiosity for human sciences (and human beings) and even if the very first weeks of my engineering studies were a little more difficult for me than for colleagues coming from scientific studies, I do not regret at all my classical education. And I have been particularly happy (and at ease) when, in 2004, I was asked to give a talk about Aristotle and the subjectivity in science in front of an audience mainly composed of statisticians in the middle of French Alps [Pasanisi 2004c].

I received my MSci in Civil Engineering (specialty: hydraulics) in 1997 from the University of Naples Federico II[1] with a final dissertation ("tesi di laurea") on the cost-effectiveness of artificial nourishment of tourist beaches suffering from erosion. The main subjects I dealt with, during my university studies have especially been hydraulic, environmental and coastal engineering.

I also had the opportunity (when I was still a student) to realize my very first engineering studies in these domains (wave motion studies and planning of marinas and coastal protection works).

After a one-year break due to military service (during which I made rather unusual working experiences as the reception and redeployment of deserters and jailed soldiers), my career actually started in Summer 1998.

Italian period (1998-2000). Between June 1998 and September 2000, I worked on my own-account as consultant engineer in the domain of hydraulic and coastal engineering. I realized a number of studies concerned with the planning of hydraulic works and especially the study of the wave climate in particular areas of interest by means of historical data (offshore wave climate) and computer simulation (onshore wave climate).

[1] Founded in 1224 by the Emperor Frederick II, it is the world's oldest state-supported and secular university, i.e. specifically intended to the training of administrative civil servants, judges and lawyers.

A particularly significant work has been the development of a software tool (named Tiresia[2]) for the study of beach erosion phenomena. In spite of the simplicity of the underlying mathematical assumptions (the so-called "one-line model", cf. [Koutitas 1988, Abbott & Price 2005] for more details about coastal modelling) this light software, written in Quick-Basic in cooperation with my brother Francesco, has proved effective and has been especially deployed (in late 1998) at the Environmental Agency of the Brazilian State of Espírito Santo, in the framework of an Italian-Brazilian cooperation program.

From February 1999 to September 2000, I mainly worked for the Department of Civil Engineering of the University of Naples 2. I was involved in a vast engineering and R&D project (supported by European Union), named "Realization of an integrated system for the control of coastal areas pollution phenomena near river mouths". In particular, I was in charge of the development of finite difference models of pollutants dispersion in rivers and estuaries. The working program was ambitious, as it aimed at the development of an integrated modelling environment coupling 1-D (river) and 2-D (estuary) shallow-water hydraulics and advection-dispersion of reactive and non-reactive pollutants in liquid bodies Cf. [De Bonis *et al.* 2002] for more details about hydraulic modelling issues in this specific context. Some models I developed were also used as case-studies of the European project ALICE-QFView (1998-2001), aiming at the effective management and visualization of computational fluid dynamics data [Vucinic *et al.* 2000, Vucinic 2007].

A more scientific and prospective work-package of this technical program was concerned with the use of meta-models based on fuzzy rules for fast predictions of pollution phenomena (namely, the arrival of the pollution front, due to an accidental release, to a given location). This activity, which was much more a R&D than an engineering work, allows me to make my very first steps in the field of research and establishing a cooperation with the École Nationale du Génie Rural, des Eaux et des Forêts (ENGREF). More details on this activities are given in Chapter V (Section 1).

Figure I.2 shows some example of works, essentially concerned with hydraulic and environmental simulation, carried between 1998 and 2000.

In September 2000, I joined the aerospace Italian company Alenia (Production Engineering Department, Head Office). For a couple of months I dealt with the problem of optimizing sub-contractors technical procedures, especially the use of raw materials for avoiding excessive manufacturing waste. In spite of the interest of the job and the comfort of a more "linear" career path, I decided to move to Paris in late 2000 for starting my PhD in the framework of a cooperation between ENGREF and the French water distribution company Génerale des Eaux.

PhD period (2001-2004). I made my PhD (defended in February 2004) within a CIFRE[3] doctoral program. The technical problem under investigation was the optimization of the renewal procedures of domestic water meters (which tend to underestimate more and more the customer's consumption when getting older). I developed and deployed in planning tools several algorithms (based on Bayesian analysis) to forecast the efficiency of in-service meters and to estimate the unaccounted-for water. For more details about the work achieved during this period, cf. Chapter III, Section 7.

Actually, I made my PhD in a very technical environment (Networks, Metering, Investments Department, Head Office of the Génerale des Eaux) in close connection with the fieldwork and, at the same time, in an applied research academic laboratory.

The experience gained during this period is priceless and helped me in the following years. It allowed me to deeply learn statistical modelling, and in particular Bayesian statistics, under the supervising of my mentor Éric Parent. These years left an indelible mark on my way of working: on the one hand, great commitment to my colleagues and to myself,

[2] In Greek mythology, Tiresias was a clairvoyant and a prophet of Apollo. He intervenes in a number of tragedies concerned with the city of Thebes (in particular, Sophocle's *Oedipus the King* and *Antigone*).

[3] CIFRE, acronym of "Convention Industrielle de Formation par la Recherche" (Research-based Industrial Training Convention) is a French program which (since the 80's) partially funds PhD works carried within an industrial context.

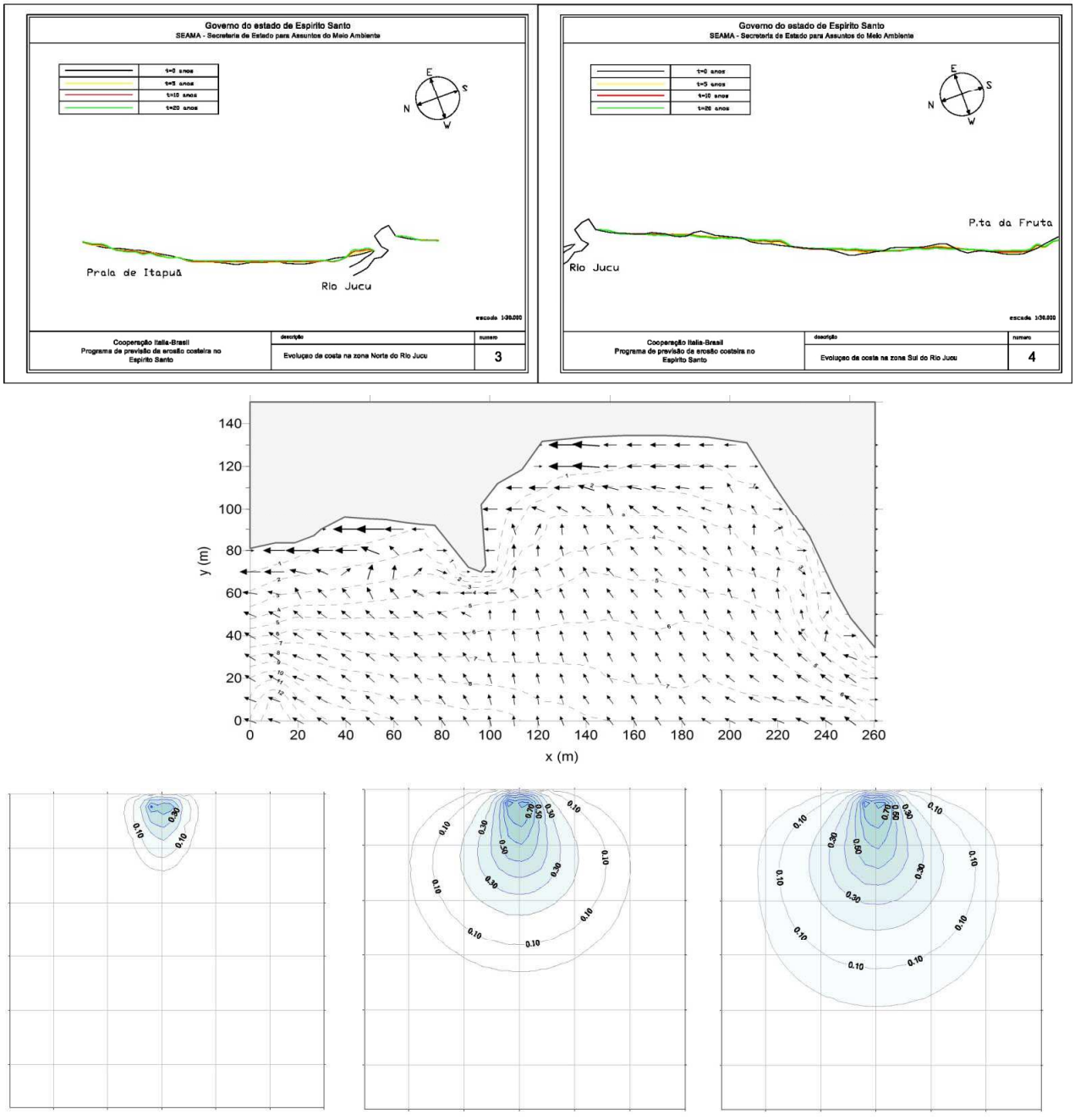


Figure I.2 – Images summarizing some works carried in the period 1998-2000. Top: exemplary study of the evolution of the shoreline in the area of the Rio Jucu estuary (Espírito Santo, Brazil) by means of the Tiresia software. Center: simulation of onshore hydrodynamic field (coast of Vico Equense, Italy). Down: Dispersion of a pollutant in an estuary zone (the river mouth is located in the center of the shoreline) in a simplified 2D domain under the effect of a symmetrical hydrodynamic field.
Remark. These figures are provided for exemplary purposes only and results (related to very particular configurations of parameters) shall not be extrapolated to draw any general conclusion.

scientific rigour, look for innovative solutions but also, on the other hand, research of clarity and brevity in both written and oral presentations, adaptation of both form and content to the technical skills of the audience, ability to formalize different technical problems (often ill-posed) in mathematical terms and reformulate them[4].

EDF R&D first period (2004-2008). After a short passage in an engineering consulting firm in Paris, I joined EDF R&D in June 2004 as research engineer at the Energy in Buildings and Territories Department (ENERBAT) in Southern Paris Area. The particular activity of ENERBAT I have been concerned with is focused on energy efficiency of buildings (envelope, thermal systems, user's behaviour). Together with other actions, these works support the Commercial Division of EDF by means of methodological and software tools, accompanying energy efficiency services for EDF customers.

As recalled in the beginning of Chapter VII, during the period from June 2004 to February 2008, I worked on the problem of defining and ranking energy efficiency solutions to be possibly proposed to individual customers. The aim of these works was to develop new concepts and ideas for inspiring the development of future tools supporting energy services. During this period I discovered and applied a class of methods concerned with multi-criteria decision-aid, and in particular those of the so-called ELECTRE family. Part of this work has been accompanied by ENGREF.

More precisely, my work mostly concerned two specific problems.

The first one was the definition and the comparison of multi-energy heating and domestic hot water solutions, based on the the effective coupling of systems using different energies, for instance electric convectors and wood-burning stoves for heating and solar thermal panels for domestic hot water. The production of this work consisted in reports, studies, as well as a software tool for the simulation, multi-criteria ranking and (preliminary) sensitivity analysis of multi-energy systems.

The other problem I coped with was the energy retrofitting of existing residential buildings. This work, achieved in the framework of a cooperation with the University of Liège, gave origin to a software tool, named REFLEX, including an expert system for identifying retrofitting alternatives, a dynamic simulation core for evaluating the building consumption (and thus the different benefits of the retrofitting alternatives) and several multi-criteria algorithms to rank the alternatives from the viewpoint of the customer and of the energy company (cf. page 139 and following).

I also took part in the development of an energy efficiency advice software for the British market, deployed by EDF Energy (EDF subsidiary in UK), named "EE Wizard", delivering customized energy audits by telephone or the Internet as well as propositions of alternatives for improving the energy efficiency of the dwelling, evaluated by means of a dynamic simulation core.

EDF R&D second period (since 2008). I joined the Industrial Risks Management Department (MRI) of EDF R&D in February 2008. This Department is focused on the study of[5] "*hazard-prone socio-technical systems operated within the EDF Group, such as nuclear and thermal power plants, hydraulic facilities and the power transmission network. This study includes various dimensions: (i) the component, (ii) the technical system, (iii) the human and organizational factors, (iv) the environment (natural, technological, organisational, regulatory, etc.).*"

More particularly, the works I have been involved in concern, on the one hand, the use of statistical and probabilistic techniques for the evaluation of the reliability of components, systems and structures, by means of feedback data analysis (cf. Chapters II and III) and/or numerical simulation (cf. Chapters IV and VI) and, on the other hand, the uncertainty and sensitivity analysis of complex computer models (cf. Chapters IV and V).

So far, this period has undoubtedly been the most intensive and creative one of my career.

[4] For knowing most of them, I have to say that this is definitely a sort of "*hallmark*" of Éric's students.

[5] Cf. the presentation leaflet of the Department at http://chercheurs.edf.com/organisation/nos-15-departements-93757.html.

From January 2009 to March 2014 I have been the leader of a large EDF R&D transverse project concerned with uncertainty analysis of computer codes. A great part of my activity has consisted in technical management. In addition to the "internal" EDF R&D project I have also been the coordinator of a multi-partners project, named OPUS (Open source Platform for Uncertainty treatment in Simulation, ended in 2011).

Firmly convinced of the positive impact (for both communities) of cooperation between industrial and academic researchers, I have been involved in several partnerships and I had the opportunity to serve in scientific societies. In particular, I founded in 2009 a thematic group "Reliability and Uncertainties" within the French Statistical Society (SFdS). In 2013 I have been elected member of the SFdS Council.

The most visible part of this collaborative activity consists in the organization of numerous and various dissemination actions: training sessions, workshops, sessions in conferences, special issues of scientific reviews.

At the same time, I had the opportunity to carry research works in particular in the fields of discrete reliability models (and actually nourish and achieve some works on Markov deterioration models initiated during my PhD) and on the use of Bayesian theory to build point estimation of safety criteria, as well as some introductory works about extra-probabilistic uncertainty assessment in computer simulation.

Chapter II

On the practical use of two discrete lifetime models

La spedizione per delimitare il confine nel tratto di frontiera rimasto scoperto partì il giorno dopo all'alba. La comandava il gigantesco capitano Monti, accompagnato dal tenente Angustina e da un sergente maggiore. A ciascuno dei tre erano state affidate la parola d'ordine di quel giorno e dei quattro giorni successivi. Era ben improbabile che tutti e tre potessero perire; ad ogni modo il più anziano dei soldati superstiti avrebbe avuto la facoltà di aprire la giubba dei superiori morti o svenuti, di frugare in un taschino interno, di trarne la busta sigillata contenente la parola d'ordine per rientrare nella Fortezza.

Dino Buzzati, Il deserto dei Tartari (1940)

The expedition to trace the unexplored stretch of frontier left the next day at dawn. In command was Monti, the huge captain, accompanied by Lieutenant Angustina and a sergeant-major. Each of the three had been entrusted with the password for that day and for the four following days. It was highly improbable that all three of them would perish; in any case the most senior surviving soldier would have had powers to open his superior officers' jackets, if they were dead or had fainted, to search in the little inside pocket and take from it the sealed envelope containing the secret pass for re-entering the Fort.

Dino Buzzati, The Tartar Steppe (translated by Stuart C. Hood, 1952)

Reading notes

Technical context. The works presented in this chapter concern the reliability of industrial components. They have been carried at the Industrial Risks Management Department of EDF R&D (MRI), mostly in 2008. Assessing the lifetime of pieces of equipment is a major input in the study of systems performances and safety, as well as in assets management. The studies carried at MRI in this field concern essentially power production facilities (nuclear, thermal, hydraulic).

Within this technical framework, I mainly focused on assessing the lifetime of industrial components which do not operate continuously but rather "on demand". In this cases, the lifetime is better expressed in "number of solicitations" (discrete variable), rather then a (continuous) calendar time.

The problem has been investigated in cooperation with INRIA-Paris Sud, in the straight continuation of other joint activities on industrial reliability. In 2008 I supervised, together with Emmanuel Remy (EDF R&D) and Gilles Celeux (INRIA-Paris Sud) the work of a young researcher (Côme Roero) concerning a deep analysis of two popular discrete lifetime models, namely the Inverse Pólya and Weibull-1 distributions. The main feature of this analysis is that it has been carried from a practical industrial viewpoint, that is we investigated the relevance of these models in engineering studies

characterised by (i) heavily censored data and (ii) failures occurring after a high number of solicitations (i.e. several hundreds or several thousands).

The conclusion is that both models, for different reasons, are not very useful tools for the reliability engineers who act in this particular context: Inverse Pólya carried an implicit assumption of decelerating ageing and Weibull-1 can be approximated by the continuous Weibull distribution, the parameters of which are more easily understandable and the handling of which is easier from a mathematical viewpoint.

Contributions. This work is described in a detailed internal EDF R&D report [Roero *et al.* 2008]. The main results have been presented in 2009 at the *41st Journées de Statistique* [Pasanisi *et al.* 2009b]. This study has been successively enriched in 2012 and 2013 and constitute the core of an article submitted in late 2013 [Pasanisi *et al.* 2013c].

As other contributions to the larger technical problem of lifetime assessment of industrial component, I am involved, as teacher, since 2008 in two training courses:

- professional training "Components reliability modelling: probabilistic and statistical methods and uncertainty analyses", hosted by the internal training institute of EDF R&D (ITECH, Institut de Transfert des Technologies),
- academic training "Reliability and industrial feedback" at the Université de Technologie de Troyes (notice also that I am coordinator since 2013 of the interventions of EDF R&D researchers in this course),

in which I essentially teach the basis of Bayesian lifetime modelling (exponential and Weibull models), with a particular focus on the role of expert opinion in case of scarce and poorly informative data.

Moreover, as president of the thematic Group "Reliability and Uncertainties" within the French Statistical Society (since 2009), I contribute to establishing bridges between researchers and engineers coping with the two strictly related problems of reliability assessment based on feedback data (topics of the present and the next chapter) and results of computer experiments (cf. Chapter IV).

Structure of the chapter and credits. After this introductory Section, the remainder of this chapter essentially excerpted from the paper "*On the practical interest of discrete lifetime models in industrial reliability studies*", available as a preprint [Pasanisi *et al.* 2013c] and submitted for publication.

1 Rationale and basic definitions

According to an usual definition, the reliability is "*the ability of an item to perform a required function under given conditions for a given time interval*" [IEC 1990]. The same norm also states that "*generally, reliability performance is quantified using appropriate measures. In some applications, these measures include an expression of reliability performance as a probability, which is also called reliability*", and finally gives another definition of reliability as "*the probability that an item can perform a required function under given conditions for a given time interval*". Other definitions (e.g. [DOD 1981]) are quite similar and, namely, are also twofold: reliability as both "ability" and "probability". Cf. also the interesting discussion in [Ahmed & Wu 2013], on the importance of terms "required function" and "given conditions" in this definition. Hence, it is interesting to see that, in practice, the reliability concept is inseparable with probabilistic considerations: failure is seen as an intrinsically random event, and the predictive tools of reliability engineers are statistical models.

There is a number of reference books, introducing the rationale of reliability engineering (e.g. [Aggarwal 1993, Zio 2007, Hamada *et al.* 2008]). Here, we will limit ourselves to the very basic definitions of survival analysis. First, let us consider the most common case where the random variable of interest "time to failure" (we denote it T) is continuous. Let $f(t)$ and $F(t)$ (with $t \in \mathbb{R}_+$), be its density and cumulative distribution function respectively. The following well-known definitions apply:

$$\begin{cases} \text{Survival function:} \quad S(t) = 1 - F(t) = \mathbb{P}[T \geq t] = \int_t^{+\infty} f(t)dt \\ \\ \text{Hazard function:} \quad \lambda(t) = \lim\limits_{dt \to 0} \dfrac{\mathbb{P}[T \in [t, t+dt] | T > t]}{dt} = \dfrac{f(t)}{S(t)}. \end{cases} \tag{II.1}$$

According to the usual definitions a component is said *ageing* if the hazard function $\lambda(t)$, also called "failure rate", is increasing. Moreover, the ageing can be increasing or decreasing, depending on the sign of the derivative $d\lambda/dt$.

For non-repairable systems, the expectation of T is usually called "mean time to failure" (MTTF):

$$\text{MTTF:} \quad \mathbb{E}[T] = \int_0^{+\infty} t\, f(t)dt. \tag{II.2}$$

The same considerations can be made for discrete survival models, the use of which is naturally considered when the lifetime of the piece of equipment under investigation cannot be properly expressed as a calendar time. It is the case of components which do not operate continuously and are only occasionally solicited. One can think of an on-off switch or an auxiliary power device: the activations of the switch or the starts of the engine can be considered as "occasional stresses" or solicitations for the equipment. In these cases, for reliability assessment purposes, the variable characterizing the lifetime of the component is not the operating time, measured as a calendar time (e.g. hours), but rather the number of solicitations that it can bear before failure. Another case of discrete lifetime data concerns pieces of equipment which only operate on cycles and the collected information is just the correct (or incorrect) behaviour at a given cycle.

In both cases, we can formalize the problem by stating that the random variable "lifetime", we aim at modelling, is discrete. Let us call it N. In the following, we will note, according to usual statistical notations, $n \in \mathbb{N}$ a particular value that could be taken by the random variable N.

In spite of the potential amount of case-studies in which discrete lifetime models could be considered as appealing tools for the engineers, surprisingly (or not?), they have been relatively not much investigated (theoretically and/or practically) in comparison with continuous models, which are nowadays largely used in industrial practice.

To our knowledge, the first scientific article proposing a discrete lifetime model (actually, derived from the Weibull distribution) dates from the mid of the 70's [Nakagawa & Osaki 1975]. [Bracquemond & Gaudoin 2003], in their reference paper, provide a quite exhaustive review of discrete distributions for lifetime data, including numerous references. Roughly speaking, they can be grouped into two categories: the ones derived from continuous models and the ones derived by urn schemes. After a concise statistical study, for various practical reasons and because of their convenient properties the authors recommend in most situations the use of the Type I discrete Weibull family (Weibull-1) defined by [Nakagawa & Osaki 1975], or the Eggenberger-Pólya distribution [Eggenberger & Pólya 1923]. In a more engineering framework, [Clarotti *et al.* 1997] stressed the importance to dispose of intuitive models, in the sense that their features have appealing meanings for the (often non-statistician) practitionner and can be interpreted by experts. Consequently, they considered that the Inverse Pólya model can be especially valuable in ageing problems. This offers an alternative solution to the difficulty of deriving discrete distributions from continuous ones, highlighted, for instance, by [Lai 2013].

Here, we will make use of the *classical* definition of the hazard function for discrete models as the conditional probability of failure after n solicitations (cf. Equations II.3 and II.5), given that the component already beared $n-1$ solicitations. Derived from the definition of the hazard function in the continuous case (where λ can be interpreted as a "conditional density function" [Lai 2013]), this is actually the most popular one in the technical community (e.g. [Salvia & Bollinger 1982, Shaked *et al.* 1995, Lawless 2003, Rinne 2008, Almalki & Nadarajah 2014]). It is worth noting that an *alternative* definition (first proposed by [Roy & Gupta 1992] and discussed by [Xie *et al.* 2002, Lai 2013]) exists of the hazard rate, defined as the logarithm of the ratio of the survival function in $n-1$ and n, that may allow for solving some theoretical and practical issues arising when using the classical definition.

This chapter is focused in particular on two probabilistic models, the use of which has been investigated in some industrial reliability studies at EDF R&D. As a major power producer and supplier, EDF is obviously concerned with the lifetime assessment of the components of its power plants as well as its transmission and distribution facilities. A huge amount of industrial feedback (observed or censored lifetime) data, often associated with the expertise provided by engineers and technicians, is available for coping with this problem. Due to the particular context of EDF business, the failure of its equipment can have cumbersome consequences, in terms of safety and availability. As a consequence, data are most of the time censored, i.e. actual failures are not observed. This makes the statistical analysis trickier. The main question we aim at addressing here is: "Are common discrete lifetime models adapted to the specific context of EDF?". Even if we do not pretend here to derive absolutely general conclusions on the interest of these models, however, we think that our conclusions can be useful for other analysts who share a similar context and/or similar data, i.e. situations in which failures are rare (data could be right-censored) and components are highly reliable (they fail after a relatively high number of solicitations).

The remainder of this chapter is organized as follows. We first discuss some properties of the popular Inverse Pólya model, which make it not suitable at all time in our context, in spite of its appealing simplicity and clarity of interpretation by the engineers' viewpoint. Actually, it appears that only situations where maintenance is known to prevent any accelerated ageing may be relevantly treated using this model. Then we discuss some properties of the so-called Weibull-1 model. Especially, some properties of this model are pointed out, that may sound odd for the practitioners and make its use and interpretation quite complicated. Theoretical results about the closeness of the Weibull-1 and the continuous Weibull distributions are given, which plead for using the latter as a robust and convenient approximation of the discrete model in the engineering practice. This viewpoint is then reinforced by numerical studies. The estimation of the considered models, in typical situations (including censoring), is investigated using simulated and real datasets. These experiments highlight the irrelevant aspects of the inverse Pólya model in concrete situations and the fair approximation made using the usual continuous Weibull distribution as a proxy of Weibull-1.

A concluding section sketches the main teachings of this study and proposes some avenues for further research.

2 Inverse Pólya model

The use of urn sampling schemes to model the behaviour of living [Marshall & Olkin 1993, Ivanova *et al.* 2000] or industrial [Alajaji & Fuja 1993, Bracquemond 2001] systems has been often considered. The basic principle of the numerous probabilistic models based on the Pólya urn scheme [Bracquemond & Gaudoin 2003, Johnson *et al.* 2005, Mahmoud 2008], first introduced in the 20's of the last century [Eggenberger & Pólya 1923, Pólya 1930], is to consider an urn in which, at the beginning of the experience, there are a black balls and b red ones, so that the probability to extract a black ball after a random trial is $a/(a+b)$. If a red ball is sampled, then z new black balls are added (together with the

sampled red ball), thus increasing the probability to sample a black ball. This scheme suggests an appealing probabilistic lifetime model [Clarotti *et al.* 1997] for discrete data: each solicitation of the piece of equipment is considered as a trial in a Pólya urn, where black balls are associated to the event "failure" and red ones to the event "correctly operating". Adding new balls can easily be interpreted as the result of a deterioration process.

The random variable N "number of the trial at which the failure occurs" follows a so-called Inverse Pólya distribution (IPD). Following [Bracquemond & Gaudoin 2003], we make use of the following parametrization for IPD:

$$\alpha = \frac{a}{a+b} \text{ and } \zeta = \frac{z}{a+b}, \quad \text{with } 0 < \alpha < 1 \text{ and } \zeta > 0.$$

Notice that parameter α can be easily interpreted as the probability of failure corresponding to the first solicitation ($n = 1$). The parameter ζ governs the ageing of the system: the higher ζ, the more severe will be the ageing.

The expression of the main reliability quantities of interest, according to this parametrization are given below. In the remainder, we will mainly focus on the hazard function (or failure rate) $\lambda(n)$:

$$\begin{cases} \text{Inverse Pólya model:} \quad N \sim \text{IPD}(\alpha, \zeta) \\ \text{Hazard function:} \quad \lambda(n) = \mathbb{P}[N = n | N > n-1] = \dfrac{\alpha + (n-1)\zeta}{1 + (n-1)\zeta} \\ \text{Prob. of failure after } n \text{ solicitations: } p(n) = \mathbb{P}[N = n] = \dfrac{(1-\alpha)^{n-1}(\alpha + (n-1)\zeta)}{\prod_{i=1}^{i=n}(1 + (i-1)\zeta)} \\ \text{Survival function: } S(n) = \mathbb{P}[N > n] = \dfrac{(1-\alpha)^n}{\prod_{i=1}^{i=n}(i + (i-1)\zeta)} \\ \text{MTTF: } \mathbb{E}[N] = \dfrac{(1-\zeta)\zeta^{(1/\zeta - 2)}}{(1-\beta)^{(1-\zeta)/\zeta}} \exp\left(\dfrac{1-\alpha}{\zeta}\right) \gamma\left(\dfrac{1-\zeta}{\zeta}, \dfrac{1-\alpha}{\zeta}\right) \end{cases} \tag{II.3}$$

In the expression of the MTTF above, $\gamma(\cdot,\cdot)$ is the so-called lower incomplete Gamma function:

$$\gamma(u, v) = \int_0^v x^{u-1} \exp(-x)\, dx.$$

The ageing of the component, i.e. the fact that the failure rate (the probability the component fails for the first time after n solicitations, given that it did not fail after $n-1$ solicitations) is an increasing function of n, is modelled by the addition of z black balls in the urn.

3 Modelling ageing by means of IPD: a major limitation

The condition $\zeta > 0$ ensures the ageing of the component under investigation. However, in practical applications one is also interested in describing situations where the ageing increases or decreases as the observed lifetime (here n) increases. This issue is solved by studying the sign of the second-order derivative of the failure rate. In the case of the Pólya model

the second-order discrete derivative of $\lambda(n)$ can be written, for $n > 2$ (after some algebra):

$$\lambda''(n) = \lambda(n) - 2\lambda(n-1) + \lambda(n-2) = \frac{2(\alpha-1)\zeta^2}{(1+(n-1)\zeta)(1+(n-2)\zeta)(1+(n-3)\zeta)}. \tag{II.4}$$

Under the conditions: $\alpha < 1$, $\zeta > 0$ and $n > 2$, it is easy to verify that the numerator and the denominator of Equation II.4 are negative and positive respectively. Thus, for any value of ζ, the second-order derivative of the failure rate is negative, i.e. the IPD can only model situations of decelerated ageing.

This result of decelerated ageing is confirmed by the intuition: if at each solicitation a number z of black balls is added into the urn, the more n increases, the more the number z of added balls becomes smaller than the number of the black balls already present. For large values of n, z becomes negligible and the added balls do not influence the failure probability anymore.

This situation can occur when a preventive maintenance is sufficient enough to prevent close breakdowns, that are typically encountered at the end of a component life. For this reason, and because the meanings of its parameters are rather intuitive, the Inverse Pólya model deserves interest in the reliability and risk community, although its use must also be strictly reserved to low ageing components or systems.

Due to the non-trivial handling of IPD, numerical computations proposed in the remainder needed to dispose of methods for simulating datasets and estimating the parameters in realistic cases. More precisely, it is needed:

- to have a view of the range of realistic values for (α, ζ), associated to various ageing situations;
- to describe a sampling method, given (α, ζ): this is done in Appendix (cf. page 153);
- to describe an appropriate estimation method; a maximum likelihood (ML) method devoted to this task is presented in Appendix (cf. page 153).

An answer to the first item is yielded in the experiment resumed in Table II.1. It is inspired by the case of both continuous and discrete Weibull distributions, in which the shape parameter β appears as an immediate indicator of the nature of ageing (see also Section 5 of this chapter). Its value can help the reliability engineers to synthetize the behavior of the considered component. It is therefore wanted to characterize the nature of ageing for the inverse Pólya model.

In a non-exhaustive way, several situations can be simulated using Weibull samples, on which inverse Pólya models are then fitted. On Table II.1, a range of such situations, from rejuvenation to accelerated ageing, are considered. In practice, the values of ζ/α shown in this table have been obtained by fitting IPD on a number of (discretized) lifetimes sampled from the usual (continuous) Weibull distribution.

Apart from providing ranges of plausible values for the ratio ζ/α in presence of rejuvenation or soft ageing, these results highlight the fact that, following engineering common sense, finding an estimate of this ratio upper than one discredits the "physical" relevance of the inverse Pólya model. Actually, a model considering that at each solicitation, the reliability decreases of an amount greater than the initial reliability, although mathematically possible, seems not coherent by an engineering viewpoint.

Except in situations, where preventive maintenance is integrated into the lifetime study, this restriction of the IPD can definitely be a concern for reliability engineers, as it is difficult to imagine (and to justify) *a priori*, in many situations, a hypothesis of decelerated ageing. The poor predictive properties of IPD in presence of data concerning systems presenting an increasing failure rate are exemplified later in the text (cf. Section 6). For this reason, the remainder of this chapter is mostly focused on another popular probabilistic model for discrete lifetimes, derived from the continuous Weibull distribution.

Scenario	Weibull shape parameter β	Ratio ζ/α
rejuvenation	$\beta \leq 0.9$	$\leq 10^{-5}$
no rejuvenation / no ageing	$\beta = 1$	$[8.10^{-5}, 10^{-4}]$
soft decelerated ageing	$\beta = 1.2$	$[5.8.10^{-4}, 7.10^{-4}]$
classical decelerated ageing (1)	$\beta = 1.5$	$[2.6.10^{-3}, 3.2.10^{-3}]$
classical decelerated ageing (2)	$\beta = 1.8$	$[2.10^{-2}, 4.10^{-2}]$
non-accelerated ageing	$\beta = 2$	$[0.25, 0.35]$
accelerated ageing	$\beta = 2.25$	$[1.28, 1.35]$
strongly accelerated ageing	$\beta = 2.5$	$[1.48, 1.85]$

Tableau II.1 – Typical magnitudes for the ratio ζ/α as a function of a Weibull shape parameter β, that indicates qualitatively the ageing behaviour of a component. These ranges of values were estimated by ML estimation from 500 discretized Weibull samples of size 1000, generated using scale parameter values in $\{10, 100, 500, 1000\}$.

4 The Weibull-1 model

The Weibull distribution (together with the exponential distribution which is actually a particular Weibull distribution) is the most popular probabilistic model for continuous lifetime data in engineering. Several discrete versions of the Weibull model for discrete data have been proposed. We focus here on the so-called "Weibull-1" distribution (or Type I Weibull distribution), which is historically the first one, proposed in 1975 [Nakagawa & Osaki 1975]. Recommended by several authors [Bracquemond & Gaudoin 2003], it can be derived from the usual (continuous) Weibull distribution by time discretization or alternatively defined by means of its survival function, which has formally the same expression as the continuous Weibull's one. Thus, the following notations and definitions apply:

$$\begin{cases} \text{Weibull-1 model:} \quad N \sim \mathrm{W}_1(\eta, \beta) \\ \text{Hazard function:} \quad \lambda(n) = 1 - \exp\left[-\left(\frac{n}{\eta}\right)^{\beta} + \left(\frac{n-1}{\eta}\right)^{\beta} \right] \\ \text{Prob. of failure after } n \text{ solicitations:} \quad p(n) = \exp\left[-\left(\frac{n-1}{\eta}\right)^{\beta} \right] - \exp\left[-\left(\frac{n}{\eta}\right)^{\beta} \right] \\ \text{Survival function:} \quad S(n) = \exp\left[-\left(\frac{n}{\eta}\right)^{\beta} \right]. \end{cases} \tag{II.5}$$

It is worth noting that, although no closed form of the MTTF exists for the Weibull-1 model, upper and lower bounds can be given, cf. Equation (II.8).

The Weibull-1 model $\mathrm{W}_1(\eta, \beta)$ is often re-parametrized as $\mathrm{W}_1(\theta, \beta)$, with $\theta = \exp\left(-1/\eta^{\beta}\right)$. This parametrization allows for a very easy interpretation of the parameter θ: actually, $1 - \theta$ is the probability of failure at the first solicitation (i.e. for $n = 1$). Nevertheless, the advantage of the parametrization (η, β) is the easiness of the comparison of the two distributions $\mathrm{W}_1(\eta, \beta)$ and $\mathrm{W}(\eta, \beta)$ (i.e. Weibull-1 and continuous Weibull having the same parameters) and, following the purposes of our study, it will be used in the remainder of this chapter.

Notice that other discrete distributions can be proposed from the continuous Weibull one, and namely the Weibull-2 [Stein & Dattero 1984], preserving the power function form of the hazard rate, and the Weibull-3 [Padgett & Spurrier 1985]. See also [Jazi *et al.* 2010, Alzaatreh *et al.* 2012, Bebbington *et al.* 2012, Lai 2013] for examples of more complex related

distributions, as well as the recent review paper of [Almalki & Nadarajah 2014] proposing several variants of both discrete and continuous Weibull distributions. However, as reminded in [Rinne 2008], no discrete distribution exists that can mimic all the functional forms and the properties, so familiar to engineers, of the continuous Weibull one.

A number of applications of the Weibull-1 distribution can be found in the technical literature, e.g. modelling the number of shocks [Sheu 1998] or the number of preventive maintenance actions [Liao *et al.* 2009, Liao & Sheu 2011] supported by a repairable system before the total loss of operability in the context of optimal replacement strategies or the number of items produced in an in-control state of a manufacturing process before shifting to an out-of-control state [Wang & Sheu 2001, Wang & Sheu 2003, Wang *et al.* 2009, Tsai & Wang 2011].

Other examples of applications exist outside the industrial reliability context, in which Weibull-1 distribution has been used for modelling: the discretized duration of wind events [Castino *et al.* 1998], the recruitments of trees in *Growth and yield* models of forests [Fortin *et al.* 2009], the distribution of polymeric particles hosting the active agent in drug release experiments [Grassi *et al.* 2000], the time to replacement of a technology option (water heaters and solar photo-voltaic panels) aiming at reducing energy consumption and greenhouse gas emissions [Higgins *et al.* 2014], the number of cells population doublings until senescence in *in vitro* experiments [Wein & Wu 2001].

5 Modelling ageing by means of the Weibull-1 distribution

One of the most interesting features of the continuous Weibull distribution from an engineer's viewpoint (and probably the reason of its success within the technical community) is the great flexibility of the hazard function $\lambda(t)$ which can model very different ageing mechanisms. Moreover, the parameters (η,β) of this model have a clearly understandable technical meaning. The first is the quantile of the lifetime corresponding to a survival probability of approximately $1/3$ (actually 0.37) and the latter rules (independently of the value of η) the ageing of the system: (i) rejuvenation if $\beta < 1$, (ii) constant hazard rate if $\beta = 1$ (exponential model), (iii) decelerated ageing if $\beta \in]1,2[$, (iv) accelerated ageing if $\beta > 2$.

These are very useful properties which become even more interesting in a Bayesian framework for eliciting formal informative priors [Bousquet 2005, Bousquet 2008, Bousquet 2010] from available expertise.

Starting from this very technical viewpoint, we investigated how these properties of the Weibull model can be transposed to Weibull-1 and we particularly focused on the relation between the value of β and ageing, i.e. the monotonicity properties of the function $\lambda(n)$.

From the expression of $\lambda(n)$ (cf. Equations II.5) it can be shown that:

- For $\beta = 1$, the hazard function is constant. A trivial calculation gives: $\lambda(n) = 1 - \exp(-1/\eta)$.
- For $\beta > 1$, the hazard function is an increasing function of n. One has just to consider, for $n \geq 2$, the argument of the exponential in the expression of $\lambda(n)$:

$$\left(\frac{n-1}{\eta}\right)^{\beta} - \left(\frac{n}{\eta}\right)^{\beta}. \tag{II.6}$$

  This function of n is decreasing for $\beta > 1$, thus $\lambda(n)$ is increasing.

- for $\beta < 1$, the hazard rate is a decreasing function of n. One has just to follow the same reasoning about the monotonicity of the function (II.6) above, which is increasing if $\beta < 1$.

For $\beta \in]1,2]$ it is possible to show analytically that $\lambda(n)$ is a concave function of n, i.e. the ageing is decelerated (proof in Appendix, page 154).

For $\beta > 2$, we did not succeed to prove any analytical results. However, as also highlighted by [Xie *et al.* 2002], following the classical definition of $\lambda(n)$ as a conditional probability, it is obvious that this function cannot be strictly convex as it must tend to 1 as $n \to \infty$, which is not the case for the Weibull continuous model.

We found empirically, by studying the convexity of $\lambda(n)$ for $(\eta, \beta) \in [1, 1000] \times [2.1, 20]$, that for a given β, a value η_0 of η exists, so that for each $\eta < \eta_0$, $\lambda(n)$ is strictly concave and for each $\eta > \eta_0$, $\lambda(n)$ is initially convex, then concave, presenting thus an inflection point. The main lesson learnt by this empirical study is that, unlike the continuous Weibull distribution, the concavity of the hazard function does not depend on β only, but also on η. As a conclusion, an interesting property of the Weibull model, particularly attractive for engineers, is actually lost when switching to Weibull-1.

Figure (II.1) shows the value of n corresponding to the inflection points, found by means of the empirical study described above.

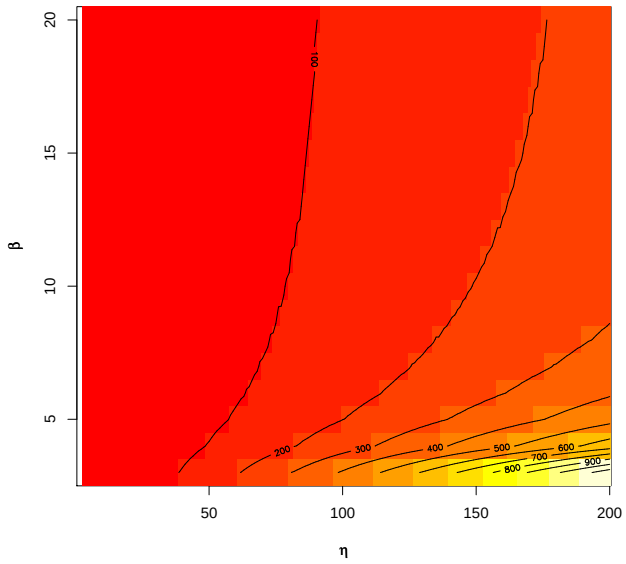


Figure II.1 – Weibull-1 model. Values of the solicitation number n corresponding to the inflection points of the hazard function $\lambda(n)$ as a function of η and β.

In practice, the presence of an inflection point could not be a serious issue in practical problems: that happens when the corresponding value of n is a quantile corresponding to a very low survival probability. In that case, we can conclude that for the set of values of n interesting for practical purposes, $\lambda(n)$ is convex (accelerated ageing).

6 Ageing: two numerical examples

In order to show the behaviour of IPD and Weibull-1 models with respect to the fundamental engineering issue of ageing, it is interesting to see how these ones are able to reproduce two known hazard functions from simulated samples. More precisely, starting from two known hazard functions, shown in Figure II.2 (blue curves), two samples of 100 uncensored discrete lifetimes for each of the components were randomly generated. Then, using the two generated data sets, we estimated by the Maximum Likelihood method the corresponding parameters of IPD and Weibull-1 models, and we plotted the hazard functions corresponding to the estimated parameters. The relevance of this method to assess the quality of the adjustment in the reliability context is defended by [Bracquemond & Gaudoin 2003]. The results are graphically

shown in the same figure II.2 (dotted curves).

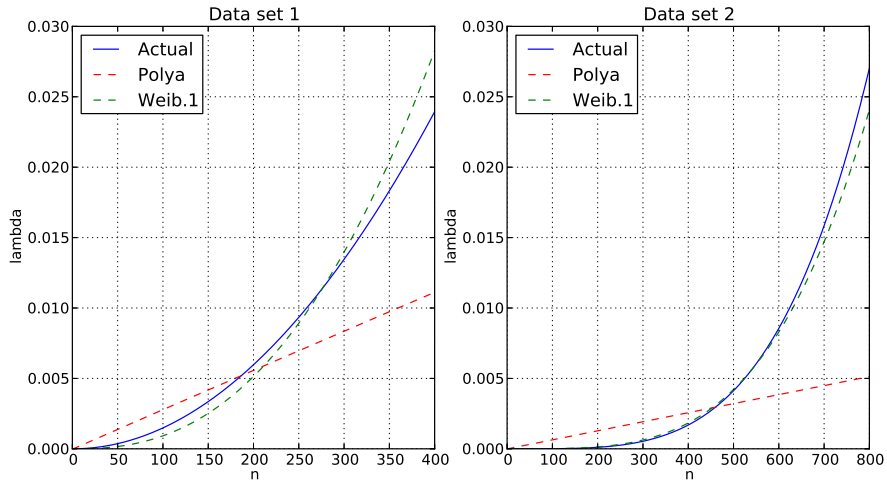


Figure II.2 – Actual (blue curves) hazard functions and estimated hazard functions according to Weibull-1 and IPD models, the parameters of which have been estimated by maximizing the likelihood of random data sets of size 100.

Not surprisingly, in presence of convex hazard functions (i.e. accelerating ageing) the performance of IPD is poor. Intuitively, as IPD can return concave hazard functions only, the best approximation it can give of a convex λ is a linear function. Instead, the flexible Weibull-1 model returns a quite fair approximation of λ.

7 Approximating Weibull-1 by continuous Weibull model

Face to the issues sketched hereinbefore, it may appear practical for the engineer to use the usual (continuous) Weibull distribution for the assessment of main functions of interest in a reliability study. If seen as a possible "continuous approximation" of the Weibull-1 distribution, its computational treatment (e.g. parameter estimation, sampling) is well known and familiar to the practitioner. As evoked in Section 4, let us focus on the two random variables:

$$N \sim \mathrm{W}_1(\eta,\beta) \quad \text{and} \quad T \sim \mathrm{W}(\eta,\beta). \tag{II.7}$$

They have the same parameters but the first is discrete and follows a Weibull-1 distribution and the latter is continuous and follows a Weibull distribution. The closeness of both models appears first in the closeness of MTTF's. The following proposition (the proof of which is given in Appendix at page 155) highlights in particular that the two MTTF's (noted $\mathbb{E}_{\mathrm{W}_1}[N|\eta,\beta]$ and $\mathbb{E}_{\mathrm{W}}[T|\eta,\beta]$ respectively) are closer and closer as both quantities are $\gg 1$.

Proposition 7.1 *Given the two random variables T and N (defined by Equation II.7), the following inequality stands:*

$$\mathbb{E}_W[T|\eta,\beta] \leq \mathbb{E}_{W_1}[N|\eta,\beta] \leq 1+\mathbb{E}_W[T|\eta,\beta].$$

By definition, the survival functions of both models have the same mathematical form, that is they lead to the same value of the survival and the failure probability for a given n. Therefore, the expression of the quantile n_q of probability q

is the same for $\mathrm{W}_1(\eta,\beta)$ and $\mathrm{W}(\eta,\beta)$:

$$n_q = t_q = \eta\left[-\log(1-q)\right]^{1/\beta}.$$

We conclude that the two distributions (which have similar means and the same quantiles) are extremely close to one another. In particular, in an engineering context, they give the same values for the main quantities of practical interest (MTTF, quantiles, probability of failure).

Moreover, when estimating (η,β) from actual industrial feedback data in presence of right-censored observations, the likelihoods of the two models tend also to be very close: the inference, thus, leads to very similar estimates for η and β for both models. Actually, it can be seen that the likelihoods of a given samples of discrete lifetime, according to $\mathrm{W}_1(\cdot)$ and $\mathrm{W}(\cdot)$ respectively, are closer and closer as (i) the rate of right-censored data increases, and (ii) the (unknown) value of η is high.

The proof of the first part of the assertion is trivial: any right-censored datum n contributes to the likelihood by means of the value of the survival function $S(n|\eta,\beta)$, which has the same expression for both Weibull and Weibull-1 distributions. As far as the second part of the proposition is concerned, if we note $f_{\mathrm{W}}(\cdot)$ the density of $\mathrm{W}(\eta,\beta)$ and $p_{\mathrm{W}_1}(\cdot)$ the probability distribution of $\mathrm{W}_1(\eta,\beta)$, the contribution of an uncensored observation n to the likelihood of the two models is equal to $f_{\mathrm{W}}(n)$ and $p_{\mathrm{W}_1}(n)$ respectively.

As the survival functions $S(\cdot)$ have the same expression for $\mathrm{W}_1(\eta,\beta)$ and $\mathrm{W}(\eta,\beta)$, one can write:

$$p_{\mathrm{W}_1}(n) = \mathbb{P}[N > n-1] - \mathbb{P}[N > n] = S(n-1) - S(n) = \int_{n-1}^{\infty} f_{\mathrm{W}}(t)dt - \int_{n}^{\infty} f_{\mathrm{W}}(t)dt = \int_{n-1}^{n} f_{\mathrm{W}}(t)dt. \tag{II.8}$$

It is easy to provide the following bounds for the last integral in the right hand side of Equation II.8:

$$\min_{t\in[n-1,n]} f_{\mathrm{W}}(t) \leq p_{\mathrm{W}_1}(n) \leq \max_{t\in[n-1,n]} f_{\mathrm{W}}(t). \tag{II.9}$$

Intuitively, the higher the values of η and n are, the closer the bounds in Equation II.8 are and, consequently, the closer $p_{\mathrm{W}_1}(n)$ and $f_{\mathrm{W}}(n)$ are. The graphs displayed on Figure II.3 confirm, empirically, this intuition.

More formally, in Appendix (page 155), it is proven the following proposition.

Proposition 7.2 *For all $\beta \geq 1$:*

$$\lim_{\eta\to\infty} \sup_{t\in\mathbb{R}^+} |p_{W_1}(t|\beta,\eta) - f_W(t|\beta,\eta)| = 0. \tag{II.10}$$

Notice that the convergence to 0 of $|p_{\mathrm{W}_1}(t|\beta,\eta) - f_{\mathrm{W}}(t|\beta,\eta)|$ when $t\to\infty$ is trivial as both $p_{\mathrm{W}_1}(t|\beta,\eta)$ and $f_{\mathrm{W}}(t|\beta,\eta)$ tend to 0.

In the remainder, we exemplify and highlight these features of Weibull and Weibull-1 distributions. An empirical study of the specific properties of the Weibull-1 distribution is carried out, in view of testing the ability of the continuous Weibull model to approximate the Weibull-1 model. The study is conducted using simulated datasets. Finally, a last study is presented, involving two industrial examples, based on real feedback data.

7.1 Empirical study

The previous results suggest that for practical industrial purposes (i.e. predicting probabilities of failure and MTTF) the Weibull-1 and Weibull models provide very similar outcomes when η is high. High values of η mean that the system

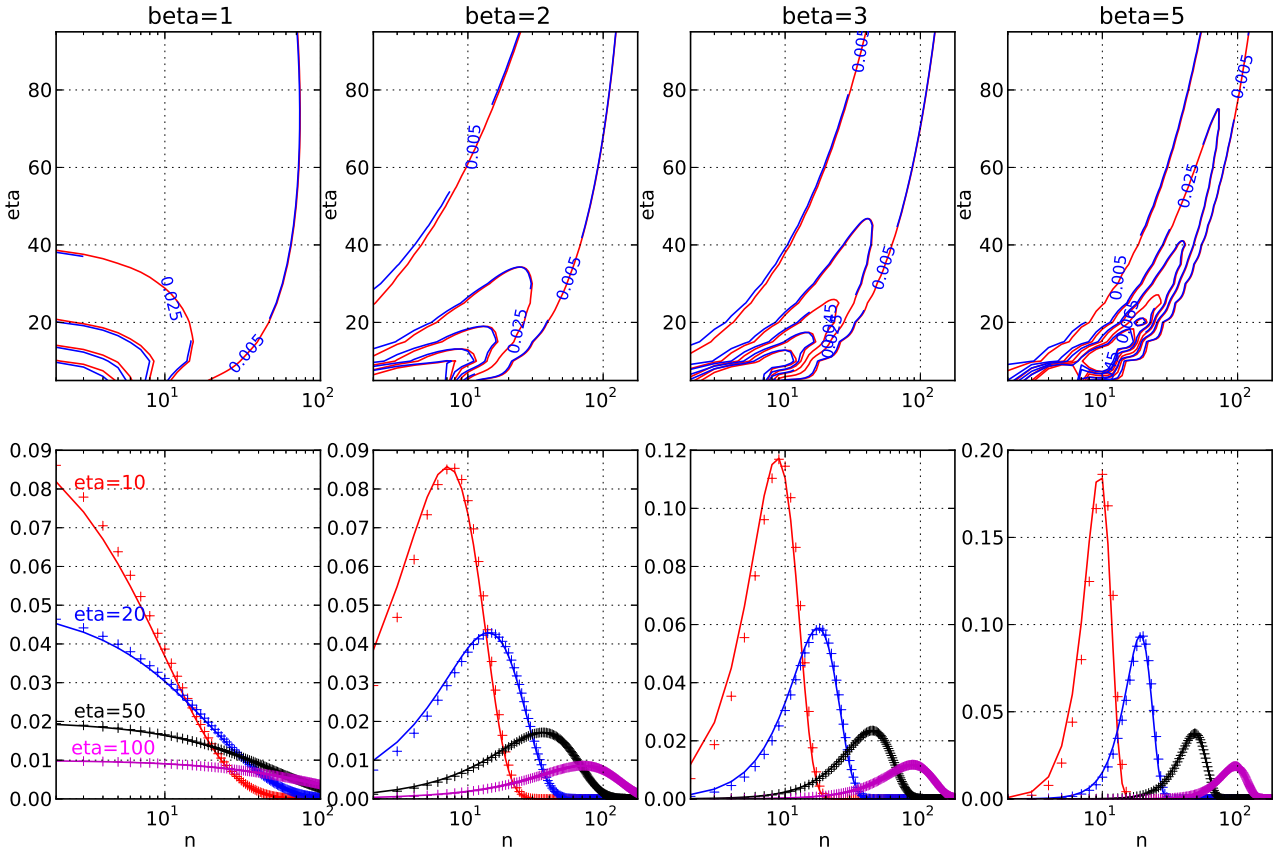


Figure II.3 – Upper part: plot of the Weibull-1 (red lines) and Weibull distribution functions (blue lines) for fixed values of β as a function of n and η; one can see that isolines are very close to one another. Lower part: plot of the discrete Weibull-1 distribution (crosses) vs. the corresponding values of the continuous Weibull for given values of η and β (continuous lines), as functions of n; as one can see, the higher η and n are, the more the crosses tend to be superposed over the lines.

under investigation is *reliable* in the common-sense meaning, that is failures occur for *high* values of n (i.e. $\gg 1$).

Moreover, industrial feedback datasets contain generally a number of censored data. In particular, in our specific industrial context, data are most of the time right-censored (and quite never left-censored) because failures are to be strictly avoided as they have a costly impact on availability of the overall production facility.

As shown in the last part of the previous section, a set of lifetimes of a reliable system with a significant number of censored data leads to a very similar likelihood under the two hypotheses of Weibull-1 and Weibull model. Hence, Maximum Likelihood estimations (MLE) of (η,β) for both models are expected to be very close. To confirm these results, we carried intensive numerical simulations, by generating datasets likely to be encountered in industrial practice and thus evaluating the MLE of (η,β) for Weibull-1 and Weibull model, noted $(\hat{\eta}_{W_1},\hat{\beta}_{W_1})$ and $(\hat{\eta}_W,\hat{\beta}_W)$ respectively.

More precisely, for $(\eta,\beta) \in \{10,50,300,500,800,1000\} \times \{0.5,1,1.5,2,2.5,3,5,10\}$, and for right-censored data rates of 0%, 25%, 50% and 75%, 5000 samples of sizes 50 and 100 were generated from the Weibull-1 distribution $W_1(\eta,\beta)$.

Based on these data, the MLE $(\hat{\eta}_{W_1},\hat{\beta}_{W_1})$ and $(\hat{\eta}_W,\hat{\beta}_W)$ were evaluated, as well as the relative errors concerning the estimations of (η,β):

$$\frac{\eta-\hat{\eta}_{W_1}}{\eta},\quad \frac{\beta-\hat{\beta}_{W_1}}{\beta},\quad \frac{\eta-\hat{\eta}_W}{\eta},\quad \frac{\beta-\hat{\beta}_W}{\beta}$$

and the relative errors of plug-in estimators of the following quantities of interest: hazard rates corresponding to the quantiles of probabilities $(0.5, 0.75, 0.90, 0.99)$ of the original distribution, MTTF and quantiles. As data have been generated from Weibull-1 distributions, one can expect the estimators $(\hat{\eta}_{W_1}, \hat{\beta}_{W_1})$ to be closer to the actual values of (η, β) than $(\hat{\eta}_W, \hat{\beta}_W)$, the first being obtained by fitting the "true" probabilistic model, the latter by fitting an approximation of it.

Figure II.4 shows some results of this empirical study. Here, the mean ML estimation error under the Weibull-1 model assumption (x-axis) is plotted against the error under the continuous assumption. One can see that the points of the scatterplot are quite close to the first bisector, showing that using the Weibull-1 model yields no significant improvement with respect to the continuous approximation.

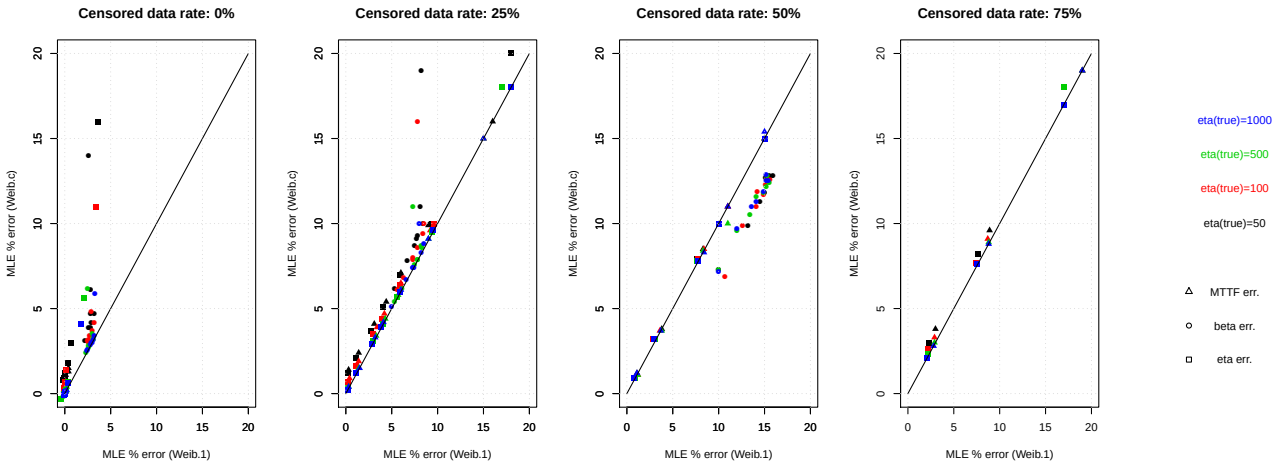


Figure II.4 – Relative errors of the MLE of η (circles) and β (squares) and of the ML plug-in estimator of the MTTF (triangles), obtained from data simulated from the Weibull-1 model under the hypothesis of different censoring rates from 0% to 75%. The estimation is carried using the Weibull-1 and the continuous Weibull model assumption. The colors correspond to different values of the *true* η.

7.2 Inference from actual feedback data

As a conclusion of this study, it is interesting to examine two datasets coming from actual industrial feedback. Even though this analysis is proposed for exemplary purposes only, nevertheless data are representative of the ones reliability engineers cope with in our specific business context. Table II.2 shows the main features of the examined datasets. A quick look at data summary allows to find out two important characteristics. First, most of the data are right-censored lifetimes: the censor rates are equal to 96% and 81% respectively. Second, the components under investigation are *reliable*, in the sense that failures are expected to occur after a (relatively) high number of solicitations: the empirical means of the (highly censored) observed data are 63.8 and 334.5 respectively.

In the same table are also shown the Maximum Likelihood estimators of the parameters of the Inverse Pólya, Weibull-1 and Weibull models.

Regarding the first data set, the estimated parameters of both Weibull-1 and Weibull models (which have very similar values) suggest an accelerated ageing. The extremely high (and hardly understandable by a technical viewpoint) value of ratio $\hat{\zeta}/\hat{\alpha}$ (order of magnitude: 10^6) highlights a poor modelling performance of Inverse Pólya. Figure II.5 shows the cumulative distribution functions (CDF) of the three estimated distributions as well as the Kaplan-Meier estimator. In

	Sample 1 (Aux. power device linings) [Clarotti *et al.* 1997])	Sample 2
Data size	497	48
Sum of observed data	31715	16058
Observed failures	18	9
Number of right-censors	479	39
Parameters estimation		
Inv. Pólya	$\hat{\alpha} = 7.037 \cdot 10^{-12}$ $\hat{\zeta} = 1.349 \cdot 10^{-5}$	$\hat{\alpha} = 5.601 \cdot 10^{-4}$ $\hat{\zeta} = 1.774 \cdot 10^{-19}$
Weibull-1	$\hat{\eta} = 306.814$ $\hat{\beta} = 2.320$	$\hat{\eta} = 1530.139$ $\hat{\beta} = 1.122$
Weibull	$\hat{\eta} = 320.580$ $\hat{\beta} = 2.320$	$\hat{\eta} = 1510.250$ $\hat{\beta} = 1.124$

Tableau II.2 – Example of analysis of data set coming from actual industrial feedback. Upper part: data summary. Lower part: Maximum Likelihood estimators of the parameters of Inverse Pólya, Weibull-1 and Weibull distribution.

spite of the issues evoked hereinbefore, the prediction properties of the three models (in terms of failure probabilities) are quite equivalent within the range of observed data. Yet, as shown in Figure II.6, the predictions given by IPD for higher values of n are more optimistic and less conservative, in the sense IPD provides lower values of the CDF (i.e. higher values of the reliability function) than the ones given by Weibull and Weibull-1, the CDF's of which are practically indistinguishable.

As far as the Sample 2 is concerned (cf. Figure II.7), the components do not show a significant ageing (the Weibull shape parameter is close to 1). The three probabilistic models return a very similar prediction in terms of CDF (and reliability function).

As a conclusion, these exemplary analyses confirm the conclusions presented in the previous Sections, by means of theoretical and empirical considerations: for engineering purposes, the continuous Weibull model is a fairly good alternative to the discrete model (Weibull-1) investigated in the framework of the present study.

Remark. We stress that, although the data come from real surveys, the study shown in this section is given for exemplary purposes only and neither results nor methodology must be extrapolated to make any general conclusion about the reliability of EDF industrial components or EDF risk assessment policies.

8 Discussion

The study shown hereinbefore has highlighted some weaknesses of both inverse Pólya (IPD) and Weibull-1 distributions as discrete models for lifetime of industrial components.

IPD carries the implicit hypothesis of decelerating ageing, that can definitely be an issue as this assumption can be hardly justified a priori in industrial studies.

As far as the Weibull-1 model is concerned, it has been shown that the popular interpretation of the shape and scale parameters of the Weibull distribution is no longer valid for its discrete version. In particular, the type of ageing does not depend on β only but also on η.

Moreover, for practical purposes, the Weibull-1 model and the Weibull model are very close. In practice, the Maximum Likelihood estimation of the parameters (η, β) computed under the hypotheses of Weibull-1 and Weibull models lead to

the same results. That is more and more true as far as the value of η is high (i.e. the piece of equipment under investigation is reliable, in the sense that it normally fails after a significantly high number of solicitations) and the rate of censored data is high. That is exactly the case of an industry like EDF: in this context, fortuitous failures can have a great impact on the availability of the production facilities and lead to high unexpected costs. For these reasons they have to be avoided: components are highly reliable and they are replaced well before that failures are likely to occur.

Thus, the practical impact of the use of Weibull-1 model for improving reliability analyses based on feedback data is quite low.

Of course, the conclusions of this study can be questionable and (we insist) they must be clearly put inside the specific context of an industry like EDF. Moreover the study is limited to IPD and Weibull-1 models, as their use has been evoked in former internal technical reports as an interesting perspective. Other probabilistic models exist and we do not pretend to give general conclusions about discrete lifetime models. Nonetheless, the easily-interpretable features of the Inverse Pólya distribution could remain valuable in practice if the phenomenon of decelerating ageing could be discarded. It is likely that adding a supplementary hypothesis of the following nature could improve the versatility of the model: the number z of balls added at solicitation n should follow an increasing pattern in function of n rather than remaining constant. Defining and comparing several patterns, from both analytical and computational viewpoints, should be a keypoint of future studies aiming at preserving the interest of IPD in reliability analysis.

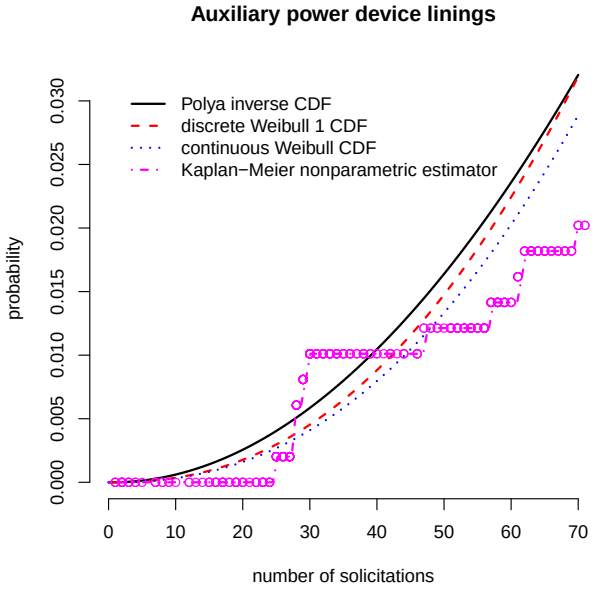


Figure II.5 – Data Sample 1: Cumulative distribution functions from ML estimations of Inverse Pólya, Weibull and Weibull-1 model and non-parametric Kaplan-Meier estimator.

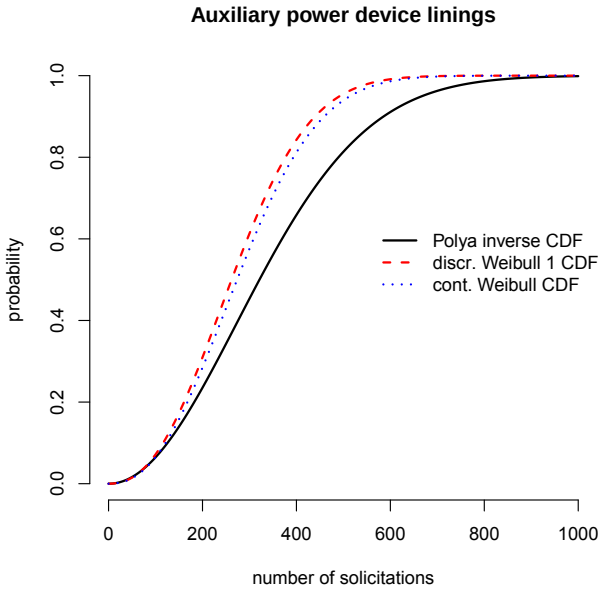


Figure II.6 – Data Sample 1: Cumulative distribution functions from ML estimations of Inverse Pólya, Weibull and Weibull-1 model. The range of the number of solicitation is here extended beyond the maximum of the observed sample to show the predictive properties of the model for high values of n.

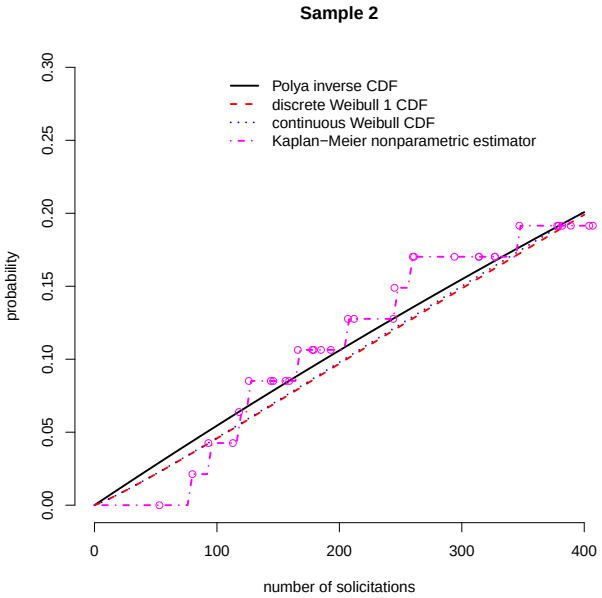


Figure II.7 – Data Sample 2: Cumulative distribution functions from ML estimations of Inverse Pólya, Weibull and Weibull-1 model and non-parametric Kaplan-Meier estimator.

Chapter III

Markov chain modelling of industrial systems deterioration

"The trees and the Ents," said Treebeard. "I do not understand all that goes on myself, so I cannot explain it to you. Some of us are still true Ents, and lively enough in our fashion, but many are growing sleepy, going tree-ish, as you might say. Most of the trees are just trees, of course; but many are half awake. Some are quite wide awake, and a few are, well, ah, well getting Entish. That is going on all the time."

John Ronald Reuel Tolkien, The Two Towers (1954)

Reading notes

Technical context. Markov chains are useful tools for the reliability engineer when analysing the behaviour of systems which are subject to deterioration. In this case, a simple and intuitive scheme consists of a discrete state model, each state describing more or less degraded operating conditions or performances. Knowing the initial state (or at least the initial probabilities for the system to be in each of the considered states), assuming a first-order stationary Markov scheme, the behaviour of the system is entirely described by the transition probabilities from a state to another in the time unit (e.g. one year).

The technical context in which I have been first interested in this class of models has been the analysis of the accuracy of water meters, for the private company Génerale des Eaux, I worked for from December 2000 to February 2004. Actually, though the company was mainly interested in forecasting the overall accuracy (ratio between billed and consumed water) for estimating the unaccounted-for water and thus planning the periodic renewal of the machines, regression-like models directly dealing with accuracy as regressand proved to be ineffective (as in fact observed accuracies seemed to be rather sampled from a mixture of populations). I proposed, instead, a four-state Markov model, as it is detailed in Section 7 (pages 49 and following).

A problem that frequently arises in engineering practice is that data are not adapted to the common inference methods for estimating the transition probabilities. Actually, although the estimation is very easy in presence of complete i.i.d. (independent and identically distributed) sequences of states for a number of systems over a given time period, the problem become trickier when the analyst has at his/her disposal, sequences with missing observations (or even constituted by only one observation) or aggregated data (i.e. the number of systems in a given state at a given time).

EDF R&D is also concerned with this kind of models. When I joined the Industrial Risks Management Dept. in 2008 I had the opportunity to work again on the same problem of estimating Markov transition probabilities from incomplete

data, for dealing with problems of cracks propagation in pieces of equipment of power plants (cf. Section 6).

Contributions. The degradation of the accuracy of water meter has been the problem at the heart of my PhD works, carried within the framework of a CIFRE partnership between the Compagnie Générale des Eaux (CGE) and ENGREF (École Nationale du Génie Rural, des Eaux et des Forets). Besides the PhD thesis [Pasanisi 2004a], different features and results of these works have also been presented at the *7th Valencia International Meeting on Bayesian Statistics* [Pasanisi *et al.* 2002] (as a poster), at the *35th Journées de Statistiques* [Pasanisi 2003] and at the first two editions of the *Rencontres Statistiques au Sommet de Rochebrune* [Pasanisi 2002, Pasanisi 2004d] and published in the *Revue de Statistique Appliquée*[1] [Pasanisi & Parent 2004].

Some years later, at EDF R&D, I resumed working on this class of models, and I attempted to better formalize the inference problem, with the help of my colleague Nicolas Bousquet. In 2009 we supervised the internship, about this topic, of Shuai Fu (who successively joined my project team on Uncertainty Analysis as PhD student) . The main result of this work is the acceleration of the Metropolis-Hastings algorithm for the Bayesian inference of transition probabilities (cf. Section 3) by introducing an instrumental density modelling the dependence of elements of different rows of the matrix by means of a Gaussian copula. Another result is the equivalence between the case of aggregated data and the one of sequences made by one single observation.

First presented at the *42nd Journées de Statistiques* [Pasanisi *et al.* 2010], these works have been later on published in the journal *Computational Statistics and Data Analysis* [Pasanisi *et al.* 2012a].

Structure of the chapter and credits. Most of the remainder of this chapter is constituted by the text of the article "Estimating Discrete Markov Models From Various Incomplete Data Schemes" [Pasanisi *et al.* 2012a] co-authored with Nicolas Bousquet and Shuai Fu (with very few adaptations).

The content of Section 7 is essentially adapted from my PhD thesis as well as other studies carried within my PhD work [Pasanisi *et al.* 2002, Pasanisi 2003, Pasanisi & Parent 2004].

1 Introduction

In many applications, the analyst is required to model and/or to predict the behaviour of a system Σ which is fully characterized (with respect to the framework and the purposes of the analysis) by a discrete variable of interest Z which takes time-dependent values within a finite (discrete) set $\mathcal{S} = \{s_1, s_2, \ldots, s_r\}$ of r classes (let us call them *states*). For instance, he/she could be interested in estimating the probability $p_{\mathcal{A}}(t)$ for Z to be in a given set of states $\mathcal{A} \subset \mathcal{S}$ as a function of time t.

In a reliability analysis context, these states can correspond to failure states, thus $1 - p_{\mathcal{A}}(t)$ is the reliability function of the system under investigation Σ at time t. Another function of interest could be the expected number of states $N_{\mathcal{A}}$ before Σ reaches $\mathcal{A}$. To do so, the analyst first has to estimate the transition probabilities from one state to another, i.e. estimate the transition matrix $\boldsymbol{\theta}$. The vector $\boldsymbol{p}(0)$ of the initial probabilities $p_1(0), \ldots p_r(0)$ for the system to be in states $s_1, \ldots s_r$ respectively, at $t = 0$, is usually assumed known in real-life applications; therefore, the knowledge of the transition matrix $\boldsymbol{\theta}$ allows to evaluate, for a given time t, the probabilities for being in each of the r states, i.e. the vector:

$$\boldsymbol{p}(t) = \boldsymbol{p}(0) \cdot \boldsymbol{\theta}^t. \tag{III.1}$$

[1] Journal edited by French Statistical Society (SFdS) up to 2006 when it has been merged with the Journal de la Société Française de Statistique.

Given some data $\boldsymbol{z}$ under the form of observed sequences of states, the statistical estimation of these probabilities is traditionally facilitated by a time-homogeneous, first-order Markov stationarity assumption about the process $\boldsymbol{\Lambda}$ which generates the data. In other words, the transition probability $\theta_{i,j}$ from any state s_i to any other state s_j (i possibly equals to j) is assumed to be independent of time and of the past trajectories before reaching s_i. Of course, this assumption is questionable and may seem restrictive regarding the external knowledge about the process and the complexity of the actual system [Sonnenberg & Beck 1993, Diebold *et al.* 1994, Faissol *et al.* 2009, Grimshaw & Alexander 2011]. However, as noted by [Jones 2005], "*using (possibly more appropriate) higher-order processes increases the complexity and data requirements quite substantially, and may not be feasible with only a limited time series*". That is often the case in practical applications and in particular in the studies this chapter is concerned with.

Here, the interest in Markov models is rooted in industrial reliability problems; cf. Section 6 and (especially) Section 7. Many other applications[2] can be found in this field: discrete Markov schemes, the states of which correspond to gradually deteriorated operating conditions, have for instance been used to assess the reliability of programmable electronic systems [Bukowski & Goble 1995], cogeneration plants [El-Nashar 2008], machineries of oil refineries [Cochran *et al.* 2001], piping systems of power plants [Cronvall & Männistö 2009], welded structures submitted to fatigue damage [Lassen 1991] and cracks propagation (cf. Section 6).

As a more recent example, a scheme based on Markov (hidden) discrete states has been used to describe the deterioration of optronic devices from the history of data collected by HUMS (Health & Usage Monitoring Systems) [Baysse *et al.* 2012].

Examples of applications in water resources engineering concern the modelling of river inflows [Parent *et al.* 1991], lake inflows [Duckstein & Bogardi 1979], water supply reservoir states [Vogel 1987] or propagation of pollutants in water courses [Zhang & Dai 2007].

In biomedical survey, Markov chains can model the health condition of patients affected by infectious or viral diseases [Sonnenberg & Beck 1993, Gentleman *et al.* 1994, Faissol *et al.* 2009]. These models are also applied to capture-recapture problems [Dupuis 1995, Dupuis & Schwarz 2007], used to describe the dynamics of an animal population.

As a last example, the financial world makes a wide use of first-order Markov transition matrices to explain a number of phenomena like economic cycle switching [Diebold *et al.* 1994], migration of credit ratings [Jones 2005, Fuertes & Kalotychou 2007] or loan defaults [Grimshaw & Alexander 2011].

1.1 Different data structures

In an ideal framework, the data $\boldsymbol{z}$ consist in m time series of observed states for m identical individuals (systems) Σ that are assumed independent. If no data is missing, the estimation of $\boldsymbol{\theta}$ is relatively straightforward. In many applied problems, however, part of data is missing. Such problems can often be divided in two classes.

(i) We call an *incomplete sequence problem* the estimation of $\boldsymbol{\theta}$ when $\boldsymbol{z}$ are observed trajectories of states:

$$\begin{array}{ccccc} z_{(1,1)} & \bullet & \cdots & z_{(1,T-1)} & z_{(1,T)} \\ \bullet & z_{(2,2)} & \cdots & z_{(2,T-1)} & z_{(2,T)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z_{(m,1)} & \bullet & \cdots & z_{(m,T-1)} & \bullet, \end{array}$$

[2] Notice that in the (certainly not exhaustive) review of applications we provide here are highlighted cases in which more or less complex "Markov schemes" are used to model "systems" that randomly pass from one state to another. This summary of examples includes also studies involving more complex statistical models than the stationary discrete Markov chain, this chapter is mostly focused on.

containing random missing items (random successions of unknown states symbolized by "•"), assuming the initial state is known. This occurs typically when the m individuals are checked at deterministic times $t = 1, \ldots, T$, independently from $\boldsymbol{\Lambda}$, as noted by [Dupuis 1995], or when the survey of all individuals at the same time is impossible (e.g. only a given proportion of the machineries can be inspected simultaneously, in order to avoid stopping the industrial production).

(ii) We call an *aggregate data problem* the estimation of $\boldsymbol{\theta}$ when the sequential data $\boldsymbol{z}$ are reduced to the numbers of individuals $n_i(t)$ being in a given state s_i at a given time t (i.e. $n_i(t) = \sum_{j=1}^{m} \mathbb{1}_{\{z_{(j,t)}=s_i\}}$). Such data are frequently [Gouno *et al.* 2011] the only ones being at disposal of the analyst, because, for instance, the full trajectories of individuals represented too much information or were not considered of primary importance during the survey process.

2 Bayesian estimation of transition probabilities

This section provides a review of Bayesian inference techniques for the estimation of the transition matrix $\boldsymbol{\theta}$ under the obvious conditions:

$$0 \leq \theta_{i,j} \leq 1, \quad \sum_{j=1}^{r} \theta_{i,j} = 1. \tag{III.2}$$

This estimation problem has thus $r(r-1)$ degrees of freedom. As stated hereinbefore, we voluntarily chose a Bayesian viewpoint. Besides the more theoretical issues pointed in a number of reference works we fully agree with (in particular, we refer to [Bernardo & Smith 1994, Robert 2001, Parent & Bernier 2007, Kadane 2011]), we motivate our choice, in an industrial context, by the possibility to explicitly (and relatively easily) quantify [Girard & Parent 2004], via predictive simulation, the uncertainty affecting some quantities of practical interest for the reliability engineer (e.g. the probability for the system to be in a failure state for a given time t, or the mean time before the system reaches one of the failure states).

Moreover, from a strictly computational point of view, the Bayesian framework allows here to deal with some issues that can be quite burdensome in frequentist inference, without any particular additional difficulty. These include the intractability of the likelihood expression in missing data schemes, the respect of constraints III.2 [Lee *et al.* 1968], the difficulty to obtain a probability distribution for the estimators $\hat{\boldsymbol{\theta}}$, which requires using (possibly costly) bootstrap approaches [Fuh 1993]. Besides, the validity of such distributions remains usually asymptotic. Finally, even if this point has not been investigated, using an informative prior could maybe solve some identifiability problems [Allman *et al.* 2009], when the dimension of $\boldsymbol{\theta}$ is high and/or data are poorly informative [Puolamäki & Kaski 2009].

A convenient, and quite natural, choice for the prior distribution of the transition matrix $\boldsymbol{\theta}$ is the product of r independent Dirichlet distributions, one for each row $\boldsymbol{\theta}_i$ of $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}_i \sim \mathrm{Dir}(\boldsymbol{\gamma}_i) \;\text{ i.e. }\; \pi(\boldsymbol{\theta}_i) \propto \prod_{j=1}^{r} \theta_{i,j}^{\gamma_{i,j}-1}. \tag{III.3}$$

Actually, as the Dirichlet density is null outside the standard $(r-1)$-simplex, it is particularly suited as a prior distribution of probabilities vectors, that must fulfil conditions III.2. Another well-known rationale for choosing a Dirichlet prior is that it can be seen as the reference posterior for a multinomial parameter given some virtual data of state-occupancy,

whose sizes $\gamma_{i,j} - 1$ can be interpreted as measures of the prior's strength [Minka 2003]. However, in absence of precise expert opinion in the remainder of this chapter, uniform priors ($\gamma_{i,j} = 1, \forall i,j$) were used, as also recently recommended by [Tuyl *et al.* 2009], based on symmetry requirements of posterior predictive distributions.

2.1 Complete sequence problem

Transition probabilities estimation can easily be performed when complete states time-series (often alternatively called *panel data*) are available for the m individuals. The estimation is based on the calculation, for every couple of states (s_i, s_j), of the number of observed one-step transitions from state s_i to state s_j:

$$w_{i,j} = \sum_{t=1}^{T} \sum_{k=1}^{m} \mathbb{1}_{\{z_{(k,t-1)}=s_i, z_{(k,t)}=s_j\}}. \tag{III.4}$$

Full data likelihood can be written as a function of the sufficient statistics $w_{i,j}$ by observing that conditional on the row vector $\boldsymbol{\theta}_i = (\theta_{i,1} \ldots \theta_{i,r})$, the vector $\boldsymbol{w}_i = (w_{i,1} \ldots w_{i,r})$ is multinomial with parameters $\boldsymbol{\theta}_i$ and $\sum_{j=1}^{r} w_{i,j}$. Therefore, the likelihood $\mathcal{L}(\boldsymbol{z}|\boldsymbol{\theta})$ can be written as the product of r multinomial terms:

$$\mathcal{L}(\boldsymbol{z}|\boldsymbol{\theta}) = \prod_{i=1}^{r} \binom{\Sigma_j w_{i,j}}{w_{i,1} \ldots w_{i,r}} \theta_{i,1}^{w_{i,1}} \ldots \theta_{i,r}^{w_{i,r}}. \tag{III.5}$$

In a Bayesian framework, the estimation of transition probabilities given complete sequences is straightforward. The inference problem consists in computing the posterior probability distribution of model parameters $\pi(\boldsymbol{\theta}|\boldsymbol{z})$ by updating the prior distribution $\pi(\boldsymbol{\theta})$ conditional to the observed data $\boldsymbol{z}$, through the Bayes formula:

$$\pi(\boldsymbol{\theta}|\boldsymbol{z}) = \frac{\mathcal{L}(\boldsymbol{z}|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})}{\int_{\Omega} \mathcal{L}(\boldsymbol{z}|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})\,d\boldsymbol{\theta}}, \tag{III.6}$$

where Ω denotes the set of all possible values of $\boldsymbol{\theta}$. From Equations III.5 and III.3, it can be seen that the prior of $\boldsymbol{\theta}$ is conjugate, i.e. the posterior distributions of the $\boldsymbol{\theta}_i$'s are also Dirichlet distributions, with parameter vectors equal to $(\gamma_{i,1} + w_{i,1}, \ldots, \gamma_{i,r} + w_{i,r})$. This is the well known Dirichlet-multinomial model.

2.2 Incomplete sequence problem, ignorable DCM

In the most general case of incomplete sequences problem, the estimation problem turns out to be more complicated. Throughout this study, we mostly consider the case where the Data Collection Mechanism (DCM) is ignorable, which means, in practice, that it can be neglected in the statistical data analysis. Besides simplicity purposes, this choice is essentially motivated by the framework and the background of our study, which is reliability analysis. Some elements about the more general cases of non-ignorable DCM will be provided in the next section.

Let $x_{(k,t)}$ be an auxiliary binary variable (missingness indicator) which is one if the observation is missing, zero if the state has been observed. The DCM is described by a complementary statistical model specifying $\mathbb{P}\left[x_{(k,t)}|\boldsymbol{z}, \boldsymbol{z}_{\text{mis}}, \boldsymbol{\eta}\right]$, i.e. the probability for an observation to be missing, depending on observed and unobserved data and (possibly) some other parameters $\boldsymbol{\eta}$.

Fulfilling two conditions is sufficient for ignorability [Gelman *et al.* 2004]: the first one states the independence between the parameters of the DCM and the main model (here $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$ respectively), the second one asserts that the probability that an observation is missing does not depend on missing data (MAR: *missing at random* condition). The

first condition is generally easily checked, while the second one highly depends on the context of the statistical study. For instance, in capture-recapture experiments the probability of recapture may depend or not on the state of the individual (e.g. younger animals can be more easily captured than older ones). In longitudinal medical surveys the health state of a patient can prevent him from going to a periodical visit (e.g. in case he/she is hospitalized). In an industrial reliability framework, and in particular in the specific context of EDF, the presence of missing data is mainly due to the impossibility of simultaneously surveying the whole population of components for cost or system availability reasons. This motivates our choice to mainly focus on ignorable DCM situations.

Let us now come back to our estimation problem. In incomplete sequences problems, the likelihood has a highly complex expression. It is the product of m terms which are the probabilities to observe each one of the m sequences. Whilst writing the term related to an incomplete sequence, one must consider all possible values of the unknown observations. For example, the probability of the sequence $(s_1, s_1, \bullet, \bullet, s_3)$ must be written by taking into account all possible three-steps paths from state s_1 to state s_3:

$$\mathbb{P}\left[s_1, s_1, \bullet, \bullet, s_3\right] \propto \sum_{i=1}^{r} \left[\theta_{1,i} \sum_{j=1}^{r} \theta_{i,j}\theta_{j,3} \right].$$

Estimation methods dealing directly with the likelihood expression may be quite tricky to perform [Deltout *et al.* 1999]. On the other hand, Bayesian inference can elegantly be performed by means of a Gibbs sampler.

This procedure is particularly adapted to the cases where the posterior distribution of model parameters would be more easily determined if data were fully observed. Missing data are considered as additional model parameters $z_{\text{mis}(k,t)}$ and, within the Gibbs sampling, an additional step is performed to simulate them, thus completing the data set. This technique is usually known as *data augmentation* [Robert & Casella 2010]. Note that Gibbs sampling may be viewed as the Bayesian mirror of Stochastic Expectation-Maximization (SEM) algorithms based on a similar mechanism [Deltout *et al.* 1999].

In our case the augmented data set, say $\boldsymbol{y}$, is the set of the completed state sequences for all individuals:

$$y_{(k,t)} = z_{(k,t)} \text{ if } z_{(k,t)} \text{ is observed}$$

and

$$y_{(k,t)} = z_{\text{mis}(k,t)} \text{ otherwise.}$$

The Gibbs sampler algorithm for the incomplete sequence problem can be viewed as a particular case of the more general method for the Arnason-Schwarz capture-recapture model [Marin & Robert 2007]. We first initialize the algorithm by arbitrarily completing state sequences. Then at each step $h = 1, 2, \ldots$, we perform the following two-step procedure:

1. drawing new parameter values, conditional on the augmented data $\boldsymbol{y}^{[h-1]}$:

$$\boldsymbol{\theta}_i^{[h]} | \boldsymbol{y}^{[h-1]} \sim \text{Dir}\left(\gamma_{i,1} + w_{i,1}^{[h-1]}, \ldots, \gamma_{i,r} + w_{i,r}^{[h-1]}\right),$$

   where $w_{i,j}^{[h-1]}$ are the sufficient statistics (III.4) evaluated from current completed sequences $\boldsymbol{y}^{[h-1]}$;

2. drawing missing data $z_{\text{mis}(k,t)}^{[h]}$ conditional to the current values $\boldsymbol{\theta}^{[h]}$ of model's parameters (data augmentation step).

This can be done by sampling from a conditional categorical distribution defined by the following probabilities:

$$\begin{cases} \mathbb{P}\left[y^{[h]}_{(k,1)} = s_j | y^{[h-1]}_{(k,2)} = s_i, \boldsymbol{\theta}^{[h]}\right] \propto \theta^{[h]}_{j,i}, \text{for } t = 1 \\ \mathbb{P}\left[y^{[h]}_{(k,T)} = s_j | y^{[h]}_{(k,T-1)} = s_i, \boldsymbol{\theta}^{[h]}\right] \propto \theta^{[h]}_{i,j}, \text{for } t = T \\ \mathbb{P}\left[y^{[h]}_{(k,t)} = s_j | y^{[h]}_{(k,t-1)} = s_{i_1}, y^{[h-1]}_{(k,t+1)} = s_{i_2}, \boldsymbol{\theta}^{[h]}\right] \propto \theta^{[h]}_{i_1,j} \cdot \theta^{[h]}_{j,i_2}, \text{otherwise.} \end{cases} \tag{III.7}$$

The computational method shown above is quite general and easy to implement. On the other hand, the more incomplete the sequences are, the more additional parameters are required and the more the data augmentation step becomes time-consuming. This issue will be illustrated later on in the example of Section 2.5. A technique to accelerate this step, consisting in simulating blocks of consecutive missing data instead of one datum at a time, is proposed by [Dupuis & Schwarz 2007].

A particularly interesting case of incomplete sequence problem occurs when each individual is observed just once over the observation period. This can happen in industrial reliability when the data come from the first survey of operating machines, as in the real-world example of Section 6, or from destructive controls (Section 7). Then let t_k (with $1 < t_k < T$) be the time when the individual k has been observed and s_j be the observed state. The state sequences takes the form:

$$\bullet, \ldots, \bullet, s_j, \bullet, \ldots, \bullet.$$

In that case, it can be shown (proof in Appendix, page 156) that the likelihood $\mathcal{L}(\boldsymbol{z}|\boldsymbol{\theta})$ has the general expression:

$$\mathcal{L}(\boldsymbol{z}|\boldsymbol{\theta}) \propto \prod_{t=1}^{T} \prod_{j=1}^{r} p_j(t)^{n'_j(t)}. \tag{III.8}$$

In the formula above, $p_j(t)$ is the unconditional probability for the system to be in state s_j at time t and $n'_j(t) = \sum_{i=1}^{m} \mathbb{1}_{\{z_{(i,t)}=s_j\}}$ is the number of times the state s_j has been observed at time t in the data sample $\boldsymbol{z}$. It has to be noticed that the expression of the likelihood depends on sufficient statistics $n'_j(t)$ and the statistical problem is equivalent to the aggregate data problem considered hereinafter. In this particular case, Bayesian estimation can be performed using the Gibbs sampler described above or the Metropolis-Hastings procedure we carry out for the aggregate data problem in Section 2.4.

2.3 Incomplete sequence problem, non-ignorable DCM

Let us now consider the more general case where DCM is non ignorable.

This problem has been studied in detail (cf. chapters 6-10 of [Little & Rubin 1987]) in particular within the framework of longitudinal medical surveys: indeed, for different reasons, patients can leave the study permanently (dropout) or temporarily (intermittent missing). Using the same notation as in the previous subsection, let $\boldsymbol{y}_k$ be a complete data sequence for the individual k (while $\boldsymbol{z}_k$ denotes the actually observed sequence). The different ways for coping with MNAR (*missing not at random*) observations rely, from a technical point of view, on the way the *full-data* likelihood

$\mathcal{L}(\boldsymbol{y}_k, \boldsymbol{x}_k|\boldsymbol{\theta}, \boldsymbol{\eta})$ is factorized. Three types of factorization are usually proposed:

$$\begin{cases} \mathcal{L}(\boldsymbol{y}_k|\boldsymbol{x}_k, \boldsymbol{\theta}) \cdot \mathcal{L}(\boldsymbol{x}_k|\boldsymbol{\eta}) \text{ (pattern mixture model)}, \\ \mathcal{L}(\boldsymbol{y}_k|\boldsymbol{\theta}) \cdot \mathcal{L}(\boldsymbol{x}_k|\boldsymbol{y}_k, \boldsymbol{\eta}) \text{ (selection model)}, \\ \int \mathcal{L}(\boldsymbol{y}_k|\boldsymbol{x}_k, \boldsymbol{v}_k, \boldsymbol{\theta}) \cdot \mathcal{L}(\boldsymbol{x}_k|\boldsymbol{v}_k, \boldsymbol{\eta}) \cdot f(\boldsymbol{v}_k|\psi)\, d\boldsymbol{v}_k \text{ (shared parameter model)}. \end{cases} \tag{III.9}$$

The formulations above can be complexified, by considering the influence of covariates in both the main and the missingness models.

In the pattern mixture framework [Little 1993], the analyst models the conditional distribution of the observable outcome, given its observation pattern, and the distribution of the different patterns. As a matter of fact, the data are stratified (each pattern determines a stratum) and the main parameters $\boldsymbol{\theta}$ are estimated in each stratum.

The selection factorization, first introduced by [Rubin 1976], instead, focuses on the dependence between the missingness and the actual value of the observable variable (in our case the state of the individual). This scheme explicitly copes with the distribution of the complete data $\boldsymbol{y}$ conditional on the main parameter of the model, here $\boldsymbol{\theta}$. The DCM parameters $\boldsymbol{\eta}$ are easy to interpret and provide additional valuable information to the analyst.

In the shared-parameter scheme [Wu & Carroll 1988], the missing mechanism is indirectly related to the observable variable through a latent variable $\boldsymbol{v}$, depending on some additional parameters ψ.

In the particular framework of the estimation of transition probabilities, [Cole *et al.* 2005] considered categorical quality-of-life data in cancer clinical trials, using a selection factorization. Transition probabilities $\theta_{i,j}$ and missingness probabilities $\eta_i = \mathbb{P}[x_{(k,t)} = 1|y_{(k,t)} = s_i]$ both depend on observable covariates.

The Arnason-Schwarz model [Dupuis 1995, Marin & Robert 2007], also based on a selection factorization, has an elegant Bayesian solution in the case where the η_i's do not depend on covariates. In this case, a natural choice of the prior for each one of the η_i's is a Beta pdf: $\text{Be}(\alpha_i, \beta_i)$. The Gibbs algorithm for estimating the posterior of $(\boldsymbol{\eta}, \boldsymbol{\theta})$ is straightforward as, conditional on the the complete data $\boldsymbol{y}$, both posterior distributions of $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$ are explicit. The detailed description of the two steps of the algorithm (data augmentation and parameters estimation) is given in Appendix, at page 156.

2.4 Aggregate data problem

In many real-life problems, we do not follow individuals passing from state to state and the only available data for estimating transition probabilities are aggregate data $\boldsymbol{n}$, i.e. the number of individuals $n_i(t)$ being in a given state s_i at a given time t. Any track of individual trajectories is lost. That may occur in practice when a population of m individuals has been followed over an observation period but the original aim of the survey was simply having the fractions of the population in particular states. State sequences have thus been considered as raw data and discarded. Examples in sociology and population dynamics were highlighted by [Bartholomew 1973] and [Pollard 1973], among others. Applications in credit rating were recently studied by [Jones 2005].

The inference problem has been formalized by [Lee *et al.* 1968]. Conditional on the probability vector $\boldsymbol{p}(t) = \boldsymbol{p}(0) \cdot \boldsymbol{\theta}^t$, the data vector $\boldsymbol{n}(t) = (n_1(t), n_2(t), ..., n_r(t))$ is multinomial with parameters $\boldsymbol{p}(t)$ and $\sum_{j=1}^{r} n_j(t)$. The likelihood $\mathcal{L}(\boldsymbol{n}|\boldsymbol{\theta})$ can then be written as the product of T independent terms:

$$\mathcal{L}(\boldsymbol{n}|\boldsymbol{\theta}) = \prod_{t=1}^{T} \frac{\left(\sum_j n_j(t)\right)!}{\prod_{j=1}^{r} n_j(t)!} \prod_{j=1}^{r} p_j(t)^{n_j(t)}. \tag{III.10}$$

[Lee *et al.* 1968] focused on obtaining point estimates of the matrix $\boldsymbol{\theta}$ and in particular the posterior mode of $\pi(\boldsymbol{\theta}|\boldsymbol{n})$ by maximizing the product of the likelihood (Equation III.10) and r independent Dirichlet priors (Equation III.3), one for each row of $\boldsymbol{\theta}$.

In the same frequentist context, [MacRae 1977] then [Kalbfleisch & Lawless 1984] were among the main authors who developed generalized least square estimators to remedy the difficulty of the maximum likelihood estimation, because of the untractability of $\mathcal{L}(\boldsymbol{n}|\boldsymbol{\theta})$. Under mild conditions on the stationary matrix $\boldsymbol{\theta}$, [Kalbfleisch & Lawless 1984] obtained general consistency results and asymptotic $r(r-1)-$variate normality (in T and $N=\sum_{j=1}^{r} n_j(t)$) for the estimated vector $\boldsymbol{\theta}_{\text{row}}$ of entries in $\boldsymbol{\theta}$ written rowwise, i.e. $\boldsymbol{\theta}_{\text{row}}=(\theta_{1,1},\ldots,\theta_{1,r-1},\theta_{2,1},\ldots,\theta_{r,r-1})$. [Lawless & McLeish 1984] gave conditions on functions of interest for which the information loss due to aggregation is asymptotically negligible with respect to the estimation based on complete sequences. In a specific reliability framework, [Gouno *et al.* 2011] recently provided a methodology to estimate such functions of interest (e.g. survival probability, sojourn time in a state) .

In a Bayesian context, the inference problem can be solved by using a Metropolis-Hastings (MH) algorithm to construct a sample of matrices of $\Omega:\boldsymbol{\theta}^{[0]},\boldsymbol{\theta}^{[1]},\ldots,\boldsymbol{\theta}^{[h]},\ldots$, asymptotically drawn from the posterior $\pi(\boldsymbol{\theta}|\boldsymbol{n})$, by sampling at each step h a candidate vector $\boldsymbol{\theta}^{[h]^*}$ from a given distribution function $J(\cdot|\boldsymbol{\theta}^{[h-1]})$. The candidate is accepted with probability:

$$\rho(\boldsymbol{\theta}^{[h]^*}|\boldsymbol{\theta}^{[h-1]})=1\wedge\frac{\pi(\boldsymbol{\theta}^{[h]^*}|\boldsymbol{n})}{\pi(\boldsymbol{\theta}^{[h-1]}|\boldsymbol{n})}\cdot\frac{J(\boldsymbol{\theta}^{[h-1]}|\boldsymbol{\theta}^{[h]^*})}{J(\boldsymbol{\theta}^{[h]^*}|\boldsymbol{\theta}^{[h-1]})}, \tag{III.11}$$

i.e. the acceptance of the candidate is the result of a Bernoulli trial of probability $\rho(\boldsymbol{\theta}^{[h]^*}|\boldsymbol{\theta}^{[h-1]})$.

The *instrumental* density function $J(\cdot|\boldsymbol{\theta}^{[h-1]})$ allows a random exploration of the space of parameters. The convergence of the chain to the target distribution is proved for any arbitrary function $J(\cdot|\cdot)$ which satisfies mild regularity conditions [Robert & Casella 2010]. In the present case, a comfortable instrumental function is the product of r independent Dirichlet distributions $\text{Dir}(d_i\cdot\boldsymbol{\theta}_i^{[h-1]})$, where d_i is a positive (scalar) constant. This is a usual case of *controlled* MCMC [Andrieu & Thoms 2008]. As the Dirichlet density is null outside the standard $(r-1)$-simplex, all candidates drawn by the instrumental functions automatically respect constraints shown in Equation III.2.

It can easily be seen that the mean of each of the r Dirichlet instrumental densities is $\boldsymbol{\theta}_i^{[h-1]}$, i.e. the candidate matrix is sampled from a probability function which is "centered" on the last retained matrix. The variance terms of the covariance matrix, equal to $\theta_{i,j}^{[h-1]}(1-\theta_{i,j}^{[h-1]})/(d_i+1)$, depend on the shape parameters d_i which can be interpreted as tuning coefficients that rule the distance of exploration from the current state of the MCMC chain to the next proposed one.

Notice that, as the expressions of the likelihoods in Equations III.8 and III.10 are formally the same, up to a proportionality constant, the MH procedure described above can also be used in the interesting case of incomplete sequences when each individual has been observed only once. Such examples are treated in the next paragraphs.

2.5 A four-dimensional simulated case study

We compare hereby the performances of Gibbs and MH algorithms in the case where individuals are observed only once. Following a case-study from [Lee *et al.* 1968], we consider the following transition matrix:

$$\boldsymbol{\theta}_o = \begin{pmatrix} 0.6 & 0.4 & 0 & 0 \\ 0.1 & 0.5 & 0.4 & 0 \\ 0 & 0.1 & 0.7 & 0.2 \\ 0 & 0 & 0.1 & 0.9 \end{pmatrix}. \tag{III.12}$$

First, complete state sequences for $m \in \{10, \ldots, 1200\}$ individuals have been generated for $T = 20$ observation periods, under the hypothesis that at $t = 0$ the initial vector probability is $(3/4, 1/4, 0, 0)$. Then, given complete sequences, a single observation per individual has been randomly selected, thus obtaining incomplete sequences. Finally, for each m we used the Gibbs and the Metropolis-Hastings algorithms described above. The convergence has been checked using the Brooks-Gelman statistic [Brooks & Gelman 1998] computed on three parallel chains and a visual inspection of the chains. A classic rule of thumb (RT) is to suppose quasi-stationarity once the statistic stably remains under 1.1 [Brooks & Gelman 1998]. The precision in estimation was measured using the relative absolute error matrix between the elements of $\boldsymbol{\theta}_o$ and a progressive Monte Carlo posterior estimate of $\boldsymbol{\theta}$. In each case, it has been obtained by using the second half run of Metropolis-Hastings iterations and Gibbs iterations after the burn-in periods determined by Brooks-Gelman RT respectively. Parameters d_i were sampled uniformly in [100, 2500]. For a same estimation error of at most 5% per element, the CPU time observed on a 2.8 GHz CPU (Xeon) machine before the RT is fulfilled has been plotted in Figure III.1 as a function of m. Plots are smoothed over 30 repetitions of the algorithms. Clearly, the increasing number of missing data makes Gibbs less competitive than MH: after $m = 700$, conditional sampling of individuals requires more CPU time than our basic MH. The number of missing data to be simulated increases linearly with the total number m of individuals, as individuals could be observed only once throughout their lifespan. This explains the linear behaviour of the Gibbs CPU time.

The efforts of the practitioner should then concentrate on improving the mixing of Gibbs and MH algorithms to diminish their burn-in period. The development of acceleration methods has been the subject of a large number of works, reviewed in [Gilks & Roberts 1996, Mira & Sargent 2003, Gentle *et al.* 2004]. Techniques such as blocking, which consists in updating multivariate blocks of (often highly correlated) parameters [Roberts & Sahu 1997], were shown to be efficient to accelerate Gibbs algorithms in conjugate models [Ischwaran & James 2001, Accoto 2009], although their implementation often remains case-specific [Sargent *et al.* 2000] and can sometimes slow the sampler's convergence [Roberts & Sahu 1997]. Alternatively, the multi-move Gibbs sampler [Carter & Kohn 1994], which was developed for Markov switching state-space models, proved to be more efficient than the single-move Gibbs sampling.More recently, cheaper approximations of the Gibbs sampler using best linear predictors have been carried out [Nott & Kohn 2005].

3 Accelerating the MH algorithm using adaptive approaches

Heuristically, implementing an adaptive MCMC consists in sequentially tuning the transition kernel using the knowledge of past iterations, in an automated way during the simulation, in order to improve the mixing rate [Andrieu & Thoms 2008]. In the particular case of our class of MH algorithms, this means modifying the product of Dirichlet densities chosen as the instrumental distribution J for the MH algorithm introduced in Section 2.4. Each successive instrumental distribution

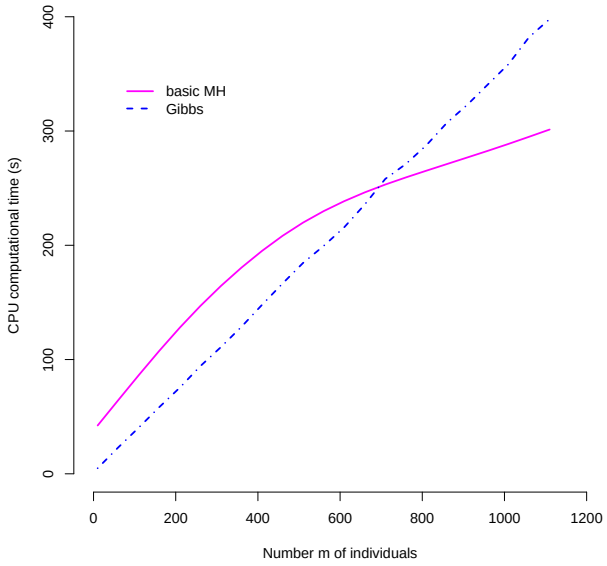


Figure III.1 – Case study of [Lee *et al.* 1968]. Mean CPU time needed to reach quasi-stationarity (in the sense of the Brooks-Gelman rule of thumb) as a function of the number m of individuals (one individual being associated to a single true observation). Data have been generated according to the four-state transition matrix (III.12).

is ideally selected such that parallel sampling can explore a large part of the parameter space, especially in the first steps of the algorithm.

Recently, a rich literature has been dedicated to these approaches, and is especially focused on the preservation of the ergodicity of the adaptive chains towards the stationary distribution, which is not automatically ensured by automated tunings. Seminal works on this subject are due to [Roberts & Rosenthal 2007, Roberts & Rosenthal 2009], as well as [Andrieu & Moulines 2006, Andrieu & Atchadé 2007, Andrieu & Thoms 2008]. These theoretical works also led to interesting software developments [Rosenthal 2007, Vihola 2010].

Assuming Γ_i are indices chosen in some collection $\mathcal{Y}_i$ based on past algorithm output, we denote by K_{Γ_i} the transition kernel updating $\boldsymbol{\theta}^{[i]}$ to $\boldsymbol{\theta}^{[i+1]}$:

$$K_{\Gamma_i}\left(\boldsymbol{\theta},\boldsymbol{\theta}'\right) \quad = \quad \rho_{\Gamma_i}\left(\boldsymbol{\theta},\boldsymbol{\theta}'\right) J_{\Gamma_i}\left(\boldsymbol{\theta}'|\boldsymbol{\theta}\right) + \int \left(1-\rho_{\Gamma_i}\left(\boldsymbol{\theta},\boldsymbol{\epsilon}\right)\right) J_{\Gamma_i}\left(\boldsymbol{\epsilon}|\boldsymbol{\theta}\right) d\boldsymbol{\epsilon}\; \delta_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}'\right), \tag{III.13}$$

where $\delta_{\boldsymbol{\theta}}$ is the Dirac measure in $\boldsymbol{\theta}$ and

$$\rho_{\Gamma_i}\left(\boldsymbol{\theta},\boldsymbol{\theta}'\right) \quad = \quad 1 \wedge \frac{\pi(\boldsymbol{\theta}'|z) J_{\Gamma_i}\left(\boldsymbol{\theta}|\boldsymbol{\theta}'\right)}{\pi(\boldsymbol{\theta}|z) J_{\Gamma_i}\left(\boldsymbol{\theta}'|\boldsymbol{\theta}\right)}.$$

Basically, the ergodicity and stationarity properties of an adaptive MH algorithm can be ensured if the amount of adapting progressively *diminishes*, in the sense that the kernel parameters are modified by smaller and smaller quantities, or if the probability of adaptation ρ_{Γ_i} decreases towards zero as $i \to \infty$ (Theorem 5 in [Roberts & Rosenthal 2007]). In the framework considered here, such adaptations could be based on eliciting vanishing adaptations for the parameters $(d_i)_{1\leq i\leq r}$. These approaches would however be limitative since each d_i characterizes the marginal distribution of row i, hence they do not explore the correlations between the rows. Therefore, the approach proposed here focuses on this particular aspect.

In the following, assuming we are at step $h > 1$ of the MH algorithm, we propose two ways of building an adaptive

instrumental distribution $\boldsymbol{\theta}^{[h]^*} \sim J_h$ (denoting $J_{\Gamma_h} = J_h$ in the following for simplicity) taking advantage of a $\sigma-$algebra $\mathcal{F}_{h-1}$ generated by the succession of sampled parameter matrices $\boldsymbol{\theta}^{[0]}, \ldots, \boldsymbol{\theta}^{[h-1]}$. Both using a (small) fixed number p of basic MH iterations, these approaches explore correlations between the rows in the instrumental sampling.

In our first approach (DCS-MH), we attempt to summarize the correlations within $(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_r)$ by simply capturing the correlations between the diagonal elements of $\boldsymbol{\theta}$.

In our second method (RCS-MH), we generalize the first method replacing the $r-$vector of diagonal elements by r elements whose position is randomly sampled within each vector $\boldsymbol{\theta}_i$. Doing so, we hope to capture more efficiently the dependency between the $\boldsymbol{\theta}_i$ and accelerate the DCS-MH algorithm.

Diagonal correlated sampling (DCS-MH)

At iteration $h \gg p$ (large enough):

1. denote $\{\tilde{\boldsymbol{\theta}}^{[1]}, \ldots, \tilde{\boldsymbol{\theta}}^{[p]}\}$ the set of last p non-identical sampled matrices in the chain $(\boldsymbol{\theta}^{[0]}, \ldots, \boldsymbol{\theta}^{[h-1]})$;
2. for $i = 1, \ldots, r$
   - **(i)** denote $\tilde{\boldsymbol{\theta}}_{i,i} = (\tilde{\theta}_{i,i}^{[1]}, \ldots, \tilde{\theta}_{i,i}^{[p]})$ the $p-$vector of replicates of the $i-th$-diagonal element;
   - **(ii)** compute $\mathbf{u_i} = \hat{F}_i(\tilde{\boldsymbol{\theta}}_{i,i})$ where $\hat{F}_i$ is the empirical marginal cdf of $\tilde{\boldsymbol{\theta}}_{i,i}$;
3. estimate the Pearson correlation $\mathbf{R^{[h]}}$ of $(\mathbf{u_1}, \ldots, \mathbf{u_r})$;
4. sample a candidate vector $\boldsymbol{\theta}_{\mathtt{diag}}^{[h]^*}$ of diagonal elements $\theta_{1,1}^{[h]^*}, \ldots, \theta_{r,r}^{[h]^*}$ using:
   - **(i)** a Gaussian copula, the parameter of which is $\mathbf{R^{[h]}}$;
   - **(ii)** Beta marginal distributions $\mathrm{Be}\left(d_i \cdot \theta_{i,i}^{[h-1]}, d_i\left(1-\theta_{i,i}^{[h-1]}\right)\right)$;
5. for $i = 1, \ldots, r$
   - **(i)** sample $\theta_{i,1}^{[h]^*}, \ldots, \theta_{i,i-1}^{[h]^*}, \theta_{i,i+1}^{[h]^*}, \ldots, \theta_{i,r}^{[h]^*}$ from:

$$\mathrm{Dir}\left(\frac{\theta_{i,1}^{[h-1]}}{1-\theta_{i,i}^{[h-1]}} \cdot d_i, \ldots, \frac{\theta_{i,i-1}^{[h-1]}}{1-\theta_{i,i}^{[h-1]}} \cdot d_i, \frac{\theta_{i,i+1}^{[h-1]}}{1-\theta_{i,i}^{[h-1]}} \cdot d_i, \ldots, \frac{\theta_{i,r}^{[h-1]}}{1-\theta_{i,i}^{[h-1]}} \cdot d_i\right);$$

   - **(ii)** for $j \neq i$, renormalize each $\theta_{i,j}^{[h]^*}$ by multiplying with $1-\theta_{i,i}^{[h]^*}$.

Randomized correlated sampling (RCS-MH)

At iteration $h \gg p$ (large enough):

1. same as step 1 in DCS-MH;
2. sample (with replacement) a $r-$vector $I \in \{1, \ldots, r\}$ of random indicators;
3. for $i = 1, \ldots, r$
   - **(i)** denote $\tilde{\boldsymbol{\theta}}_{i,I_i} = (\tilde{\theta}_{i,I_i}^{[1]}, \ldots, \tilde{\theta}_{i,I_i}^{[p]})$ the $p-$vector of replicates of the $(i,I_i)-th$ matrix element;
   - **(ii)** compute $\mathbf{u_i} = \hat{F}_i(\tilde{\boldsymbol{\theta}}_{i,I_i})$ where $\hat{F}_i$ is the empirical marginal cdf of $\tilde{\boldsymbol{\theta}}_{i,I_i}$;
4. same as step 3 in DCS-MH;

5. sample a candidate vector $\boldsymbol{\theta}^{[h]^*}_{\texttt{rand}}$ of elements $\theta^{[h]^*}_{1,I_1},\ldots,\theta^{[h]^*}_{r,I_r}$ following the same main idea as in DCS-MH method;

6. For $i=1,\ldots,r$

   (i) sample $\theta^{[h]^*}_{i,1},\ldots,\theta^{[h]^*}_{i,I_i-1},\theta^{[h]^*}_{i,I_i+1},\ldots,\theta^{[h]^*}_{i,r}$ from:

$$\texttt{Dir}\left(\frac{\theta^{[h-1]}_{i,1}}{1-\theta^{[h-1]}_{i,I_i}}\cdot d_i,\ldots,\frac{\theta^{[h-1]}_{i,I_i-1}}{1-\theta^{[h-1]}_{i,I_i}}\cdot d_i,\frac{\theta^{[h-1]}_{i,I_i+1}}{1-\theta^{[h-1]}_{i,I_i}}\cdot d_i,\ldots,\frac{\theta^{[h-1]}_{i,r}}{1-\theta^{[h-1]}_{i,I_i}}\cdot d_i\right);$$

   (ii) for $j\neq i$, renormalize each $\theta^{[h]^*}_{i,j}$ by multiplying with $1-\theta^{[h]^*}_{i,I_i}$.

In our experiments, we used a Gaussian copula to sample the new diagonal parameters, mainly because of its symmetric properties and its simplicity of calibration using a correlation matrix $\mathbf{R}$ [Marshall & Olkin 1988]. Note that one has to consider and check up with great care the p previously simulated matrices $\{\tilde{\boldsymbol{\theta}}^{[1]},\ldots,\tilde{\boldsymbol{\theta}}^{[p]}\}$ to make sure that a robust empirical estimator of $\mathbf{R}$ can be defined, in the sense that its Cholesky decomposition is numerically stable during the sampling process [Marshall & Olkin 1988]. The *condition number* can be used to do so [El Ghaoui 2002]. Conditionally on correlated sampled parameters, Dirichlet distributions appear necessary to get coherent instrumental sampling of remaining elements within each row vector $\boldsymbol{\theta}^{[h]^*}_i$.

For a more general introduction to copulas, see for instance [Nelsen 2006] or [Genest & Favre 2007], as well as [Genest *et al.* 2006, Kim *et al.* 2007] for more specific issues about copulas fitting.

Theoretical behaviour. Despite the large amount of existing work aiming to simplify the conditions ensuring ergodicity and stationarity of the target distribution [Nott & Kohn 2005, Roberts & Rosenthal 2007, Roberts & Rosenthal 2009, Atchadé *et al.* 2011], theoretical descriptions of kernels based on Dirichlet products compounded with Gaussian copulas turn out to be technically complex, and their study deserves a specific work which remains outside the scope of this chapter. Since our primary aim is to assess the interest of exploring the correlations between the rows of $\boldsymbol{\theta}$, we adopt the simplest approach of a *finite sampling scheme* when choosing J, as proposed by [Roberts & Rosenthal 2007]: given a time $\tau<\infty$, $J_{\Gamma_n}=J_{\Gamma_\tau}$ for any $n\geq\tau$. Here, this approach is carried out at each sweep of the algorithm after a given mixing period, selecting the final J_{Γ_τ} as the basic product of Dirichlet's described hereinbefore. In substance, τ has the sense of an exploration time, and in practice is selected as the minimum time between the time required for a fixed number of iterations and the time until the Brooks-Gelman RT is fulfilled.

Nonetheless, this explorative study fits into recent schemes shared by several authors, who tested copula-based methods to improve the efficiency of their sampling algorithms. In their seminal work on the optimization of the adaptation, [Haario *et al.* 2001] considered Gaussian copula instrumental distributions calibrated over the full past of the chains. See [Andrieu & Thoms 2008] for a review of this particular major field of adaptive MCMC. [Strid *et al.* 2010] used the sampling history to continuously calibrate a $t-$copula proposal distribution, in order to sample from dynamic stochastic equilibrium models. Finally, [Craiu 2011] used products of bivariate copulas to tune MCMC during an initialization period only, in the same spirit as the finite sampling approach used in the present paper.

Illustration. Continuing the four-dimensional simulated example from Section 2.5, we applied the DCS-MH and RCS-MH methods with $p=30$, still augmenting the number m of individuals and using three parallel chains per experiment.

Parameters d_i remain similarly sampled at each iteration. Results are smoothed over 50 similar runs of algorithms. The comparison of Gibbs and MH burn-in periods in Figure III.2, in the sense of the Brooks-Gelman RT, illustrates the improvement yielded by RCS-MH. On the other hand, in this case DCS-MH performs worse than basic MH and even Gibbs sampler. As we could expect, RCS-MH does clearly better than DCS-MH because of its widest exploration of the parameter space. RCS-MH strongly beats Gibbs even for relative low numbers of individuals.

The poor performance of DCS-MH is due to the computational cost of the selection of p past matrices $\{\tilde{\boldsymbol{\theta}}^{[1]}, \ldots, \tilde{\boldsymbol{\theta}}^{[p]}\}$ sufficiently different to allow for a robust Cholesky inversion. This cost clearly increases with the progression towards stationarity since sampled matrices become more and more similar and many among them must be rejected in the calibration task of the instrumental distribution. The RCS-MH algorithm suffers of course from the same defect, but the much better mixing counterbalances the increase of the computational cost, with respect to the basic MH algorithm, in a significant way.

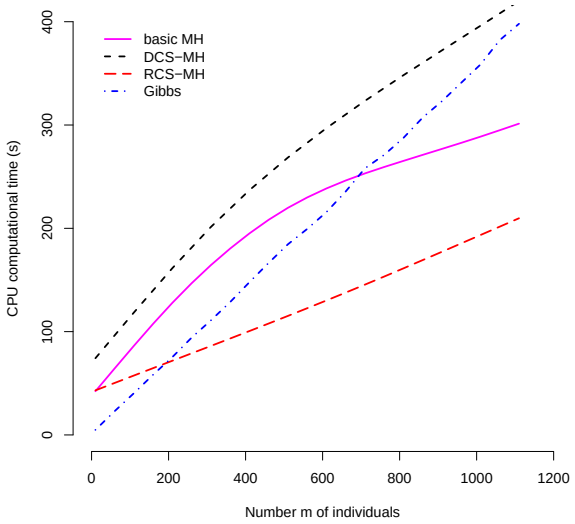


Figure III.2 – Case study of [Lee *et al.* 1968]. Mean CPU time needed to reach quasi-stationarity as a function of the number m of individuals (same simulations as Figure III.1). With respect to the Figure III.1, Gibbs and basic MH are also compared to DCS-MH and RCS-MH algorithms.

4 Numerical experiments

This section deals with simulation studies to test the potentialities of our adaptive proposals to a wide class of transition matrices commonly encountered in reliability and risk assessment (RRA). In RRA, it often occurs that the degradation of a system Σ is described using r separated states (for instance defined by a scale of crack sizes), ordered from minor defect to major failure (replacement cause). To be conservative, one may assume that potential repairs following a running failure are, at best, *as bad as old*, namely Σ remains in the same state than before the failure. In other cases, one might assume these repairs bring actually more complications than real improvement (for instance if Σ is old), so that Σ is more deteriorated after the repair than before (*worth than old* repair). See [Basile *et al.* 2007] for more details about these notions. Under a stationarity assumption, the transition matrix $\boldsymbol{\theta}$ is necessarily upper triangular, with $\boldsymbol{\theta}_r = (0, \ldots, 0, 1)$.

Simulation features. In the following experiments, we test the potentialities of Gibbs and the three MH algorithms described hereinbefore (basic, DCS-MH and RCS-MH) as a function of r. We vary the dimension r between 2 and $r_{\max}$ (in practice, we consider $r_{\max} = 6$ to remain realistic). To start with, we need a rule to sample realistic matrices with decreasing dimension:

1. denote $\boldsymbol{\theta}^{(r)}$ a $r \times r$ upper triangular matrix.

2. create $\boldsymbol{\theta}^{(r-1)}$ matrix as follows: for $i = 1, \ldots, r-1$,

$$\theta_{i,j}^{(r-1)} \quad = \quad \theta_{i,j}^{(r)} \text{ for } j = 1, \ldots, r-2$$

and

$$\theta_{i,r-1}^{(r-1)} \quad = \quad \theta_{i,r-1}^{(r)} + \theta_{i,r}^{(r)}.$$

Doing so we automatically ensure that $\boldsymbol{\theta}_{r-1}^{(r-1)} = (0, \ldots, 0, 1)$. The rationale for this construction is obviously to increase the probability of a major failure event when simplifying the model. Thus we simply need to sample $\boldsymbol{\theta}^{(r_{\max})}$ to get all other matrices considered for simulation tests. Pursuing our wish of realism, we assume that *worth than old* repairs are less probable than *as bad as old* ones. Therefore, for $i = 1, \ldots, r_{\max} - 2$ and $k = 1, \ldots, r_{\max} - i - 1$, we assume in the sampling:

$$\theta_{i,i}^{(r_{\max})} \; > \; \theta_{i,i+k}^{(r_{\max})} > \sum_{p=k+1}^{r_{\max}-i} \theta_{i,i+p}^{(r_{\max})},$$

and especially for $i = r_{\max} - 1$, $\theta_{r_{\max}-1,r_{\max}-1}^{(r_{\max})} > \theta_{r_{\max}-1,r_{\max}}^{(r_{\max})}$ to ensure a constant decreasing of values $\theta_{i,i}^{(r)}, \theta_{i,i+1}^{(r)}, \ldots, \theta_{i,r}^{(r)}$ for any $r \leq r_{\max}$. Finally, we selected matrices $\boldsymbol{\theta}^{(r_{\max})}$ for which:

$$\theta_{i,i}^{(r_{\max})} \leq \theta_{i+1,i+1}^{(r_{\max})}.$$

That models the following case: the closer to a major failure state, the better (the more cautious) the repair. Notice that we do not take into account any of our simulation constraints in the following estimation procedures, except the presence of zeros beneath the diagonal of $\boldsymbol{\theta}$ (by reducing the length of Dirichlet distributed vectors in the instrumental sampling). We consider it as a minimal knowledge assumable in real case-studies (cf. Sections 6 and 7). Finally, per simulated matrix, a complete sequence for $m = 1000$ individuals was generated for $T = 20$ observation times. As we are in the particular case of "one single observation per individual", only one observation is randomly retained in each sequence for the inference exercise.

Estimation. As in Section 2.5, each experiment for a given $r \in [3, r_{\max} = 6]$ consists in running three parallel chains for each method and monitoring them using the Brooks-Gelman statistic. Relative Euclidian errors on posterior means of matrix components (computed using 1000 iterations after a burn-in period determined by the Brooks-Gelman RT) are fixed at most at 5%, involving preliminary tests for fixing the total number of iterations. Again, parameters d_i are sampled uniformly in $[100, 2500]$. Finally, each experiment is repeated 100 times to average the results (each time a new family of matrices $\boldsymbol{\theta}^{(r_{\max})}, \ldots, \boldsymbol{\theta}^{(3)}$ being simulated).

Results. Boxplots and mean CPU times before quasi-stationarity (in the sense of the Brooks-Gelman RT) are plotted in Figures III.3 and III.4. Results obtained on the simulated example from Section 2.5 can be generalized: RCS-MH provides for all dimensions a significant improvement in mixing. Similar results have been obtained when carrying out an empirical approach to calibrate the mean acceptance rate to a standard nominal value of 50% then 25%.

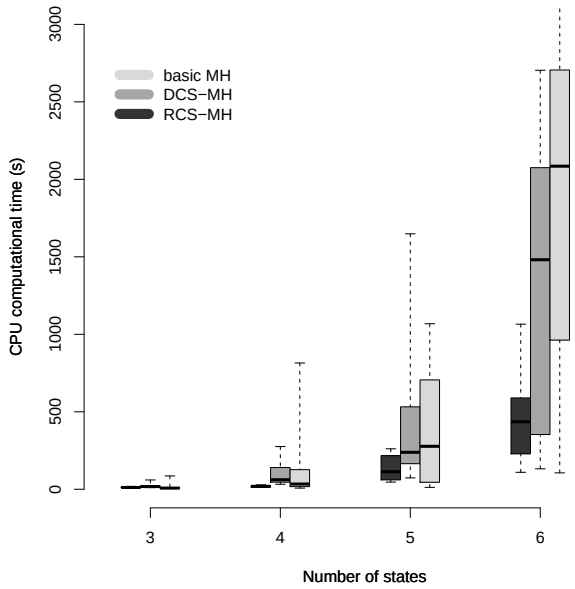


Figure III.3 – RRA case study. Boxplots of CPU times needed to reach quasi-stationarity as a function of the dimension r. Half lines indicate median and bounds indicate most extreme values. Data have been generated by upper-triangular transition matrices.

5 Discussion

5.1 Main ideas and results

This work first aims to provide a general review and technical advises about the Bayesian estimation of finite-state transition matrices $\boldsymbol{\theta}$ in discrete Markovian models under various missing data schemes, which appear to be of particular interest in several domains, especially in engineering. Actually, reliability practitioners may frequently deal with classes of upper-triangular transition matrices that have been chosen for most of the experiments presented here. Depending on the nature of available data, the practitioner may have to choose between Gibbs or Metropolis-Hastings (MH) algorithms. The time-consuming features of these algorithms, depending on the size of missing data and the dimension of the problem, appear as limiting factors in practice. Therefore, the second part of this study focuses on a first exploration of two adaptive mechanisms (DCS-MH and RCS-MH) likely to accelerate the MH algorithms.

Numerical experiments have highlighted, on this specific class of examples, that using instrumental distributions based on Gaussian copulas to account for the correlations between the rows of $\boldsymbol{\theta}$ yields a better mixing of the chains, implying a significant reduction of the computational cost. The gap with basic MH strategies, based on the independent sampling of the rows of $\boldsymbol{\theta}$, increases with the number m of individuals or the number r of states. The simplicity of the approaches

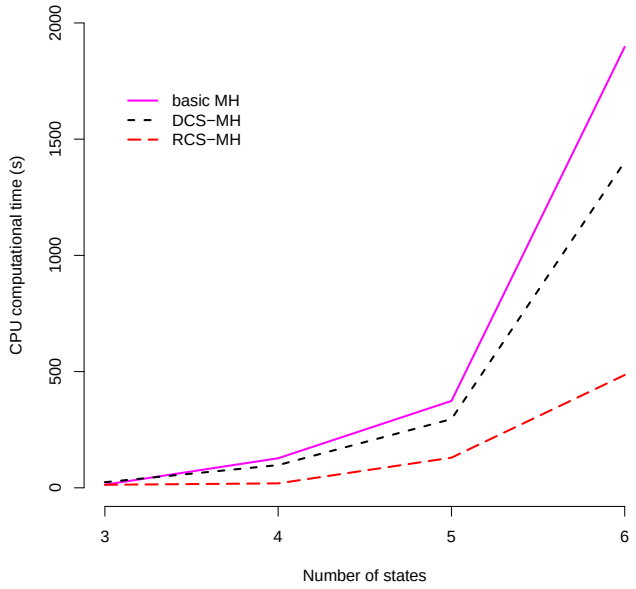


Figure III.4 – RRA case study. Mean CPU times needed to reach quasi-stationarity.

proposed here lets us think that any practitioner dealing with aggregate data could easily implement the DCS-MH and RCS-MH mechanisms and reduce the computational time.

Supplementary experiments have highlighted that the CPU time can be still diminished by using two "coarse" versions of the DCS-MH and RCS-MH mechanisms. They consist in estimating the copula parameter **R** directly from the Pearson correlation of the matrix elements, namely removing the step 2.(**ii**) in each mechanism. These coarse approaches (we call them DCS-C-MH and RCS-C-MH) have been be compared to the previous ones in Figure III.5. Here, the difference in CPU time is mainly due to the cost of empirical inversions in the DCS-MH and RCS-MH methods.

The adaptive schemes proposed here (especially the most powerful RCS-MH and RCS-C-MH), which remain only empirically studied, deserve a more specific study from both theoretical and applied viewpoints. This point is more widely discussed in the following subsection.

As a take-home message, in the most general case of incomplete data problems with several observations per individual, the Gibbs sampler based on the data augmentation technique seems to be the only possible alternative. In the particular case of a single observation per individual, the adaptive MH algorithms (and especially RCS-MH) are valid alternatives to the Gibbs sampler if the number of individuals is greater than a few hundred, say 200, and the number of states is greater than three. In low dimensional problems (two or three) the practical interest of adaptive MH methods, with respect to the simpler Gibbs sampler, is less obvious.

5.2 Directions of further research

The adaptation processes proposed here remain empirical, and theoretical studies are needed to build copula-based strategies ensuring the ergodicity and the stationarity of the chains less crudely than imposing a finite adaptation time, based on principles initiated by [Roberts & Rosenthal 2007] and [Andrieu & Moulines 2006]. Indeed, fully adaptive MCMC

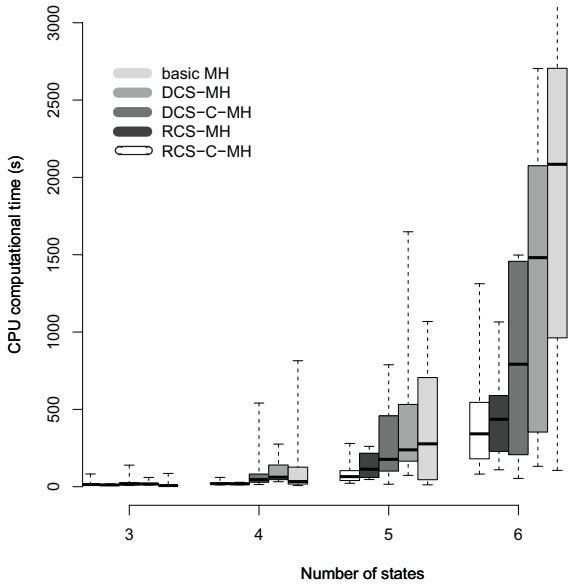


Figure III.5 – RRA case study. Boxplots of CPU times needed to reach quasi-stationarity as a function of the dimension r. Half lines indicate median and bounds indicate most extreme values. Data have been generated by upper-triangular transition matrices. The DCS-C-MH and RCS-C-MH abbrevations indicate two "coarse" versions of the DCS-MH and RCS-MH mechanisms.

should be build on *infinite* adaptations which continuously modify the choice of the transition kernel using the past values of $\boldsymbol{\theta}$ along the chains, quasi-stationarity occurring when these kernel modifications become imperceptible. These adaptations should be led on both correlated and marginal features of the matrix elements. To this first aim, future studies could focus on the mechanism of state permutation, inspired by similar ones carried out in the framework of variable selection [Nott & Kohn 2005], and on removing the strong assumption made by using a Gaussian copula to model correlations within the elements of $\boldsymbol{\theta}$. This choice can appear oversimplified since it does not take into account possible correlations between extreme values in the instrumental distribution of $\boldsymbol{\theta}$. Therefore a copula selection procedure should be carried out at different times of the adaptive chain, for instance using frequentist tests (e.g. Cramer-von Mises), based on distances between estimated and simulated copulas [Genest *et al.* 2006, Nikoloulopoulos & Karlis 2008] or Bayesian posterior odds [Huard *et al.* 2006]. As those procedures remain time-consuming in dimensions $r \geq 2$, this approach was not implemented here in this exploratory work.

Furthermore, it is necessary that such more sophisticated adaptive Metropolis-Hastings algorithms be compared in practice to refined Gibbs algorithms, evoked at the end of Section 2.5, that could benefit from the stick-breaking properties of Dirichlet distributions.

Another point of interest could be the adaptation of the methods reviewed here to the case of non stationary Markov chains. A simple way for doing this could be to stratify the data on the time t or on groups of values of t [Urakabe *et al.* 1975, Sendi *et al.* 1999]. The use of logit or proportional odds models [Cole *et al.* 2005, Grimshaw & Alexander 2011] to include also the effect of additional covariates is another perspective for this work.

6 An application to turbine cracks propagation

In the example shown hereby, a discrete Markov model has been used to describe the propagation of transverse cracks on steam turbine shafts. This phenomenon has been first observed on EDF facilities in late 90's and since then periodical non-destructive controls are made to measure crack depths. For a description of the technical problem and available survey data, see [Garnero & Montgomery 2006]. The most important identified explanatory variable is the time spent by the turbine in hot shutdown condition. For the purpose of our study, the time has been discretized in equally long intervals. Cracks depths are classified in four states $s_1 \dots s_4$ associated to growing crack lengths. The modelling of cracks growth by discrete Markov schemes is quite common, e.g. [Roh & Xi 2000].

It is worth noting that more sophisticated models can be proposed for cracks propagation, and namely the so-called "Piecewise-deterministic Markov process" (PDMP) which allow accounting for both deterministic evolution of cracks (ruled by differential equations) and stochastic jumps between deteriorated states [Azaïs *et al.* 2010, Gégout-Petit 2012].

Coming back to our simple Markov-chain model, we assume that the process is irreversible, which is physically correct as crack lengths cannot decrease. Thus, the transition matrix is upper-triangular and consequently, $\theta_{4,4} = 1$. We made the hypothesis that all turbines are in state s_1 when putting-into-service at the beginning of the study. Manufacture and acceptance controls justify this hypothesis. A set of data collected between 1998 and 2001 has been analyzed. The data come from 68 turbines from 24 EDF power plants. Each turbine is observed only once for a given value of t between 2 and 7. Given the uniformity of EDF French generation facilities (same design, operating conditions and maintenance policy for all units), we can assume that observed data are i.i.d.

The results of MCMC estimation, using the Gibbs sampler described in Section 2.2 (second half run of 10 000 iterations), are shown in Table III.1 (left). The application of the MH algorithm described above leads to the same results.

The data set has been enriched between 2001 and 2004 with new crack measures (t between 2 and 7). 38 turbines among the 68 previously observed were inspected for the second time and two for the first time. Some of the collected data are redundant: this happens when for the first and the second observation the corresponding times spent in hot shutdown condition fall into the same interval. Finally, 17 new exploitable observations can be added to the data set. The estimation of transition probabilities gives the results shown in Table III.1 (right).

	Data set 1			Data set 2		
	Mean	St. Dev.	95% CI	Mean	St. Dev.	95% CI
$\theta_{1,1}$	0.637	0.042	[0.551, 0.719]	0.655	0.040	[0.573, 0.728]
$\theta_{1,2}$	0.306	0.050	[0.208, 0.405]	0.278	0.049	[0.185, 0.374]
$\theta_{1,3}$	0.044	0.034	[0.002, 0.127]	0.056	0.036	[0.003, 0.133]
$\theta_{1,4}$	0.012	0.011	[0.000, 0.041]	0.012	0.011	[0.000, 0.041]
$\theta_{2,2}$	0.713	0.088	[0.538, 0.884]	0.774	0.075	[0.636, 0.921]
$\theta_{2,3}$	0.250	0.087	[0.079, 0.418]	0.197	0.075	[0.054, 0.341]
$\theta_{2,4}$	0.037	0.032	[0.001, 0.119]	0.029	0.024	[0.001, 0.087]
$\theta_{3,3}$	0.872	0.097	[0.627, 0.995]	0.910	0.071	[0.730, 0.996]
$\theta_{3,4}$	0.128	0.097	[0.005, 0.373]	0.090	0.071	[0.004, 0.270]

Tableau III.1 – Turbine cracks example. MCMC estimations of transition matrix $\boldsymbol{\theta}$ using the first data set (left, individuals observed only one time) and the second data set (right). Here, the bounds of the posterior 95% credibility intervals (CI) are the quantiles of probabilities 0.025 and 0.975 respectively.

We can notice that in this case the posterior variance has been very lightly reduced by incorporating the information conveyed by the new data. Given the posterior samples of transition probabilities, some quantities of practical interest in industrial reliability have been sampled: the unconditional probabilities of the four states, as a function of time, and the expected number of steps before the system reaches the absorbing state s_4. As s_4 can be interpreted as a "failure"

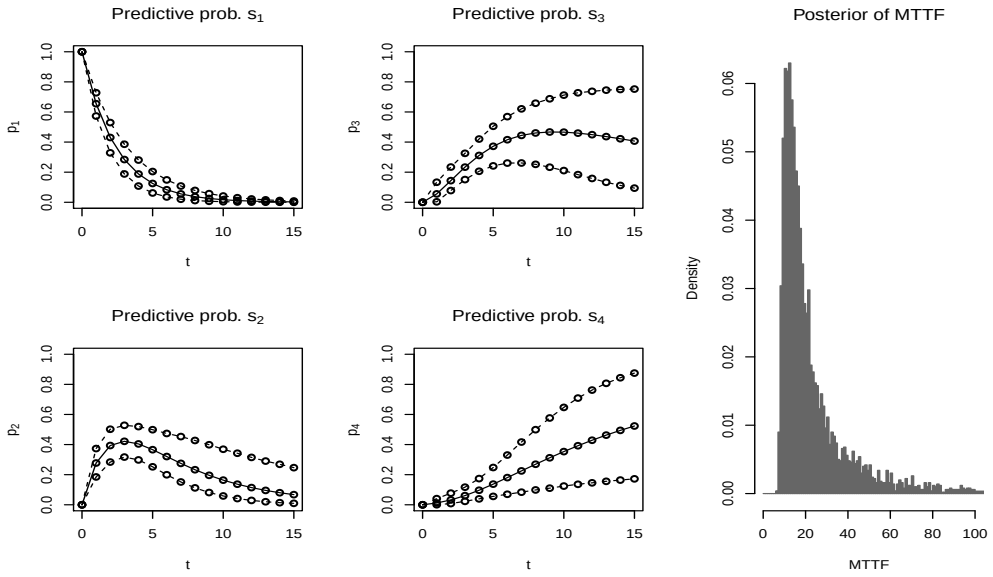


Figure III.6 – Turbine cracks example. Predictive 95% credibility intervals of state probabilities (left) and predictive distribution of the MTTF (right).

condition, the expected time to absorption is here the classical MTTF (Mean Time To Failure). Notice that here the term "failure" just means that the crack has reached a given length, arbitrarily chosen for the purposes of this study.

The calculation of state probabilities using Equation (III.1) is straightforward. To evaluate the MTTF we made use of a well known property of absorbing Markov chains (Chapter 11 in [Grinstead & Laurie Snell 1997]). If we consider the matrices:

$$\boldsymbol{\zeta} = \begin{pmatrix} \theta_{1,1} & \theta_{1,2} & \theta_{1,3} \\ 0 & \theta_{2,2} & \theta_{2,3} \\ 0 & 0 & \theta_{3,3} \end{pmatrix} \text{ and } \boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

the matrix $\boldsymbol{I} - \boldsymbol{\zeta}$ has an inverse and each component t_i^* of the row vector

$$\boldsymbol{t}^* = (1,1,1) \cdot (\boldsymbol{I} - \boldsymbol{\zeta})^{-1}$$

is the expected number of steps before absorption, given that the initial state was s_i. In our case the MTTF is then the first component of the vector $\boldsymbol{t}^*$.

Figure III.6 shows the 95% credibility intervals of the predictive state probabilities for discretized time t extended up to 15 and the histogram of 5000 samples from the predictive distribution of MTTF. Concerning state probabilities, we can notice that p_1 credibility intervals are narrower than other state probabilities as, according to our hypotheses of an irreversible process and initial state s_1, $p_1(t) = \theta_{1,1}^t$ which mean that the uncertainty over p_1 only depends on uncertainty over $\theta_{1,1}$ (and no other transition probability). The long tail in the MTTF distribution (which is even longer than shown in the figure) is due to the high values (close to 1) of the posterior distribution of $\theta_{3,3}$.

Remark. We stress that, even if the data come from real surveys, the study shown hereinbefore is given for exemplary purposes only and neither results nor methodology must be extrapolated to make any general conclusion about EDF risk assessment policies.

7 Modelling water meters deterioration

7.1 Rationale and modelling assumptions

The accuracy of a water meter is expressed by an *accuracy curve*, relying the flow rate Q to the relative error (in %):

$$\mathrm{Err}(Q) = (V_{\mathrm{mes}} - V_{\mathrm{act}})/V_{\mathrm{act}},$$

where V_{mes} and V_{act} are the measured and the actually flowed volume of water during a standard test (at constant flow rate). The dependency of the error on the flow rate is due to the fact that meters definitely act as hydraulic machines and the flow rate determines the hydraulic load on the measuring device. For low flow rates, the hydraulic load is not strong enough to let the machineries correctly operate, so that errors are high (in absolute value) and negative.

The mean of the the errors $\mathrm{Err}(Q_i)$, weighted with respect to the part of the consumption δ_i which takes place in each interval of flow rates centred in Q_i, is used to evaluate the *overall accuracy* R of the meter: $R = 1 + \sum_i \mathrm{Err}(Q_i) \cdot \delta_i$, which can be seen as the efficiency of the meter.

Water meters tend to be more and more inaccurate when getting older, which gives rise to an under-estimation of the actual customer's consumption. As a consequence, a part of the consumed water is not billed and that originates financial losses for the water utility. For well-run exploitations the losses due to unaccounted-for water are generally limited, with respect for instance to piping leaks; nevertheless, at the scale of a large distributor (as Générale des Eaux[3]) they can generate relevant financial losses. As an example, they were grossly estimated around 50 M€/year in 2004.

In addition, the loss of accuracy can also let the meter be non-compliant with respect to national standards or local regulations.

Each water distributor have a replacement policy (more or less complex) intended to cover these risks. Whatever the policy, it is obvious that cornerstone of the methodology is the mathematical model describing the degradation of meters' accuracy.

The statistical models used in this technical framework are normally regression-like: the overall accuracy is explained by continuous regressors like the age and the registered volume and/or categorical variables, e.g. meter's brand, type, location etc. The issue with these models is that: (i) they strongly depend on the assumption on the water consumption profile, i.e. the values of the δ_i's (cf. Figure III.7), that can be very different from one customer to another, and (ii) data structures shows generally a more complex dependency of the accuracy on the age t, suggesting rather a mixing of different populations of meters, the proportion of which depends on t.

Hence, a 4-state Markov model has been proposed to cope with this problem. The definition of the states, numbered from 1 to 4 in decreasing order of quality, is inspired by the ISO 4064-1 standard [ISO 1993], defining tolerance bounds of the accuracy curves for different quality-classes of meters. The model is based on the assumption of irreversible degradation; as in the case of the cracks propagation (Section 6 of this chapter), the transition matrix is upper-triangular and s_4 is an absorbing state, corresponding to stuck meters (no volume is recorded). Figure III.7 graphically sketches the assumption of the statistical model.

[3] Key figures in 2004: about 6 millions of water meters and $2{\cdot}10^9$ m^3 of water distributed in France.

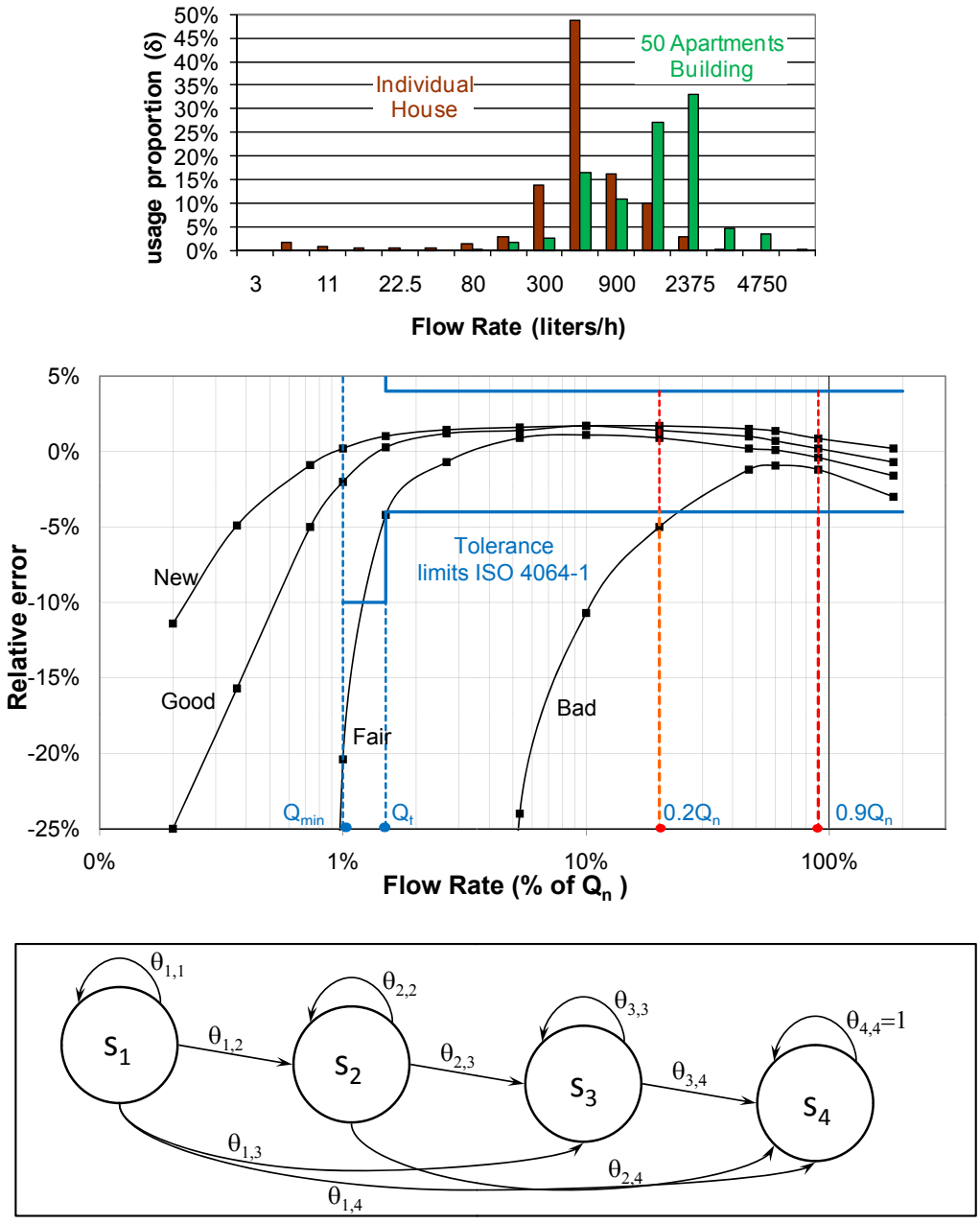


Figure III.7 – Modelling of water meters' accuracy degradation. Top: Typical *consumption histograms* (ratios δ_i of the consumption occurring at different flow rates ranges, centred in Q_i) for an individual house (red) and a 50 apartments building (green). Center: Typical error curves of in-service water meters. The *state* of a meter is determined by the respect of more or less severe tolerance bounds: *good* meters (s_1): the curve stays inside the blue bounds, *fair* meters (s_2): the error exceed tolerance limits of ISO 4064-1, but it still complies with minimal in-service standards, *bad* meters (s_3): to be replaced as soon as possible, *stuck* meters (s_4). The flow rate is expressed in % of the *nominal* flow rate, a characteristic of the meter, defined by the norm ISO 4064-1 [ISO 1993] (e.g. 1.5 m^3/h for a typical domestic meter with nominal diameter of 15 mm). As an example, it is also shown the accuracy curve of a new meter of the same type.
Bottom: Graphical representation of the 4-state irreversible Markov model.

It has to be noticed that, as in many applications, the initial state is assumed to be known: all devices start their operating lifetime in state s_1. Actually, new meters are submitted to a severe quality control, namely their curves[4] must stay within even tighter bound (5% at $Q_{\min}$ and 2.5% at Q_t and $Q_{\max}$) than in-service ones. That assures they largely fulfil conditions required for state s_1. The state model is completed by the statistical model of the overall efficiency R within each state. As, for technical reason, the overall accuracy does not exceed 1.04 and, except for stuck meters, it is well above 0, a Beta distribution bounded between 0.1 and 1.04 was chosen.

$$R|z = s_i \ \sim \ \mathrm{Be}_{[0.1,1.04]}(\alpha_i, \beta_i) \qquad i = 1,2,3 \tag{III.14}$$

Of course for $i = 4$ (stuck meter), $R = 0$.

Main data available for this study were accuracy curves of in-service meters, obtained experimentally by means of an *ad hoc* facility, owned by the water company. We insist again on the fact that the great advantage of this model is that it directly copes with the accuracy curve and not with the overall efficiency, which strongly depends on the consumption profile (the δ_i's introduced hereinbefore). The Markov state model proved to be more robust than the ordinary regression-like approaches, previously used by the company.

7.2 Different sources of information

It is interesting to highlight the following features of available data.

- Data are stratified by meter's *model* (type, brand, nominal flow rate).

- Sampled meters are never put into operation again, for practical and financial reasons. The cost of a domestic meter is comparable to the cost of the human intervention on the customer's connection, so that it would be too expensive to remove, test, repair and put into service the same meter some days later. Instead, when a meter is removed, it immediately replaced with a new one. As a main consequences, data are incomplete and, according to the classification of Section 1.1, this is an *incomplete sequence problem* in which each individual can be observed only once. As far as states s_1, s_2, s_3 are concerned, the DCM is ignorable: the missingness only depends on the fact that the test is destructive. State s_4 raises more tricky issues, as discussed in the next item of this list. In practice, we are here in the situation of data only once observed; as it is equivalent to the one of *aggregate data*, the notations concerning this kind of problem are used below (cf. Section 2.4).

- A major issue in data structure concerns stuck meters. Indeed, stuck meters are easily detected by meters reading personnel (as the recorded volume does not increase between two readings) and immediately replaced: for this reason there are very few stuck meters among the tested ones, and their proportion in the experimental database is absolutely not representative of the actual one among in-service meters.

- A second source of data was available: customers database (CDB) actually can be used to obtain information about stuck meters, because when a meter is removed the reason of the replacement is recorded in customer's billing file (among the possible options in the form used by the personnel, one can find "stuck meter"). However the use of this piece of information is tricky because this field is not systematically filled, as often considered as "not essential". Actually, customers database is intended to billing purposes, not to statistical analysis!

[4] In practice, they are tested at three characteristic values, named $Q_{\min}$, Q_t, $Q_{\max}$ defined by the norm ISO 4064-1 (cf. also Figure III.7).

7.3 Variants of the main model

Different ways to cope with the problem of stuck meters have been proposed. That has led to different variants of the main model (listed below).

- Imposing fixed stuck probability, following expert's advice [Pasanisi *et al.* 2002]; namely, the annual probability for a meter to get stuck, independently on its state (s_1, s_2 or s_3) was fixed at 0.04%. The probabilities, $p_j(t)$ in the multinomial likelihood (Equation III.10) are replaced by the conditional probabilities $p_{j\neg s_4} = p_j(t)/[1-p_4(t)]$, with $j = 1,2,3$:

$$\mathcal{L}(\boldsymbol{n}|\boldsymbol{\theta}) \propto \prod_{j=1}^{3} p_{j\neg s_4}(t)^{n_j(t)}.$$

  In a similar variant, a zero-mean Gaussian noise on the the failure probability provided by the experts was also introduced [Pasanisi 2004a].

- Exploiting the customer database (CDB) information [Pasanisi 2003, Pasanisi & Parent 2004]; the number $n_4(t)$ of recorded stuck meters of age t, among the overall population $n_{\text{cdb}}(t)$, is Binomial distributed with probability $p_4(t) \cdot p_{\text{obs}}$, the latter being the probability to actually observe the failure in the database:

$$\mathcal{L}(\boldsymbol{n}|\boldsymbol{\theta}) \propto \prod_{j=1}^{3} p_{j\neg s_4}(t)^{n_j(t)} \cdot [p_4(t) \cdot p_{\text{obs}}]^{n_4(t)} [1 - p_4(t) \cdot p_{\text{obs}}]^{n_{\text{cdb}}(t) - n_4(t)}.$$

- Working on a reduced 3-state model; definitely, the main technical question the water company is concerned with is estimating the accuracy of in-service meters, in order to define an optimal replacement policy. Stuck meters are (i) extremely rare (annual probabilities around 0.05%) and, above all, (ii) immediately recognized by the meters reading personnel. As a matter of fact, the interest of the distributor was essentially focused on improving the main model to obtain more tailored prediction of the behaviour of operating meters.
  As an example, Figure III.8 shows the 95% predictive interval of the state probabilities evaluated with respect to the 4-state model (taking into account the CDB information on stuck meters) and with respect to a 3-state model. Posterior means as well as high and low credibility bounds are extremely close and equivalent, in practice, by the distributor's viewpoint. The research effort was then put on the search for covariates, rather then improving the failure model for accurately estimating the proportion of stuck meters.

7.4 Search for explanatory variables

According to water metering specialists, many other explanatory variables (besides meter's type) can be proposed. A first group of factors concerns local effects, depending on the particular features of more or less extended geographical zones of exploitation: water hardness and temperature, casual presence of solid particles in case of works on the network etc. Another well known variable which can have an effect on meters deterioration is the annual consumption: one can easily figure that, at fixed age t, the higher the consumption, the more severe the wear of the meter (as for any other hydraulic machine).

A stratification of the available data according to all the possible explanatory factors was not possible, as most of the groups of data would have been empty or of not-significant size. We decided to define as additional covariates (i) the mean annual consumption (i.e. the ratio between the total recorded water volume and the age) and (ii) a latent geographically-based variable named *aggressiveness* of the operating conditions, which has been defined at the level of the smallest

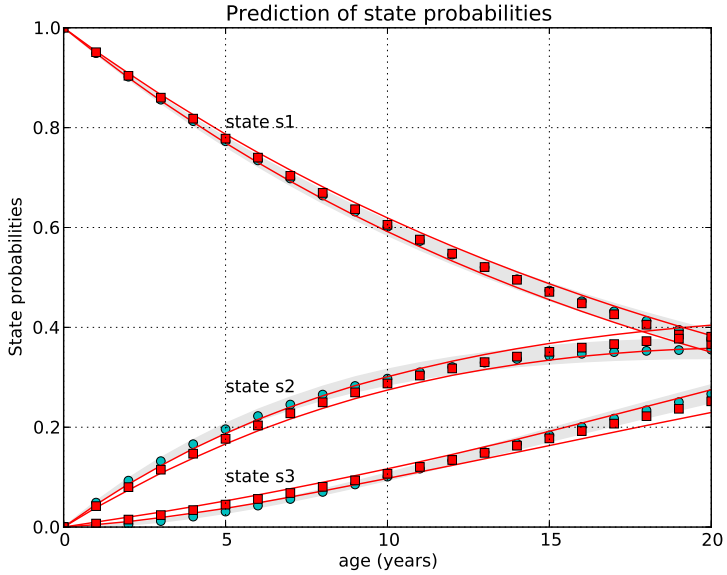


Figure III.8 – Results obtained for the same dataset of accuracy curves of domestic meters aged up to 20 years, according to a 4-state and a 3-state model respectively. The shades grey zones correspond to 95% posterior credibility intervals for the 4-state model of the conditional probabilities $p_i/(1-p_4)$ The posterior means are represented by cyan circles. The solid red lines correspond to the bounds of the same credibility intervals, evaluated according to the 3-state model. The red dots correspond to the posterior means. As one can see means and bounds of the credibility interval are very close, and most of time practically indistinguishable.

territorial unit available in both databases (meters experimental accuracies and customers data base), named *contract*. This unit covers, in practice, the perimeter of a given public service delegation contract, established between the company and the public authorities. As the public delegator could be a Municipality or a group of Municipalities, the *contracts* geographical units are very different from one another, with respect to the number of customers served and territorial extent. However, it is reasonable to assume that many of the factors affecting the deterioration of meters are relatively homogeneous inside a contract unit. In the remainder, we first give some details on the way the aggressiveness variable has been defined, then we will explain how this covariate was introduced in the main model.

Characterizing local aggressiveness. We defined a methodology, entirely based on the pragmatic will to exploit the available data in the best way possible, to characterize a certain number of zones with different aggressiveness. This method is based on the steps listed below (more details are given hereinafter).

(a) First, assessing a local aggressiveness parameter (λ_i), following the information provided by experimental accuracy curves (stratified by contracts), to contracts sufficiently represented in the experimental testing database. Three groups of aggressiveness are then defined by discretizing this parameter (i.e. by defining three intervals for the values of the λ_i's).

(b) Second, for each of the three groups, evaluate the failure ratio (number of recorded failures, divided by the overall population) provided by the customers database.

(c) Finally, for zones not adequately represented in the accuracy curves database, assigning a group of aggressiveness

under the basis of the failure ratio only, by comparing this ratio with the reference value of each of the three aggressiveness groups defined in Steps (a) and (b). The zone is assigned to the group with the closest failure ratio.

More precisely, first, the results of experimental tests were exploited. Available experimental accuracy curves were stratified by contract and age. To avoid empty and/or too small groups of data, the curves were not stratified by meter types but rather by groups of (similar) meter types. As this choice can actually introduce a significant bias in the analysis, we only considered a group of meters with very similar technical features (volumetric type with rotary piston and dry register), which correspond, in practice, to four models from two brands (which also represent the most common devices installed in France).

Ages were discretized in classes of five years: 0-4 years, 5-9 years and so on.

The statistical analysis of the accuracy curves was made with respect to a 2-state model, in which the states s_2 and s_3 were grouped (and s_4 ignored). That leads to a very simplified formulation of the model, as the upper-triangular transition matrix is completely defined by one of the two probabilities of the first row

$$\boldsymbol{\theta}(i) = \begin{pmatrix} \theta_{1,1,(i)} & 1-\theta_{1,1,(i)} \\ 0 & 1 \end{pmatrix}.$$

In the expression above, the transition matrix is now also indexed by the contract geographical unit (i). For sake of simplicity the model has been reparametrized in an exponential form, which also defines the parameter λ_i:

$$p_{1,(i)}(t) = \theta^t_{1,1,(i)} = \exp(-\lambda_i \cdot t) \quad \text{with} \quad \lambda_i = \log(1/\theta_{1,1,(i)}).$$

Notice that here $p_{1,(i)}(t)$ can be seen as a survival function, that is the lifetime of the meter in state s_1 is exponentially distributed.

Conditionally to $p_{1,(i)}(t)$ the number $n_{1,(i)}(t)$ of *good* meter from contract i of age t is binomial:

$$n_{1,(i)}(t) \sim \text{Bin}(\exp(-\lambda_i \cdot t), n_{(i)}(t)). \tag{III.15}$$

Instead of performing Bayesian inference techniques separately on each data set from a specific contract (i), the statistical model of Equation III.15 has been provided with a hierarchical structure: all the λ_i's have a common (Gamma) prior distribution. Figure III.9 presents the result of the analysis.

This analysis concerned 78 territorial units, covering approximately 20% of the operating meters. The value of λ is clearly representative of the aggressiveness of the operating conditions: the higher λ, the faster the deterioration and, hence, the more aggressive the location.

That suggested the definition of three groups of aggressiveness, each one being defined by an interval of values of λ:

$$\begin{aligned} &\text{Group A}_1 : \lambda \leq 0.145 \\ &\text{Group A}_2 : \lambda \in]0.145, 0.21[\\ &\text{Group A}_3 : \lambda \geq 0.21. \end{aligned}$$

Of course this method is empirical. Nevertheless the obtained groups of contracts showed an interesting relationship between the values of λ (obtained from the accuracy curves database) and the "failure ratios" (the ratios of meters uninstalled because they have been found stuck by the reading personnel). Actually, if one considers the same kind of meters concerned by the study above, the mean failure ratio observed in the years 2000, 2001 and 2002 was found to be higher

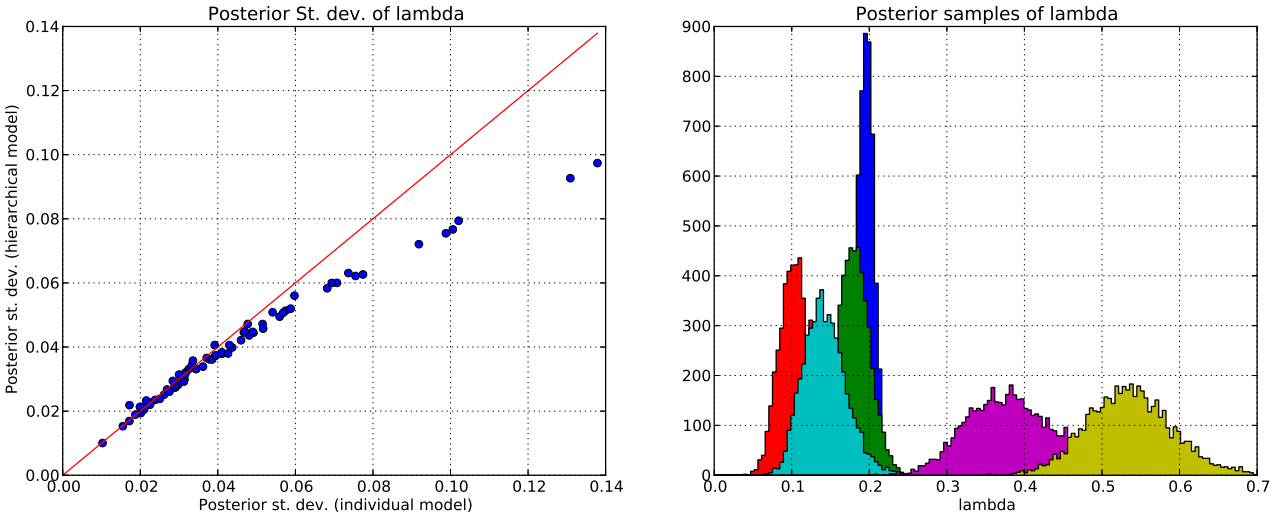


Figure III.9 – Some results of the statistical analysis of accuracy curves stratified by *contract* (exponential-binomial model of Equation III.15). Left: Posterior standard deviations of λ_i's evaluated with respect to independent (x-axis) and hierarchical (y-axis) models respectively. Most of the points stand below the first bisector, showing that the hierarchical model tends to give less dispersed results. Right: posterior samples of λ_i's corresponding to five different contracts.

and higher as one moves from Group A_1 to Group A_2 and Group A_3. More precisely, if one considers the variable:

$$h_{i,j} = \frac{1}{3}\frac{n_{s_{i,j}}}{(1-u_i)\cdot n_{\mathrm{cdb}_{i,j}}},$$

where $n_{s_{i,j}}$ is the number of meters of type j found stuck in the zone i in the three years considered, $n_{\mathrm{cdb}_{i,j}}$ is the overall number of meters of type j in the zone i and u_i is the ratio of meters removals for which the cause was not recorded in the zone i, one can find the results shown in Figure III.10.

The correspondence, highlighted by results of Figure III.10, between aggressiveness indicators obtained from the analyses of accuracy curve database and customers database suggested an empirical way to assign to zones, for which no experimental accuracy curves were available, their aggressiveness A_1, A_2 or A_3. The extremely simple procedure consists in considering for each zone the point of $[0,1]^5$ (say K_i) the coordinates of which are the ratios $h_{i,j}$ (with $j = 1,...4$) and the *overall* stuck ratio (obtained considering all the types of meters together).

This point characterizes the aggressiveness of the zone with respect to meters failure. The same characteristic points were determined for the three groups A_1, A_2, A_3. Then the three Euclidean distances between the point K_i and the points K_{A_1}, K_{A_2} and K_{A_3} were used as a measure of the *distance* between the unknown aggressiveness of the contract i and the aggressiveness of each group: the contracts were assigned to the closest group, i.e. the group the characteristic point was the closest one.

This empirical procedure allowed covering more than 50% of the overall meters population.

Taking into account the consumption level by *deforming* the transition matrix. Let us consider the other covariate: the annual water consumption. According to the experts (and after some exploratory analyses) it has been found that it could be adequately represented, for the purposes of this study, by a binary variable, indicating if the consumption is *ordinary* or *high* (in practice exceeding or not a fixed threshold). Its role with respect to metering accuracy deterioration

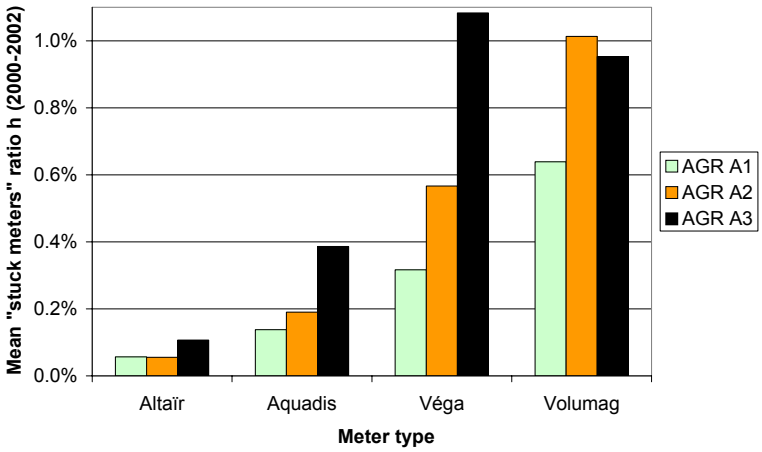


Figure III.10 – Empirical mean "failure ratio" (i.e. mean proportion of stuck meters recorded in the customers database in the years 2000-2002) for the three groups of aggressiveness A_1, A_2 and A_3 determined by the analysis of the accuracy curves thanks to the exponential-binomial model of Equation III.15. The analysis shows a quite fair agreement between the indicators coming from accuracy curves database and customers database: the higher the aggressiveness, the higher the mean failure ratio.

is known: degradation will be faster when the consumption is high. In other terms transition probabilities from states s_1 to s_2 and s_3 and from s_2 to s_3 will be higher for high consumption than for ordinary consumption.

To account for this variable, additional parameters were introduced to *deform* the transition matrix θ and accelerating transition towards deteriorated states, in case of high consumption:

$$\underbrace{\begin{pmatrix} \theta_{1,1} & \theta_{1,2} & \theta_{1,3} \\ 0 & \theta_{2,2} & \theta_{2,3} \\ 0 & 0 & 1 \end{pmatrix}}_{\text{Ordinary consumption}} \qquad \underbrace{\begin{pmatrix} \xi \cdot \theta_{1,1} & \nu \cdot \theta_{1,2} & \nu \cdot \theta_{1,3} \\ 0 & \omega \cdot \theta_{2,2} & \kappa \cdot \theta_{2,3} \\ 0 & 0 & 1 \end{pmatrix}}_{\text{High Consumption}} \tag{III.16}$$

Notice that, as the sum of the probabilities in each row must sum to 1 it is easy to show that: $\nu = (1-\xi \cdot \theta_{1,1})/(1-\theta_{1,1})$ and $\kappa = (1-\omega \cdot \theta_{2,2})/(1-\theta_{2,2})$. That is, the additional parameters are actually two: ξ and ω. A variant of this model, involving only one additional parameter (namely ξ) was also tested with good results.

Some results. The results shown in Figure III.11 and Table III.2 concern the statistical analysis of 3800 accuracy curves of the same type of meter (namely the *Volumag*). Ages were regrouped in groups of five years. Table III.2 presents posterior summaries of the probability distribution of the parameters (5000 MCMC samples).

Figure III.11 shows the posterior mean of predictive overall accuracies as a function of the age for the three groups of aggressiveness. The vertical bars give the 95% credibility intervals.

Here, the threshold separating high and low consumption has been fixed at 200 m^3/year. This value (definitely arbitrary) is close to the empirical 90^{th} percentile of the recorded consumption in the CGE customers database and is considered, in spite of the great variability of the water consumption, as a quite high value for a French dwelling, the reference value being normally assumed as 120 m^3/year (even if this assumption is questionable, cf. [Montginoul 2002] for a summary of a wide number of studies on this topic).

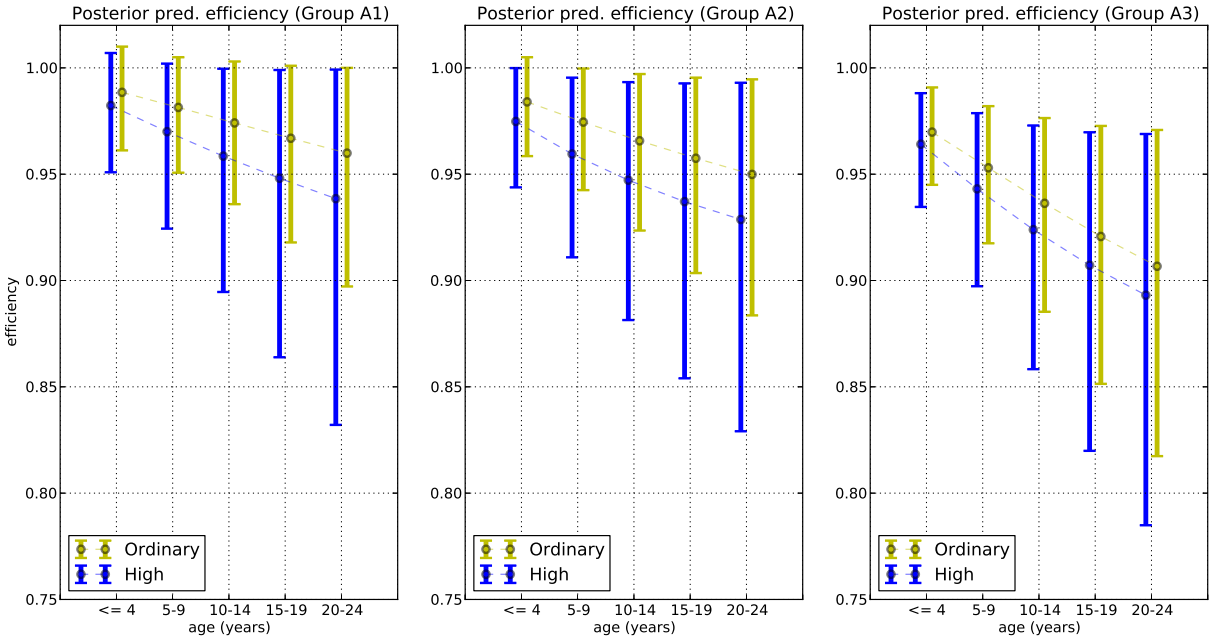


Figure III.11 – Posterior 95% credibility intervals of the overall efficiency as a function of the aggressiveness (A1, A2 and A3) and the consumption level (ordinary or high). The efficiency is obtained by mixing population of different states, in proportions depending by the age, based on results of Table III.2.

			Agr. A_1			Agr. A_2			Agr. A_3
	Mean	St. Dev.	95% CI	Mean	St. Dev.	95% CI	Mean	St. Dev.	95% CI
$\theta_{1,1}$	0.8995	0.0067	[0.8859,0.9121]	0.8473	0.0054	[0.8366 , 0.8578]	0.7304	0.0109	[0.7087 , 0.7515]
$\theta_{1,2}$	0.0859	0.0082	[0.0702, 0.1021]	0.1251	0.0065	[0.1126 , 0.1381]	0.2362	0.0147	[0.2070 , 0.2645]
$\theta_{1,3}$	0.0146	0.0059	[0.0041, 0.0269]	0.0276	0.0049	[0.0179 , 0.0373]	0.0333	0.0112	[0.0136 , 0.0570]
$\theta_{2,2}$	0.8671	0.0489	[0.7744, 0.9640]	0.9279	0.0288	[0.8697 , 0.9835]	0.7807	0.0345	[0.7172 , 0.8517]
$\theta_{2,3}$	0.1329	0.0489	[0.0360 , 0.2256]	0.0721	0.0288	[0.0165 , 0.1303]	0.2193	0.0345	[0.1483 , 0.2828]
ξ	0.9003	0.0419	[0.8115 , 0.9756]	0.8387	0.0272	[0.7845 , 0.8913]	0.8616	0.0536	[0.7531 , 0.9619]
α_1	258.00	21.10	[219.40 , 301.40]	300.10	21.05	[260.70 , 343.50]	422.20	74.26	[290.30 , 584.90]
α_2	39.22	5.90	[28.44 , 51.69]	61.57	6.39	[49.75 , 74.57]	84.07	14.25	[58.51 , 114.60]
α_3	5.27	1.16	[3.21 , 7.74]	6.08	1.03	[4.24 , 8.24]	17.59	4.39	[10.00 , 26.88]
β_1	11.55	0.93	[9.85 , 13.45]	13.79	0.95	[12.01 , 15.75]	23.51	4.10	[16.17 , 32.52]
β_2	4.24	0.61	[3.14 , 5.52]	6.28	0.63	[5.13 , 7.56]	8.75	1.44	[6.15 , 11.83]
β_3	1.11	0.21	[0.75 , 1.55]	1.35	0.20	[0.99 , 1.77]	4.30	1.03	[2.52 , 6.48]

Tableau III.2 – Posterior samples summary of the statistical analysis of 3800 accuracy curves, stratified by groups of territorial units. The $\theta_{i,j}$'s are the transition probabilities, the parameter ξ accounts for the level of consumption (ordinary or high) and (α_i, β_i), with $i = 1,2,3$ are the the parameters of the Beta distribution of the overall accuracies for each of the three groups of aggressiveness.

Chapter IV

Uncertainties in numerical simulation

On 22 August 1946, Stalin listened to the weather forecast and was infuriated to hear that it was completely wrong. He therefore ordered Voroshilov to investigate the weather forecasters to discover if there was 'sabotage' among the weathermen. It was an absurd job that reflected Stalin's disdain for the First Marshal who reported the next day that it was unjust to blame the weather forecasters for the mistakes.
Simon Sebag Montefiore, Stalin: The Court of the Red Tsar (2003)

Reading notes

Technical context. This chapter is quite different from the others in its form and content. Actually, the topics and the activities this chapter is concerned with correspond to a more mature phase of my professional career (since 2008-2009) at the Industrial Risks Management Dept. of EDF R&D, during which I had the opportunity to give not only a technical/scientific contribution, but also a contribution in terms of project and research management (cf. next paragraph "Contributions").

The subject of this chapter is the "generic" quantitative assessment of uncertainties tainting engineering studies based on computer simulation. Here the term "generic" is particularly important in an industrial R&D framework as methods and tools under investigation are intended to be reused in many different domains of application, based (or not) on different physics.

Motivations and issues for these works are largely discussed in Sections 1 to 3. They are rooted:

- in a larger activity carried by EDF R&D concerned with computer simulation (cf. in particular Section 2),
- in a particularly rich collaborative framework, gathering (in different forms) industrial and academic partners (cf. also Section 7.1),
- in crucial issues for an energy provider as EDF, as computer simulation is one of the ingredients of safety demonstration studies.

Hence, the main technical context of this work is industrial risk assessment and particularly safety studies. However, many other application fields as measurement science, reliability, asset management, energy trade, "smart cities" simulation can be considered as consolidated or future targets.

Contributions. The main contribution highlighted in this chapter is the technical management of the "Uncertainty Analysis" activity at EDF R&D. Actually, the EDF R&D works concerned with this topic are organized around an unifying

project (named "*Incertitudes*") I have been manager of from 2009 to 2014. Basically, the project deals with three groups of activities:

- the development and the distribution of the OpenTURNS software (cf. Section 6),
- the proper diffusion towards different disciplinary R&D and engineering Departments of methods and tools: OpenTURNS (of course) but also methodological reports accompanied by specific training sessions,
- a deep activity of scientific watch and development, generally carried within the frameworks of multipartners projects - as OPUS (cf. Section 7.2), DICE (2006-2008), CSDL (2009-2012) and ReDICE (2011-2015) - or bilateral academic partnerships associated to PhD (terminated [Blatman 2009, Fu 2012] or on-going [Damblin *et al.* 2013b, Butucea *et al.* 2013]) or post-doctoral programs (cf. [Limbourg & de Rocquigny 2010], [Limbourg *et al.* 2010], [Keller *et al.* 2011c], [Pasanisi *et al.* 2012c], [Ancelet *et al.* 2012a, Ancelet *et al.* 2012b], [Lamboni *et al.* 2013] and [Le Gratiet *et al.* 2014]) as well as methodological works carried at the own initiative of researchers of the project team, possibly in cooperation with colleagues of other industries or research institutions (for instance, one can refer to [Lebrun & Dutfoy 2009a, Dutfoy *et al.* 2012], [Blatman & Sudret 2011, Sudret *et al.* 2011], [Bousquet 2012], [Marrel *et al.* 2012] or [Faivre *et al.* 2013]).
  For having an idea of the scientific production of the project, on average 4 or 5 peer-reviewed articles or book-chapters per year are published by a project team made of approximately 4 full-time equivalent researchers.

I also had the opportunity to coordinate several collaborative working frameworks and namely the ANR OPUS project, as well as the working groups "*Fiabilité et Incertitudes*" (Reliability and Uncertainties) within the French Statistical Society (SFdS) and "*Incertitudes et Industrie*" (Uncertainties and Industry) within the French Institute for Risk Management (IMdR). Cf. Curriculum vitae, page 163, for further details.

As far as my own communications and publications related with these topics and activities are concerned, they can be classified into several groups:

- Diffusion and general presentation of methods and tools: a book chapter [Pasanisi & Dutfoy 2012], two invited conferences [Pasanisi 2011, Pasanisi 2012a], some talks in national congresses [Pasanisi 2010, Ardillon *et al.* 2012, Caruso & Pasanisi 2013] as well as the update of an internal EDF R&D methodological guide [Pasanisi *et al.* 2013a] and a technical report giving an overview of uncertainty analysis settings, with some specific considerations concerning design problems [Iooss *et al.* 2010].
- OpenTURNS: a general talk [Dutfoy *et al.* 2009], as well as an invited conference [Gaudier *et al.* 2011] and two internal reports concerning more largely software tools for uncertainty analysis [Iooss *et al.* 2011, Baudin *et al.* 2013].
- The coordination (as guest editor) of two special issues of peer-reviewed journals: [Prieur *et al.* 2011], specifically focused on stochastic methods for sensitivity analysis, and [Antoniadis & Pasanisi Eds. 2012], partly inspired from the OPUS workshops.
- Communications and internal reports on applications of uncertainty analysis in simulation: forecasting the behaviour of steam generators in specific testing conditions [Pasanisi 2008b], estimation of low-probability quantiles by means of kriging metamodels [Arnaud *et al.* 2010], inverse estimation of Strickler's roughness parameters of a shallow-water hydraulic model [Couplet *et al.* 2010], sensitivity analysis of models forecasting fouling phenomena in cooling loops of nuclear power plants equipped with cooling towers [Baudin & Pasanisi 2012, Rapenne *et al.* 2013].

Important remark: the technical and methodological contributions concerning the problems (i) of propagation of hybrid possibilistic-probabilistic uncertainties through a computer code and (ii) of point estimation of quantities of interest in uncertainy analysis study (in the sense of the framework described in Section 3), and namely to probabilistic safety criteria, will be widely sketched in dedicated chapters (Chapters V and VI).

Structure of the chapter and credits. Unlike the other chapters, most of the text shown hereby, even if it is definitely the summary of already published works, is not directly excerpted from previous documents.

The introductory considerations and motivations (Section 1 about the need for accounting for uncertainties in advanced computer simulation, as well as part of Section 5 and 6) are inspired from the book chapter [Pasanisi & Dutfoy 2012]. The brief presentation of the OPUS project (Section 7.2) is inspired (and partially excerpted) from the final report of the project [OPUS 2011].

1 Computer simulation: opportunities and issues

Computer simulation is undoubtedly a fundamental topic in modern engineering. Whatever the purpose of the study, computer models help the analysts to forecast the behaviour of the system under investigation in conditions which cannot be reproduced in physical experiments (e.g. accidental scenarios) or when physical experiments are theoretically possible but at a very high cost.

The increasing need for simulating and forecasting gave indeed a dramatic momentum in the last decades to the growth of computers' power and vice-versa. Since the very first large scale numerical experiments carried out in the 40's, the development of computers (and computer science) has gone pairwise with the will of simulating more and more deeply, more and more precisely, physical, industrial, biological, economic systems. A deep change in science and engineering has gone on in the last decades in which the role of the computer has been compared to the one of the steam engine in the first industrial revolution [Schweber & Wachter 2000]. Together with formulating theories and carrying physical experiments, computer simulation has become a "third way to Science" [Heymann 2010] that allows to solve problems which were absolutely unaffordable in a not so far past.

That raises some epistemic issues. The following quotation from [Sundberg 2010], highlighting the difference between "calculation" and "simulation", is particularly interesting and stimulating: *"The culture of calculation is modern and characterized by linearity, logic and depth, and there is a promise to explain, unpack, reduce and clarify its outcomes. Postmodern culture of simulation is fluid, decentered, and opaque and search for mechanisms and depth is futile [Turkle 1995]. The boundary between the virtual and the real is eroded, both in everyday life and in scientific fields."*

A quite negative vision of computer models, seen as sorts of magic boxes one can play with to obtain whatever desired result, arose, so that the credibility of the models themselves as tools for guiding decisions can be put under discussion: *"Most simulation models will be complex, with many parameters, state-variables and non linear relations. Under the best circumstances, such models have many degrees of freedom and, with judicious fiddling, can be made to produce virtually any desired behaviour, often with both plausible structure and parameter values."* ([Hornberger & Spear 1981] quoted by [Saltelli 2002]).

Following this reasoning, [Pilkney & Pilkney-Jarvis 2007] raised the issue of the honesty itself of forecasters, and policy makers which are often the final users of forecasts: *"The reliance on mathematical models has done tangible damage to our society in many ways. Bureaucrats who don't understand the limitations of modeled predictions often use them. [...] Models act as convenient fig leaves for politicians, allowing them to put off needed action on controversial*

issues. [...] Agencies that depend on project approvals for their very survival can and frequently do find ways to adjust models to come up with correct answers that will ensure project funding."

Adopting a more pragmatic viewpoint, we firmly believe in computer simulation as a major tool in daily scientists' and engineers' work; simulation is a great tool for understanding, for forecasting, for guiding decision. We think that the possibility to simulate more and more complex phenomena, taking into account the effect of more and more input parameters, must be rather seen as a chance than a threat. The "success stories" evoked below witness, among many others, the fundamental place (probably unsuspected for non-specialists) of advanced simulation in different domains of fundamental science and engineering.

A particularly stimulating recommendation for consolidating the credibility of computer models for risk assessment and regulation purposes is provided by [Loizou *et al.* 2008] (in the particular case of physiologically based pharmacokinetic models - cf. page 63 - but this general idea easily applies to many other domains); here it is highlighted the importance of "*increasing the understanding of regulators and risk assessors through increased transparency and accessibility to user-friendly modelling techniques*." In our opinion, the use of open source software (cf. Section 2) is a significant step towards transparency, but understandability is definitely a challenging issue as more and more refined models turn inevitably also to be more and more complex.

Even if we are convinced of the key role that computer simulation plays and will play in the years to come, at the same time, we are aware of the fact that quantitative uncertainty assessment of results is a fundamental issue for assuring the credibility of computer model based studies ... and a challenge too. Besides technical and theoretical difficulties, maybe the most challenging point is, in industrial practice, to bridge the cultural gap between a traditional engineering deterministic viewpoint and the probabilistic and statistical approaches which consider the result of a model as an "uncertain" variable.

Computers vs. test-tubes? As one of the major recent stricking facts confirming the role that computer simulation plays nowadays in science, one can think at the awarding of the Nobel prize 2013 in chemistry to Martin Karplus, Michael Levitt and Arieh Warshel "for the development of multiscale models for complex chemical systems". What is remarkable in this award, apart from the quality of the works of the three scientist, is the nature of their work: establishing computer models for deeply simulating (at the subatomic level) chemical reactions. The following quotation, from the *Information for the public* [Fernholm 2013] , motivating the award, is highly instructive: *"Using this kind of software you can calculate various plausible reaction pathways. [...] In this manner you can get an idea of what role specific atoms play at different stages of the chemical reaction. And when you have a plausible reaction path it is easier to carry out real experiments that can confirm whether the computer is right or not. These experiments, in turn, can yield new clues that lead to even better simulations; theory and practice cross-fertilize each other. As a consequence, chemists now spend as much time in front of their computers as they do among test-tubes."*

Virtual airplanes. "Virtual testing" is a popular term in manufacturing industry (and particularly in the aerospace domain) denoting the simulation of a complex system (e.g. the structure of an aircraft) for design and certification purposes. The idea is to use, as much as possible, numerical simulation instead of real tests which are expensive and time-consuming. Some basic concepts concerning virtual testing in aerospace industry are given by [Ostergaard *et al.* 2011]. In particular this practice relies on three pillars: the modelling and analyses processes, the software used and (most important) engineers skills and experience. In order to let the analysis be robust, uncertainty quantification and sensitivity analysis are considered indispensable.

Thanks to advanced simulation and CAD (Computer-Aided Design) codes, extremely detailed 3D representation of

aircrafts (named *Digital Mock-Ups* or DMU) are used by aerospace engineers in the whole project life cycle since the 90's [Sabbagh 1996, Garbade & Dolezal 2007]. A complete overview of the main industrial aspects of DMU, from the technical (requirements, data exchange, data quality), management (team organization, early warning, risk management) and communication (visualization, documentation) viewpoints is given by [Dolezal 2008].

Nevertheless, real experiments are still necessary, as the final product is a real object, so that virtual and real testing must be integrated in a coherent design framework currently known as "virtual hybrid testing" [Garcia 2013]. After all, a *flying* aircraft is real, not virtual!

Virtual nuclear reactors. The CASL project (Consortium for Advanced Simulation of Light Water Reactors), launched in 2010, is among the most ambitious R&D projects in nuclear industry. With an annual allocation of 25 M$, funded by the US Department of Energy, it gathers several key partners of nuclear R&D, among which Oak Ridge National Laboratory (leader), Idaho National Laboratory, Los Alamos National Laboratory and SANDIA National Laboratories [Michal 2011]. The main mission of the initiative is to develop and put at the disposal of the nuclear industry a multi-physics simulation environment for the simulation of a whole nuclear reactor, named VERA for *Virtual Environment for Reactor Applications*. Among the final goals, it is particularly interesting in the context of this section to highlight the following three ones [CASL 2011]:

- *"promote an enhanced scientific basis and understanding by replacing empirically based design and analysis tools with predictive capabilities,*
- *incorporate uncertainty quantification as a basis for developing priorities and supporting application of the platform tools for predictive simulation,*
- *engage the nuclear regulator to obtain guidance and direction on the use and deployment of VERA to support licensing applications."*

Interpreting rather freely and personally the three points listed above, the first one goes in the sense that up-to-date advanced simulation tools clearly enhance engineering practice. The two others are more intended to face criticisms and limitations of the use of a "virtual" software platform as a predictive (and consequently decision-aid) tool. Just like an aircraft (cf. previous example), nuclear reactors are not virtual and the use of simulation should be supervised and validated. And the uncertainty quantification could help to support the application of the simulation tools, outside R&D frameworks and purposes.

Modelling human body. Like all complex systems, the behaviour of human body and human organs can be implemented in more or less complex computer models to be used for a number of different purposes, and in particular for risk assessments (e.g. effect of chemicals or electromagnetic fields on human health).

Models can be directly based on the finite-elements resolution of physically based partial derivative equations, e.g. Navier-Stokes (simulation of vesicles [Doyeux *et al.* 2012] or blood flow in arteries [Prud'Homme *et al.* 2011]) or Maxwell equations (exposure of organs or foetuses to electromagnetic fields [Wong & Wiart 2005, Jala *et al.* 2013]).

A very important class of models is rather phisiologically based (PB). They are widely used to forecast the complex process of absorption, distribution, metabolism and excretion (ADME) of chemical substances (typically drugs or toxics). Depending on the nature of the substance, these models are usually referred as "physiologically based pharmacokinetic" (PBPK) or "toxicokinetic models" (PBTK). The body is subdivided in a number of compartments through which sub-

stances move and are transformed (e.g. well-perfused tissues, poorly perfused tissues and fat), connected by blood and lymphatic circulation.

For both types of models the importance of accounting for uncertainties due to lack of knowledge and/or variability among human beings is considered paramount. In particular, an established practice of Bayesian calibration (e.g. [Bois 2001, Micallef *et al.* 2005]) exist for PBPK and PBTK models. Advanced methods of uncertainty propagation and sensitivity analysis are also commonly used for both kinds of models [Brochot *et al.* 2007, Loizou *et al.* 2008, Jala 2013].

2 Advanced numerical simulation at EDF R&D

Among the French industrial companies, EDF has one of the largest R&D Units, with a permanent staff of about 2000 engineers and researchers and 150 PhD students, organized in 15 thematic Departments. One of the distinctive features of the EDF R&D activity is the great number of areas of interest it copes with. That is due to the great variety of EDF's activity: energy production, transmission, distribution and sale, as well as to the great complexity of the nuclear production process, involving a number of different physics: neutron transport, solid and fluid mechanics, thermo-hydraulics, shallow-water hydraulics, electromagnetism.

Consequently, EDF R&D makes an intensive use of computer modelling and simulation. More than being *simple* users, research teams develop most of the codes used in the applied studies and put them at the disposal of the engineering and business Units of the EDF Group and (as far as most of the codes are concerned) of the technical and scientific community. More than being just working tools, numerical codes play a paramount role in the organization and the structuring of R&D activities.

The motivations of this strong effort are of different natures [Andrieux 2011]:

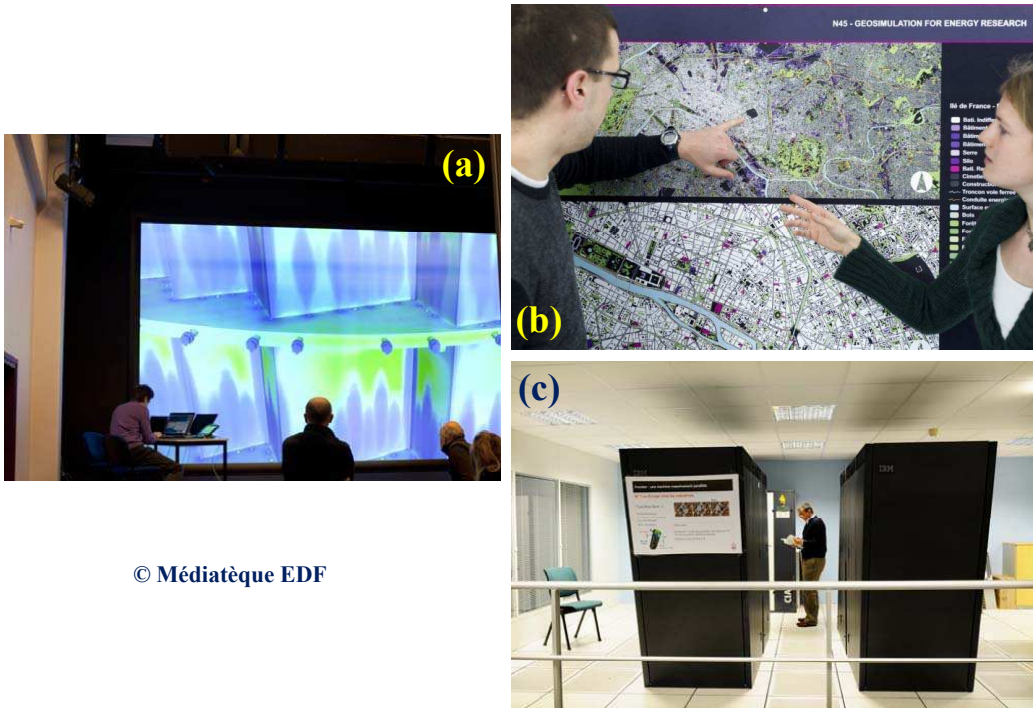
- First, the necessity to realize very complex simulations, anticipating and eventually prototyping industrial studies, which demands proper and specific models, often not yet developed nor implemented in software "available on the shelf".

- Second, the need for capitalizing R&D modelling efforts and making them available for future engineering studies.

This double goal can be reached by implementing and referencing specific software libraries and platforms which naturally become repositories of knowledge and skills. As a consequence, software also becomes a structuring tool for researchers and engineers working in the same disciplinary field; the thematic community meets up around different software-related activities: code implementation, documentation, diffusion, presentation of examples and studies (for instance) during informal and formal meetings (seminars, project reviews, users' days).

Most of the software developed by EDF R&D is open source[1]; the motivations for this strong open source positioning are numerous. First, this allows the possibility of external contributions of different nature (development of new features and algorithms, case studies and examples, bugs reporting) and thus sharing the R&D effort with the technical and scientific community. Second, this facilitates the cooperation with industrial and academic partners in collaborative frameworks, e.g. funded projects, PhD or post-doctoral programs. Third, it facilitates the dissemination (and consequently the acceptability) of methods and tools. In few words, the open source positioning gives rise to a technical environment naturally more attractive, collaborative, international.

[1] Cf. http://chercheurs.edf.com/logiciels/tous-les-logiciels-41436.html for an overview of the main open source codes of EDF R&D and the links to the download web-pages.



© Médiatèque EDF

Figure IV.1 – Some examples witnessing the R&D efforts in computer simulation carried by a great industrial company as EDF in the last years: (a) 23.9 million pixels *Visualization Wall* in Clamart for large scale visualization of complex computer experiments, (b) geo-planning simulation in Karlsruhe, (c) one of the IBM BlueGene® supercomputers of EDF R&D.

The open source codes of EDF R&D cover a large variety of physics: solid mechanics and structural analysis (*Code_Aster*), computational fluid dynamics (*Code_Saturne*), conduction and radiation heat transfers (*Syrthès*), acoustics (*Code_TYMPAN*), shallow-water hydraulics (*Mascaret* and *Telemac*).

Together with these physical calculation codes, EDF R&D also develops (in cooperation with CEA and Open Cascade Company) the *SALOME* platform [Ribes & Caremoli 2007, Chauliac *et al.* 2011], a fully integrated open source environment for numerical simulation, integrating a number of generic tools for pre- and post-processing (importing, modifying and exporting CAD models, meshing , parallel visualisation, models couplings, supervision of distributed computations) as well as specific tools for data assimilation and uncertainty analysis (in particular OpenTURNS).

In the domain of uncertainty and sensitivity analysis, the home-made software is OpenTURNS, key product of the projects *Incertitudes* since 2005. Its main features are sketched in Section 6.

It is worth noting to conclude this short introduction to advanced simulation at EDF R&D that the effort on numerical codes development is completed by the putting at the disposal of researchers and engineers powerful high performance computing (HPC) facilities. Thanks to a long-standing cooperation with IBM [Vezolle & Berthou 2008], three different families of the BlueGene® supercomputers, BG-L (25 Tflops, 2007), BG-P (110 Tflops, 2008), BG-Q (800 Tflops, 2012), have been successively deployed, as well as the x86 technology machines Ivanohe (200 Tflops in 2010) and Athos (400 Tflops, 2013).

In November 2013 (last data available when writing this manuscript), three EDF supercomputers appear in the list of 500 world most powerful machines ranked by their performance on the a standardized benchmark (`http://www.top500.org/`) at the rank number 46 (*Zumbrota* machine), 91 (*Athos*) and 257 (*Ivanohe*), respectively.

3 Different kinds of problems

Even though the fundamentals of these topics are rooted since decades in probabilistic and statistic literature, in the last years there has been a considerable rise of interest in industries and academia in the uncertainty quantification (UQ) of computer models' results.

At least in theory, one can think that the (statistical) uncertainty analysis of outputs of predictive computer models is essentially done for supporting decisions. For instance, a shallow-water hydraulic model which returns predictions of the water levels in different sections of a river, for a given discharge, can be used for evaluating the height of the protection embankment as the solution of a decision problem: the optimal height is the one that minimizes the expected loss, sum of the cost of the dike and the (possible) damages in case of overflowing. In practice, this vision is quite theoretical: in general, analysts coping with computer simulation are quite far from the actual decision problem and/or this problem may be not enough well-posed to be solved in a strict mathematical decision setting; one can think at climate modelling or the modelling of very rare events of human and/or natural origin with catastrophic consequences. In both cases, it is hardly possible to encapsulate the stakes motivating the study into a (even complicated) loss function.

In many cases, the study to be realised by means of computer simulation is clearly not related to a decision problem (at least in its initial stages). For instance:

- In modern measurement science, computer codes for indirectly measuring quantities which cannot be directly measured are more and more used [Désenfant *et al.* 2007]; the issue is to find out the uncertainty tainting the mesurand (output) value by propagating the measurement uncertainties of the inputs. Since 2008, the reference handbook within the measurement science community [JCGM 2008a], the "Guide to the expression of uncertainty in measurement" (also known as GUM, first edited in 1993), has a specific Supplement [JCGM 2008b] concerning the use of Monte Carlo methods to propagate uncertainties in measurement models, presented as an alternative to the standard approach in metrology (the so called "law of propagation of uncertainty", actually based on the first-order Taylor decomposition of the measurement model). The publication of this reference document witnesses the place that more and more complex (namely highly non-linear and of non-trivial differentiation) computer codes take in metrology today. Definitely, metrologists are today more and more concerned with intensive numerical simulation [Cox *et al.* 2012].

- In safety studies, one has often to answer to specific questions about the probability distribution function of the output variable of a generally highly complex code, e.g. what is the probability for the output to be greater than a fixed threshold? Actually, probabilistic studies are more and more used for regulatory purposes [Cunningham 2012], often as a complement of classical penalized deterministic safety assessments. It is worth noting that in some cases, the initial questions asked by the regulation authority are not fully formalized by a mathematical viewpoint and it is up to the analyst to propose the proper mathematical framework for coping with the issue. See, for instance the example shown by [Helton & Sallaberry 2012] concerning the safety requirement for the Yucca Mountain nuclear waste repository (Nevada, USA). One of the requirements of the regulation authority (Nuclear Regulatory Commission, NRC) was formally expressed as follows:
  "Department of Energy must demonstrate, using performance assessment, that there is a reasonable expectation that the reasonably maximally exposed individual receives no more than the following annual dose from releases from the undisturbed Yucca Mountain disposal system: (1) 0.15 mSv (15 mrem) for 10,000 years following disposal; and (2) 3.5 mSv (350 mrem) after 10,000 years, but within the period of geologic stability."
  Here, an important part of the analysts' work has been to propose and justify a mathematical formulation of terms

like *"reasonable expectation"* or *"reasonably maximally exposed individual"* following the requirement formulated above. It is interesting to see that methods for quantitative uncertainty assessment may sometimes provide not only the answer but also (part of) the question. This approach (compliance with a more or less initially formalized risk criterion) could be qualified as "normative", even if the term "norm" is to be interpreted in a broader sense.
A quite similar situation arises in financial studies, in which the analyst is asked to provide a very specific risk criterion, as the popular "*Value at Risk*" (VaR), i.e., roughly speaking, a quantile of the probability distribution of gains (or losses) of a given portfolio, over a given time period (cf. for instance [Linsmeier & Pearson 2000]). Even if the relevance of this "simple" criterion can be challenged by other risk summaries for theoretical and practical reasons in real-world complex situations [Rockafeller & Uryasev 2002], VaR has rapidly become very common in the financial community and is largely used for both internal studies and regulatory purposes [Lopez 1996].

- In some other cases, one just wants to "explore" the code for better understanding its behaviour with respect to the variability of the inputs. The purpose of the study is here mainly a sensitivity analysis of the computer model (and of the underlying phenomenon). Actually, when the complexity of the code increases, even though it is reasonable to expect that the model provides a more and more accurate representation of the reality, the analyst more and more needs effective mathematical tools to identify and summarize results. Which are the actually influential input parameters? Which are the sets of the inputs that transform into sets of high (and/or low) values of the output? Hence, sensitivity analysis is a precious tool to let the analyst be more aware of his/her model, to eventually suggest modifications or simplifications and to identify the input variables on which further R&D efforts must be put to improve the quality of the results (cf., for instance, the review article of [Iooss 2011] and the numerous references therein, as well as [Saltelli *et al.* 2004] and [Faivre *et al.* 2013]).

An interest taxonomy of the main initial questions and expected goals of uncertainty analyses, strongly dependent on the context of the study, can be found in the collective work edited by [de Rocquigny *et al.* 2008] summarizing the activity of a thematic working group within the scientific society ESReDA (European Safety, Reliability and Data Association). Here, four different goals are distinguished:

- **Understand.** Better understand the behaviour of the model and rank the input variables with respect to their contribution to the output "*variable of interest*" (this term will be clarified hereinafter) in order to prioritize further engineering or R&D efforts.

- **Accredit.** Give credit to a numerical predictive code or to a measurement model by: properly assessing uncertainties of input measures or input variables, simplifying the model, fixing some values of the inputs and finally validating the model with respect to the expected outcomes and the stakes of the problem (not-exhaustive list of actions to be made depending on the context).

- **Select.** Compare performances of systems and strategies and choose among alternative options in both early (e.g. design, deployment) and mature (e.g. operating and/or maintenance policy) stages of the lifetime.

- **Comply.** Demonstrate the system complies with a more or less explicit regulation criterion, typically formulating as an inequality (e.g. the annual CO_2 emissions are below a fixed threshold).

Of course, in real problems, the distinction between the goals listed above may be fuzzy: analysts can strive towards different nested goals in the same study. For instance, in a safety study aiming at verifying the compliance with a regulatory fixed threshold, one first has to show the relevance of the used codes and methodologies. In any case, as a

final recommendation, the contributors to the ESReDA working group particularly highlighted the importance of the prior identification of the goal of the study, before planning any calculation efforts. Actually that is a fundamental step for choosing the mathematical methods to be put into practice for solving the problem.

From a more methodological viewpoint, it can be interesting, for several reasons, to define a common and simple step-by-step framework. It focuses on so-called parametric uncertainties, i.e. the ones affecting the input parameters of a model, whatever it is: a complex numerical code which requires an approximated resolution or an analytical expression. It does not question explicitly uncertainties attached to the computer model itself, coming from the necessarily simplified modelling of the physical phenomenon under investigation, nor numerical uncertainties due to its practical implementation into a computer code. The methodology is based on the probabilistic paradigm, i.e. uncertainties are modelled by means of probability distribution functions (pdf).

This common framework of uncertainty management is conveniently seen as a four-steps process (Figure IV.2): (i) *Step A "Problem's Specification"* defines the structure of the study by selecting the random parameters, the outcomes of interest and the features of the output's pdf which are relevant for the analysis; (ii) *Step B "Input Uncertainty Quantification"* defines the probabilistic modelling of the random inputs; (iii) *Step C "Uncertainty Propagation"* evaluates the criteria defined at Step A; (iv) *Step C' "Uncertainty Importance Ranking"* determines which uncertainty sources have the greatest impact on the outcome (sensitivity analysis).

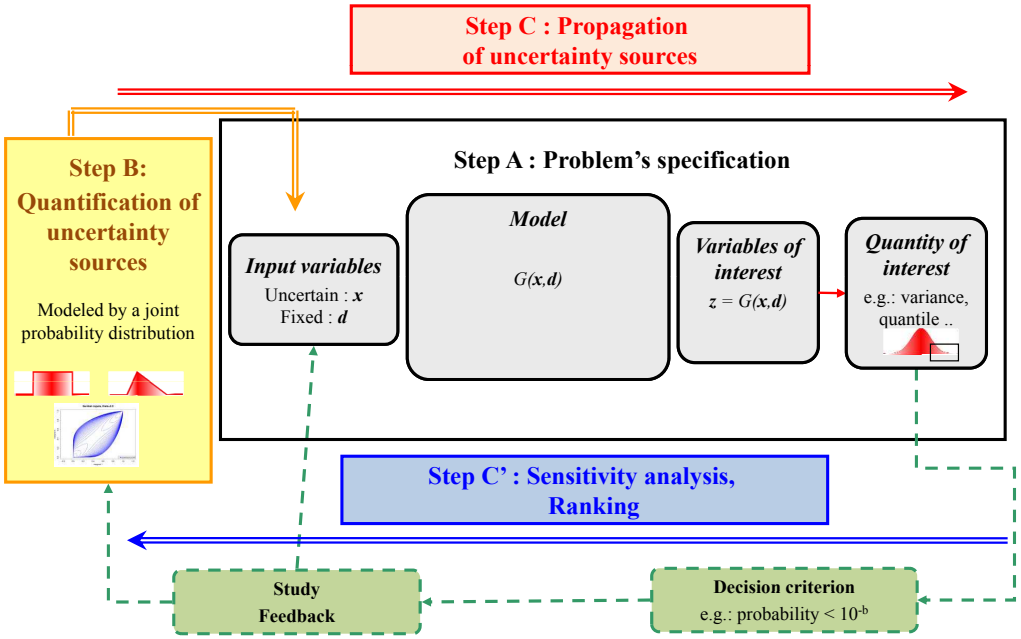


Figure IV.2 – The common methodological four-steps framework for uncertainty analysis, used in common practice.

The code $G(\boldsymbol{x}, \boldsymbol{d})$ relying the output variable of interest, noted $\boldsymbol{Z}$ to uncertain ($\boldsymbol{X}$) and fixed inputs ($\boldsymbol{d}$) is at the heart of the study. In the most general case $\boldsymbol{Z}$ is a vector. However, in the remainder we will consider the case of a scalar variable of interest, noted Z, for the sake of simplicity and also because, most of the times, it is the case in actual industrial studies.

It is worth noting that the separation between uncertain and fixed inputs proves useful in practice. The vector $\boldsymbol{d}$ (depending on the particular problem to be solved) can include "*certain*" variables (which are known with absolute certainty) and/or "*not-significant*" variables (which do not contribute significantly to the output uncertainty) and/or "*scenario vari-*

ables", the values of which are fixed a priori to describe a scenario (e.g. penalizing values of temperatures and pressures in a safety study) conditional to which the probabilistic calculation is made.

Of course this scheme is somehow reductive with respect to the complexity of most of the underlying problems to be solved. Nevertheless, it proves adapted to industrial practice for different reasons. In particular, the separation of the steps is clear for the engineers and useful (in project management) to identify different work-stages demanding different skills; of course, this kind of study require a multidisciplinary team but the cooperation of engineers specialist of the domain of application (e.g. a specific physics) during step A is extremely important, while the other steps, more technical, demand rather mathematics and software engineering skills.

Another important feature of this scheme is its genericity. Actually, it is not intended to specific problems and serves as a common base for treating a number of very different questions. This scheme is actually inspired by the organization of the uncertainty analysis activities in many industrial companies that, because of the complexity and the variety of their business core (let us think for instance of EDF, CEA or Airbus-Group), deal with problems involving several initial questions and stakes and different physics. This generic approach is consistent with the work organization of these companies which have set in the last year specific "uncertainty analysis" project-teams. Actually this organization proved to be more effective than the one consisting in having inside each disciplinary team one or two experts in statistics and probabilities for coping with uncertainty quantification in a given well-specified technical framework. For this same reason of genericity, this scheme is "non-intrusive", i.e. it considers the numerical code $G(\boldsymbol{x},\boldsymbol{d})$ as a "black box" transferring uncertainties from the inputs to the outputs.

This non-intrusive setting inspired the works of a generation of researchers [Sudret 2007, Iooss 2009, Roustant 2011, Morio 2013] at the interface between industrial R&D and academic research. The software OpenTURNS (cf. Section 6) is also fully consistent with this methodology.

As also sketched above, the methodology is also rooted into the practice of metrologists (uncertainty propagation of measurement errors). Here, Z is the final measurand and $\boldsymbol{X}$ the directly measured variables, related to Z by the measurement function $G(\cdot)$. The problem is generally posed as the estimation of the variance or the standard deviation of Z, given the (joint) pdf $p(\boldsymbol{x})$, by means of or Monte Carlo simulation or using the so-called "law of propagation of uncertainty", as advocated by the GUM [JCGM 2008a]:

$$\sigma_z^2 = \sum_{i=1}^{m} \left(\frac{\partial G}{\partial x_i}\right)^2 \sigma_i^2 + 2 \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} \frac{\partial G}{\partial x_i} \frac{\partial G}{\partial x_j} \sigma_i \sigma_j \rho_{ij}, \tag{IV.1}$$

in which the partial derivatives are evaluated for $\boldsymbol{x} = \mathbb{E}(\boldsymbol{X})$, σ_i is the standard deviation of X_i and ρ_{ij} the correlation coefficient of the random pair (X_i, X_j).

The methodology sketched in Figure IV.2 also owes a lot to a well-known technical and research framework: the *Structural Reliability* (see the works of [Ditlevsen & Madsen 1996] and [Lemaire *et al.* 2010] or [Sudret 2007] for an overview of stakes and mathematical methods). Developed within the probabilistic mechanics community since the 70's, this discipline focuses on the reliability of mechanical and civil structures and in particular on the probability for a given structure to attend a *limit state*, beyond which it does not fulfil anymore its safety or capability requirements. One can think for instance to the elementary case of the "$R-S$ state limit function": the structure fails when the the mechanical solicitation S is greater than the resistance R. The probability of failure is then the probability for the state limit function, $G(x,d)$, to be ≤ 0:

$$P_f = \int_{\mathcal{D}_f} p(\boldsymbol{x}) d\boldsymbol{x}, \quad \text{with: } \mathcal{D}_f := \{\boldsymbol{x}; G(\boldsymbol{x},\boldsymbol{d}) \leq 0\}, \tag{IV.2}$$

in which $p(\boldsymbol{x})$ is the joint pdf of the uncertain input variables and $\mathcal{D}_f$ is the so-called *domain of failure*. Generally, the probability P_f is small (orders of magnitude from 10^{-2} down to 10^{-7}) and a single evaluation of the function $G(\cdot)$ is computationally expensive. As a consequence, very specific mathematical and software tools shall be used.

It is worth noting that the scheme represented in Figure IV.2 contains both the evaluation of a measurement uncertainty and a probability of failure. From a methodological viewpoint, the main difference between the two problems consists in the choice of the quantity of interest: a standard deviation in the first case, a probability for the output to exceed a fixed threshold in the second case. According to the introduced methodological framework the goal of an uncertainty analysis is to estimate a given quantity of interest; far from solving a decision problem, the analyst is asked to solve a much more "humble" estimation problem. Notice also that in the seminal paper of [Kennedy & O'Hagan 2001], the uncertainty analysis is defined as *"the study of the distribution of the code output that is induced by probability distributions on input."* The formalization of the problem as the estimation of a quantity of interest, somehow specifies what the "study of the distribution" actually is.

4 The pioneering works of Jacques Bernier

The rapid expansion in industrial engineering and R&D of the discipline nowadays called "uncertainty analysis" or "computer experiments" started in the 90's. This is mainly due to the more and more easy access to powerful computing machines and the large diffusion of reference works on fundamental mathematical tools as response surfaces, based on polynomial chaos expansion [Ghanem & Spanos 1990] or kriging [Sacks *et al.* 1989], FORM/SORM[2] structural reliability methods [Dolinski 1983, Madsen *et al.* 1986], accelerated Monte Carlo sampling [McKay *et al.* 1979, Ditlevsen *et al.* 1988, Bucher 1988, Melchers 1990], sensitivity analysis [Cacuci 1981, Morris 1991, Sobol 1993].

In the industrial R&D community, the most known intensive uncertainty analysis studies involving complex models have been made since the late 80's by Jon Helton and his colleagues of SANDIA National Laboratory (USA). Among the very first studies published in this domain in international journals, one can find the uncertainty and sensitivity analysis of numerical codes modelling: (i) the movements of chemical releases between different zones of a boiling water nuclear reactor in accidental conditions (station blackout) [Helton & Johnson 1989], (ii) the consequences of a major nuclear accident in terms of atmospheric dispersion, dry and wet deposition, biospheric transport of radioactive materials as well as health effects (short and long term fatalities and injuries) and costs, given the weather conditions and the "*source term*" of pollution (i.e. amounts, heat content and timing of the accidental release) [Helton *et al.* 1992], (iii) the release of radionuclides to the accessible environment (i.e. atmosphere, land surfaces, water bodies etc.) for 10000 years after disposal from a number of possible events (typically drilling intrusions) that may affect the Waste Isolation Pilot Plant (Carlsbad, New Mexico, USA), a deep geological repository of transuranic waste from military applications [Helton 1996].

The works evoked hereinbefore are well known within the technical and scientific community. Here, I would rather insist on other early and pioneering works, much less known and actually extremely interesting, carried in the 70's and early 80's by Jacques Bernier, former senior researcher at EDF R&D and retired since 1991.

J. Bernier is mainly known for his numerous works on stochastic hydrology. A synthetic survey of his career is given in [Jacquet *et al.* 1998], the preface of a collective work collecting the proceedings of a conference organized in his honour in Paris in September 1998. The authors insist in particular on three key contributions of Prof. Bernier, the paternity of which was quite poorly known: the use of the Fréchet distribution (which can in some cases prove better than the Gumbel one) for modelling flood discharges since the mid of the 50's, the use of the so-called "renewal methods" in hydrology

[2] First/Second Order Reliability method: approximate methods for fast computing a failure probability (in the sense of structural reliability).

since the end of the 60's, the application of Bayesian analysis methods in stochastic hydrology since the end of the 60's for effectively coping with the two engineering concerns of incomplete or poorly informative data and decision making.

When, some years later, I decided to organise, together with É. Parent (AgroParisTech) and J.J. Boreux (University of Liège) a new conference in the honour of J. Bernier, ("*Decision statistics and engineering under uncertainties*", Paris, October 2012), I had the opportunity to discover a number of quite unknown works of him, where he explicitly copes with the problem of uncertainties tainting the output of predictive (deterministic) models. The great difference between this works and the ones (much more known) of J. Helton in the 90's cited above stands in the methodology used. SANDIA's works are much more similar to the ones carried today in the industry, as they involve complex computer models and intensive simulation-based techniques.

Nevertheless, it seems interesting to briefly remind some works not only for historical reasons but also because their are rich in teachings and a source of inspiration. Some of them, directly dealing with partial differential equations, made stochastic by adding a random error term or by making aleatory some of their coefficients, were rooted in a quite common practice at that time (cf. the works in the domain of the *statistical mechanics* cited by [Frisch 1968]). Other ones, more perspective (and quite innovative), are more concerned with sampling-based techniques.

Water quality forecast. In [Bernier & Sabaton 1972] the technical problem under investigation is the simulation of the water quality of a river, in terms of biological oxygen demand (BOD) and dissolved oxygen (DO). According to the popular Streeter and Phelps model, the phenomenon is governed by the following equations:

$$\begin{cases} \dfrac{dz}{dt} = k_1\, y - k_2\, z \\ \\ \dfrac{dy}{dt} = -k_3\, y, \end{cases} \tag{IV.3}$$

z being the oxygen deficit (i.e. the difference between the saturation and actual DO concentration), y the current BOD and k_1, k_2, k_3 model's parameters.

The authors insist on the importance to account for uncertainties when making prediction using the model above and distinguish the case of long term prediction (here the dominant source of uncertainties is the variability of parameters intervening in Equations IV.3 and more generally environmental parameters) and the case of short term prediction, in which the main errors may come from the discrepancy between the model and the physical reality. For coping with uncertainties in short term predictions the authors propose a stochastic version of Equations IV.3 by adding a random error term to each of the two original equations. These errors, for each value of the time t, follow a joint bivariate Gaussian distribution with means and standard deviations equal to (μ_1, μ_2) and (σ_1, σ_2), respectively (they can be correlated for a given value of the time t):

$$\begin{cases} dz = (k_1\, y - k_2\, z + \mu_1)\, dt + \sigma_1 \sqrt{dt}\, \varepsilon_1(t) \\ \\ dy = (-k_3\, y + \mu_2)\, dt + \sigma_2 \sqrt{dt}\, \varepsilon_2(t), \end{cases} \tag{IV.4}$$

in which $\varepsilon_1(t)$ and $\varepsilon_2(t)$ follow a standard Gaussian distribution. By definition, $\varepsilon_1(t)$ and $\varepsilon_2(t)$ can be dependent for a fixed t but both are independent on $\varepsilon_1(t+dt)$ and $\varepsilon_2(t+dt)$ respectively. According to these assumptions, the authors provided the solution of the stochastic differential Equations IV.4, i.e. the (bivariate Gaussian) probability distribution of the couple (y,z). The model is used for evaluating the probability for the maximum value of the deficit to be greater than a

given (and unacceptable) threshold value $z^\star$, under the initial conditions on the pollution source terms: $y = y_0$ and $z = z_0$. This maximum is reached, according to Equations IV.4, for $t = t_c$ (critical time):

$$t_c = \frac{1}{k_2 - k_3} \log\left(\frac{k_2}{k_3}\right)\left(1 - \frac{z_0\,(k_2 - k_3)}{k_1\,y_0}\right). \tag{IV.5}$$

This kind of study is made for design purposes, i.e. fixing $y = y_0$ and $z = z_0$ under the probabilistic constraint: $\mathbb{P}[z(t_c) \geq z^\star] \leq \alpha$.

Thanks to the assumptions above, the criterion can be written as:

$$\frac{z^\star - \mu_z(t_c)}{\sigma_z(t_c)} \geq u(\alpha), \tag{IV.6}$$

where $\mu_z(t_c)$ and $\sigma_z(t_c)$ are the mean and standard deviation of the oxygen deficit for $t = t_c$ and $u(\alpha)$ the α-quantile of the standard Gaussian distribution.

The authors also discuss the problem of the uncertainties tainting the criterion formulated in Equation IV.6 (nowadays qualified as "epistemic"). The source of these uncertainties is the imperfect knowledge of the so-called *state of Nature* θ, i.e. the set of the overall parameters of the model (here θ is made by the k_i's, the μ_j's and the σ_j's, estimated by data coming from in situ measures). A Maximum Likelihood and a Bayesian plug-in estimator (cf. Chapter VI) are proposed, i.e. the criterion is estimated by replacing θ with its Maximum Likelihood estimator or its Bayesian posterior mean.

Thermal pollution forecast. The same intrusive approach for coping with uncertainties has been used, in the same period in some studies concerning the evaluation of thermal pollution caused by the releases of power stations, located seashore [Bernier 1975]. If one notes z the sea temperature increase in the 2-dimensional (x, y) domain of interest, the phenomenon is ruled by the convection-dispersion equation:

$$\frac{\partial z}{\partial t} + v_x\frac{\partial z}{\partial x} + v_y\frac{\partial z}{\partial y} - \frac{1}{h}\left[\frac{\partial}{\partial x}\left(D_x h\frac{\partial z}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_y h\frac{\partial z}{\partial y}\right)\right] = \frac{A}{C_p h}z, \tag{IV.7}$$

in which the first terms model the convection (depending on the component v_x and v_y of the velocity vector $\boldsymbol{v}$) and the three terms in the brackets governs the dispersion (depending on the dispersion coefficients D_x and D_y and the water level h). The right-hand-side term concerns the atmospheric heat exchange (A is the exchange coefficient and C_p the water volumetric heat capacity). The limit conditions are that in the release point, the released power W is known, and that very far from the rejection point, the temperature increase tends to be null:

$$\begin{cases} C_p\,Q\,(z - z_p) = W \\ \\ z \to 0 \;\text{ for }\; x, y \to \infty. \end{cases} \tag{IV.8}$$

The problem becomes stochastic by considering the uncertainty tainting the velocity field, the dispersion coefficients and the atmospheric exchange coefficient. Taking into account the complexity of the problem of finding the probability distribution of the spatial and temporal process $z(x, y, t)$, the recommended approach is to solve the differential equations for the moments $\mathbb{E}[B^n]$ of the random variable defined by the equation $b = z \cdot \exp\left(\frac{A}{C_p h}t\right)$.

In particular, in [Lencioni *et al.* 1979], the problem is solved under the assumption that the velocity fields can be

expressed by the product of a deterministic function $H(x,y)$ and of a time-dependent Markov process $\boldsymbol{q}(t)$. This seemingly restrictive condition was considered as satisfied, under the basis of the available data, for the Channel seashore.

In this case, the introduced stochastic term has a physical interpretation, as the random convection is equivalent to an additional dispersion term, which proved to be much greater than the physical dispersion.

Methodological issues and recommendations. Methodological issues and perspectives are discussed in a more methodological paper [Bernier 1980] which also gives the impression of the end of this early period of interest for uncertainties in simulation. Together with the approaches (qualified as *probabilistic*) consisting in randomizing differential equations governing physical phenomena, the *statistical* methods, actually sampling-based, are proposed as a more and more attractive way to cope with complex models and namely for estimating the probability for a model output $z = G(\boldsymbol{x})$ to be in a given set of values, $P_{\mathcal{A}} = \mathbb{P}[Z \in \mathcal{A}]$.

A method, apparently used in the 70's for the study of the atmospheric dispersion of cooling tower plumes is described. It is based on the partition of the input space in m classes $\mathcal{C}_i$ of probability $p_i = \int_{\mathcal{C}_i} p(\boldsymbol{x})\,d\boldsymbol{x}$. An approximate, but demanding a limited number of simulations, technique consists in (i) choosing in each class one representative point $\boldsymbol{x}_\star^{(i)}$, then (ii) evaluating $G(\boldsymbol{x}_\star^{(i)})$ and finally (iii) using the following estimator for $P_{\mathcal{A}}$:

$$\hat{P}_{\mathcal{A}} = \sum_{i=1}^{m} p_i \cdot \mathbb{1}_{\{G(\boldsymbol{x}_\star^{(i)}) \in \mathcal{A}\}}, \tag{IV.9}$$

which actually presents a non-controlled bias.

From this method which "deterministically" states the class $\mathcal{C}_i$ contributes to the event $\{Z \in \mathcal{A}\}$ under the basis of a single evaluation in a chosen point of $\mathcal{C}_i$, a method based on random allocations of points of $\mathcal{C}_j$ is proposed. In practice, it consists in random sampling n_i points $(\boldsymbol{x}^{(i,1)}, ..., x^{(i,n_i)})$ within each class $\mathcal{C}_i$ and evaluating the probability $P_{\mathcal{A}}$ as:

$$\hat{P}_{\mathcal{A}} = \sum_{i=1}^{m} p_i \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbb{1}_{\{G(\boldsymbol{x}^{(i,j)}) \in \mathcal{A}\}}. \tag{IV.10}$$

The properties of this sampling method, as well as the optimal allocation of the number of random points between the m classes are discussed. The reader familiar with Monte Carlo variance reduction technique can recognize the proposed technique as a "stratified sampling", a method which will become popular in the computer simulation and structural reliability communities between late 80's and 90's (e.g. [Schuëller *et al.* 1989, Decker 1991, Ye *et al.* 1993]). Even though the method is rooted in the fundamentals of Monte Carlo simulation, it is worth noting that its use in the industrial "computer experiments" community was quite innovative in that period. As an historical point of reference, the seminal paper (we do not think J. Bernier was aware of when writing the cited report) introducing the Latin Hypercube Sampling as a variant of the stratified sampling in the context of uncertainty analysis dates from the end of the 70's [McKay *et al.* 1979].

This report ends with a discussion concerning the difficult compromise between using the best model available for representing the physics of a phenomenon and the quantitative uncertainty assessment of its results. The more complex the model is, the more expensive a single run will be and thus the more difficult the probabilistic calculations will be. The use of response surfaces, named "semi-empirical" or "statistical" models in [Bernier 1980], fitted on data generated by the original model is presented as a methodological prospective for the years to come. That will be actually one of the mainstream of the work of scientists and engineers in the following three decades! Notice that the first known article explicitly concerned with response surfaces in computer experiments dates from the mid of the 80's [Downing *et al.* 1985].

Of course, the discussion about who was the first one to discover or rather apply one method or another to a given

technical domain is questionable and actually not interesting. Indisputably, the diffusion of these works, published in French and as internal reports, has been limited and that is why they have been rather forgotten. Nevertheless, it was interesting to give them here a tribute which, to the best of my knowledge, has not been done yet.

5 The need for specific methods

In the remainder, relying on the terms of the scheme represented in Figure IV.2, some elements about the practical implementation of the different steps of uncertainty analyses are given.

Step A - Problem's specification. This step first requires to select the input parameters to be modelled as uncertain variables. The remaining parameters are considered as fixed either because they are supposed to be known with a negligible uncertainty or (as it is typical in safety study) because they are given values, generally conservative, which are characteristic of a given accidental scenario.

Step A requires also to select the relevant features of the outputs' pdf, depending on the stakes which motivated the study (the so-called *quantities of interest*). In most cases they formalize, in a quite normative and simplified way, some decision criteria. For instance, within the design stage of a system, the analyst is often required to provide the mean and the standard deviation (or the range) of a given performance indicator of the system (e.g. fuel consumption), whereas in operating stages, one has to check if the system meets or not some regulatory requirements attached to licensing or certification. Then, depending on the context of the study, the decision criteria may be: (i) a *min-max* criterion, i.e. the range of the outcomes given the variability of the inputs; (ii) a *central dispersion* criterion, i.e. central tendency and dispersion measures; (iii) a *threshold exceedance* criteria, i.e. the probability for a state variable of the system to be greater than a threshold safety value.

A short analysis of the computer code is also necessary: does it require a high CPU time for a single run, does it provide a precise evaluation of its gradient with respect to the probabilistic input parameters are primordial questions.

Depending on the previous specifications, the methodology will be implemented through different algorithms.

Step B - Input's uncertainty quantification. The methods used for the probabilistic modelling of the inputs depend on the nature and the amount of available information.

In case of scarce information, the analyst first needs to interview experts. The literature proposes numerous protocols (e.g. [O'Hagan *et al.* 2006]) that can help to get (hopefully) unbiased and relevant informations which are then translated into a pdf. In addition, a very simple and commonly used approach [JCGM 2008a] consists in applying the Maximum Entropy Principle, that leads to the pdf maximizing the lack of information (encoded by the Shannon's entropy [Shannon 1948]), given the available expertise on the variable to be modelled. Whatever the chosen model, it is important to come back to the experts and "validate" it by establishing a dialogue on some clearly understandable key features of the established pdf (e.g. mean and quantiles, rather than shape or scale parameters).

When data set are available, the analyst can use the classical statistical inference tools following a parametric or non parametric approach.

We insist on the fact that the random input parameters $X_1, ..., X_m$ form a random vector $\boldsymbol{X}$ with a multivariate pdf, the dependence structure of which must be taken into account. A common way is to define the multivariate pdf $p(\boldsymbol{X})$ by means of its univariate marginal distributions $p_1(x_1), \ldots, p_m(x_m)$ and its copula $\mathbf{C}(\cdot)$, encoding the dependence structure [Genest & Favre 2007]. In practice, the inference on copula's parameters could be tricky and Kendall's τ or Spearman's

ρ coefficients are not sufficient to completely determine the dependence structure, as shown in [Dutfoy & Lebrun 2009]. Mismodelling the dependence structure is potentially dangerous as it can lead to an error of several orders of magnitude in the estimate of a threshold exceedance probability [Lebrun & Dutfoy 2009b]. Actually, the copula inference should be performed using the same techniques (e.g. Maximum Likelihood Estimation) as those for the univariate marginals.

Step C - The Uncertainty Propagation Once quantified, uncertainties are propagated to the model outcomes. The choice of the appropriate propagation algorithms depends on the quantity of interest to be estimated and on the model characteristics specified in Step A.

In case of a *min-max* analysis, the range of the outcome is determined either thanks to an optimization algorithm or by sampling techniques. The input sample may come from a deterministic scheme (factorial, axial or composite grid) or randomly generated according to the input vector distribution. The choice of the method is imposed by the CPU time the model $G(\cdot)$ requires for a single run.

In case of a *central dispersion* analysis, the mean value and the variance of the outcome can be evaluated using Monte Carlo sampling, which also provides confidence intervals of the estimated values. As an alternative, it is possible to evaluate the mean of the outcomes thanks to the Taylor variance decomposition method (Equation IV.1) that requires the additional evaluation of the partial derivatives of the model $G(\cdot)$.

Finally, in case of a *threshold exceedance* criteria $\mathbb{P}[Z \geq z^\star]$, the most widespread techniques are the sampling-based ones, that is the Monte Carlo method and its variants that reduce the variance of the probability estimator (LHS, importance sampling, stratified sampling, directional sampling, subset simulation ...). All sampling-based techniques provide confidence intervals. The variance reduction techniques (or accelerated Monte Carlo methods) constitute a very active research field, the interest of which has gained importance in the last ten-twenty years, thanks to the rapid expansion of the computer experiments discipline (cf. for instance [Morio 2013] and the numerous references therein).

In case of high CPU runtime, popular alternatives (namely FORM and SORM methods) exist to estimate the exceedance probability.

They are based on isoprobabilistic transformations (the generalized Nataf transformation [Lebrun & Dutfoy 2009a, Lebrun & Dutfoy 2009b] in case the copula of the input random vector belongs to the elliptical family and the Rosenblatt one [Lebrun & Dutfoy 2009c] in the other cases) which maps the input random vector into a standard space of spherical standard Gaussian distributions. In that space, the integral defining the exceedance probability or failure probability (Equation IV.2) is approximated thanks to geometrical considerations [Dolinski 1983, Madsen *et al.* 1986]. These popular techniques provide approximations of very low exceedance probabilities with very few calls to the model, but no confidence interval is provided in order to validate the geometrical approximations.

Step C' - Sensitivity Analysis The ranking of the uncertainty's sources is based on the evaluation of some importance factors, correlation coefficients and sensitivity factors, the choice of which varies according to the quantities of interest specified in Step A. Sensitivity analysis is a wide area of investigation in the technical and scientific community. One can refer to review articles and books provided by experts in the domain (e.g. [Saltelli *et al.* 2004, Helton *et al.* 2006b, Iooss 2011, Faivre *et al.* 2013]).

Following [Iooss 2011], we distinguish the following families of methods.

- Screening methods, aiming at establishing a coarse hierarchy between a large number of input variables, with respect to their contribution to the output by means of a relatively small amounts of model's runs. Among them, it is worth noting the popular method, first proposed by [Morris 1991], consisting in randomly repeating a limited

number of times (4 or 5) one-at-a-time experimental designs (in which the values of each input are modified one-by-one, independently on the other inputs) in the (previously discretized) input space.

- Methods based on the evaluation of importance measures, based on (i) regression indices, e.g. standard regression coefficients (SRC), standard rank regression coefficients (SRRC), which provide useful informations when the relation between Z and X is linear or (at least) monotonic, or on (ii) the functional decomposition of the variance of the output Z [Hoeffding 1948, Sobol 1993]:

$$\begin{aligned} &\mathbb{V}[Z] = \textstyle\sum_i V_i[Z] + \sum_{i<j} \mathbb{V}_{ij}[Z] + \ldots \\ &V_i[Z] = \mathbb{V}\left[\mathbb{E}[Z|X_i]\right] \\ &V_{ij}[Z] = \mathbb{V}\left[\mathbb{E}[Z|X_i,X_j]\right] - V_j[Z] - V_j[Z]. \end{aligned} \tag{IV.11}$$

  According to this decomposition (which nevertheless requires the independence of the X_i's), the interpretation of the so-called Sobol sensitivity indices [Sobol 1993] $S_i = V_i[Z]/\mathbb{V}[Z]$, $S_{i,j} = V_{i,j}[Z]/\mathbb{V}[Z]$ is immediate: the first quantifies the part of the variance coming from the contribution of each of the inputs X_i's taken one-by-one, the second the part coming from the contribution of couples (X_i, X_j) and so on.
  The main issue in practice when evaluating these indices is the computational burden as many runs of the model are required to obtain a good Monte Carlo estimation of the variances of conditional expectations above. Among the computational methods proposed to cope with these issues: the FAST method, based on Fourier transformation of the function $G(\cdot)$, or the use of quasi-random sequences.
  Another class of powerful and popular methods for dealing with the computational burden of estimating Sobol indices and more generally quantities of interest theoretically demanding a very high number of Monte Carlo simulations (as probabilities of failure or low-probability quantiles) is based on building response surfaces, as sketched hereinafter.

Metamodelling. Building a response surface or a metamodel or an emulator (the three terms are equivalent in practice) of the actual CPU time-consuming model $G(\cdot)$ is another viewpoint of challenging computational issues: the computational budget is used for building an analytical function $\tilde{G}(\cdot)$ (demanding a negligible time for a single run) which provides an approximation of the model, considered satisfactory for the purposes of the study. The idea is quite intuitive and started diffusing in the computer experiments community since the 80's [Downing *et al.* 1985, Box & Draper 1987]; cf. also Section 4.

Many families of surface responses exist. They can be based on polynomials, splines, neural networks, support vector machines or fuzzy rule-based techniques. The most popular ones in the modern practice of computer experiments are polynomial chaos expansion and kriging.

- **Polynomial chaos.** The principle [Wiener 1938, Ghanem & Spanos 1990] at the base of the so-called polynomial chaos expansion is that, under the assumptions that X_i's are independent, their pdf's belong to the same parametric family and the second order moment of Z is finite, Z can be expressed as:

$$Z = \sum_{k=0}^{\infty} \alpha_k \cdot \Phi_k(\boldsymbol{x}), \tag{IV.12}$$

  where the $\Phi_k(\cdot)$ are orthogonal polynomials, belonging to a base that depends on the parametric family of the X_i's

pdf (e.g. the Hermite's family if the pdf's are Gaussian or the Legendre's family it the pdf's are uniform). The term "orthogonal" here means that:

$$\int \Phi_k(\boldsymbol{x}) \cdot \Phi_l(\boldsymbol{x}) \cdot p(\boldsymbol{x}) dx = 1 \text{ if } k = l \text{ and } 0 \text{ otherwise.}$$

In practice, the infinite sum of Equation IV.12, is truncated. The use of polynomial chaos approximation is particularly suited for the evaluation of the Sobol indices: thanks to the orthogonality properties of the polynomials $\Phi_k(\cdot)$, the indices are evaluated by means of elementary algebraic operations of the coefficients α_k [Sudret 2008]:

$$S_i = \frac{1}{\mathbb{V}[Z]} \sum_{k \in \mathcal{I}_i} \alpha_k^2, \quad S_{i,j} = \frac{1}{\mathbb{V}[Z]} \sum_{k \in \mathcal{I}_{i,j}} \alpha_k^2 \quad \ldots$$

where $\mathcal{I}_i$ is the set of the indices of all polynomials containing only terms in x_i, $\mathcal{I}_{i,j}$ is the set of the the indices of of all polynomials containing only terms in (x_i, x_j) and so on.

Thus, the only problem to be solved is the estimation of the coefficients α_k. In the non-intrusive setting, popular in the computer experiments community (cf. Section 3), once a sample of numerical experiments results has been obtained, that can be done by using projection or regression (also known as collocation) techniques, as summarized in [Sudret 2007]. Among the recent research works in this field, one can notice the ones focused on defining optimal strategies for the design of numerical experiments taking advantage of the sparsity of the polynomial chaos coefficients [Blatman & Sudret 2010, Crestaux 2011].

- **Kriging.** Initially developed in the framework of geostatistics [Matheron 1963], the kriging method[3] is applied to computer experiments since the late 80's [Sacks *et al.* 1989]. According to this method the output $Z = G(\boldsymbol{x})$ of a computer model with random input is modelled as:

$$Z = m(\boldsymbol{x}) + Y(\boldsymbol{x}), \tag{IV.13}$$

where $m(\boldsymbol{x})$ is a deterministic function of $\boldsymbol{x}$, typically a polynomial of degree 0 or 1, and $Y(\boldsymbol{x})$ is a zero-mean Gaussian process, characterized by its correlation function $R(\boldsymbol{x}, \boldsymbol{u})$, such as:

$$\mathrm{Cov}[Y(\boldsymbol{x}), Y(\boldsymbol{u})] = \sigma^2 R(\boldsymbol{x}, \boldsymbol{u}),$$

where σ^2 is the variance of $Y(\boldsymbol{x})$. A typical choices for the correlation function is the power exponential: $R(\boldsymbol{x}, \boldsymbol{u}) = \exp\left(-\sum_i \theta_i |x_i - u_i|^{q_i}\right)$. When a set of realizations of the actual model $\{\boldsymbol{x}^{(i)}, G(\boldsymbol{x}^{(i)})\}_{i=1,\ldots,n}$ is available, the kriging predictor of the function $G(\cdot)$ in a given point x is evaluated by first considering the $(n+1)$-dimensional Gaussian joint distribution of the vector $\left(G(\boldsymbol{x}), G(\boldsymbol{x}^{(1)}), \ldots, G(\boldsymbol{x}^{(n)})\right)$, then writing the (Gaussian) conditional distribution of $G(\boldsymbol{x}) | G(\boldsymbol{x}^{(1)}), \ldots, G(\boldsymbol{x}^{(n)})$. The mean and the variance of this conditional distribution, which can be explicitly written as a function of the parameters of the kriging model and of the $G(\boldsymbol{x}^{(i)})$, are taken as the predictor of $G(\boldsymbol{x})$ and as the prediction error respectively.
The kriging predictor is unbiased and its variance can be estimated. Thus, it gives an interesting additional information (the prediction error) with respect to other classes of metamodels. The Gaussian hypothesis simplifies the estimation problem in both frequentist and Bayesian settings. Moreover, as for the polynomial chaos, the ex-

[3] The method is named *kriging* in honor of Prof. Danie Gerhardus Krige (1919-2013), South-African mining engineer who first proposed it in the 50's in the domain of mining exploration.

pression of some quantities of interest (as the Sobol indices) can be explicitly obtained [Oakley & O'Hagan 2004, Marrel *et al.* 2009, Le Gratiet *et al.* 2013]. Many research works are carried on the use of Gaussian processes metamodels in computer experiments. In Sections 7.1 and 7.2 some elements are given about specific collaborative R&D frameworks, in which researches on this class of metamodels are carried.

It is worth noting, among current research topics [Roustant 2011], works concerned with the problem of effectively coping with high dimension problems. In these cases, common techniques, based on exponential or Matérn kernels may not be not suitable and a valuable alternative can be the decomposition of the kernel in a sum of kernels of lower dimension. Inference properties of additive kriging models are discussed in [Durrande *et al.* 2012]. In this context, the data-driven methodology proposed by [Muehlenstaedt *et al.* 2012] is also of particular interest: the idea is to represent interactions between variables as edges of a graph (named FANOVA graph), the cliques of which identify groups of variables with respect to which, the kriging model is additive.

Another vast area of research is concerned with the problem of effectively planning the numerical experiments needed to fit kriging metamodels. As also recalled in Chapter VIII, the strategies currently investigated aimed at exploring the input space for better approximating the response of the model in a particular area of interest (e.g. close to the limit threshold in a problem of structural reliability) [Picheny *et al.* 2010, Bect *et al.* 2012] or at satisfying space-filling properties [Roustant *et al.* 2010, Pronzato & Müller 2012].

The kriging metamodelling is also one of the main ingredient of the Bayesian statistical framework for model's validation proposed by [Kennedy & O'Hagan 2001] (cf. Chapter VIII).

6 The need for specific tools

By a practical viewpoint, the main difficulty in uncertainty analysis is the computational burden. Actually, adding a "probabilistic layer" to a deterministic calculation results in multiplying the number of the necessary calculations for obtaining the desired results. In the previous section a quick overview has been given of the specific mathematical tools needed for effectively cope with the problem. For the practical implementation of computations, it seems essential to have access to a software which at the same time (i) includes in its library advanced methods for Monte Carlo sampling, uncertainty analysis and metamodelling, (ii) makes as easy as possible the link between the probabilistic models of the inputs $\boldsymbol{X}$ and the black-box numerical code $G(\cdot)$, (iii) makes as easy at possible the distribution of computations on HPC facilities.

Actually, the experience of several years of industrial practice is that do-it-yourself solutions prove to be inefficient and may lead to very time-consuming studies. Nowadays, several software tools specifically intended to uncertainty analysis in simulation exist. A quite exhaustive review is given by [Iooss *et al.* 2011]. The software tools can be compared with respect to a number of different features: methodological content of the library, easiness to use, licence (commercial, free, open source), users community, software interface with HPC facilities.

Among the software platforms, specifically intended to non-intrusive uncertainty analysis, we particularly remind here: (i) DAKOTA[4], developed by SANDIA National Laboratory [Eldred *et al.* 1996], (ii) Uranie[5], developed by CEA [Gaudier 2010], as well as (iii) Sunset [6] [Chojnacki & Ounsy 1996] and (iv) Promethee[7] [Richet & Munoz-Zuniga 2013] developed by IRSN . They are all distributed under free licenses.

[4] http://dakota.sandia.gov
[5] http://sourceforge.net/projects/uranie/
[6] https://gforge.irsn.fr/gf/project/sunset/
[7] http://promethee.irsn.org

In the remainder of this section, the focus is put on the OpenTURNS software.

OpenTURNS: Treatmet of Uncertainties, Risk'N Statistics. OpenTURNS[8] [Dutfoy *et al.* 2009] is an open source software specifically designed to put into practice the methodology sketched in Section 3 (Figure IV.2). It is jointly developed since 2005 by EDF R&D, Airbus Group and Phimeca, and distributed since 2007. Running under the Windows and Linux environments, OpenTURNS is basically a C++ library proposing a Python textual interface. It can be linked to any code communicating through input-output files (thanks to generic wrapping files) or to any Python-written functions. It also proposes standard interface for complex coupling with external black-box computer codes.

Gradients of the external code are taken into account when available and otherwise can be approximated automatically by finite differences schemes. In addition to its more than 40 continuous/discrete univariate/multivariate distributions, OpenTURNS proposes several dependence models based on copulas (independent, empirical, Clayton, Frank, Gaussian, Gumbel, Sklar). It offers a great variety of definitions of a multivariate distribution, in particular: list of univariate marginals and the copula and linear combination of probability density functions or random variables. The propagation step is covered through numerous simulation algorithms (importance and directional sampling, subset simulation, Latin hypercube sampling, quasi Monte Carlo sampling etc.). The innovative Generalized Nataf and Rosenblatt isoprobabilistic transformations are implemented for performing the FORM/SORM methods and, more generally, methods based on sampling in the standard Gaussian space. For the ranking analysis, Sobol indices, and the usual statistical correlation coefficients are available.

A rich documentation of more than 1000 pages is at users' disposal; it is dispatched within a number of documents covering all the aspects of the platform: scientific guidelines (Reference Guide), end-user guides (Use Cases Guide, User Manual and Example Guide) and some software documentations (Architecture Guide, Wrapper Guide, Contribution Guide and Windows port Guide).

OpenTURNS implements some high performance computing facilities such as the parallelism of algorithms manipulating large data set (up to 10^8 scalars) using the threading building blocks (TBB) technology. For the distribution of computations, many strategies are possible [Barate 2013]; among them: (i) the distributed Python wrapper and (ii) the use of the supervision module YACS of the SALOME platform (cf. Section 2).

The software is innovative by its input data model, based on the multivariate cumulative distribution function (cdf), which enables the usual sampling-based approach (statistical manipulation of large data set) but also the analytical approach: if possible, the exact final cdf is determined (thanks to characteristic functions implemented for each distribution, the Poisson summation formula, the Cauchy integral formula ...). Furthermore, different sophisticated mechanisms are proposed: aggregation of copulas, composition of functions from $\mathbb{R}^n$ into $\mathbb{R}^p$, extraction of copula and marginals from any distribution.

OpenTURNS implements some up-to-date efficient sampling algorithms: it uses the Mersenne Twister Algorithm to generate uniform random variables [Saito & Matsumoto 2006], the Tsang & Marsaglia method for Gamma variables [Marsaglia & Tsang 2000], the Ziggurat method for normal variables [Doornik 2005] and the Sequential Rejection Method for binomial variables. The exact Kolmogorov statistics is evaluated with the Marsaglia Method and the Non Central Student and Non Central χ^2 distribution with the Benton & Krishnamoorthy method [Benton & Krishnamoorthy 2003].

OpenTURNS is the repository of some recent results of PhD researches carried at EDF R&D as sparse polynomial chaos expansion based on the Least-angle regression (LARS) method [Blatman 2009, Blatman & Sudret 2010] and Adaptive Directional Stratified sampling for estimating failure probabilities [Munoz-Zuniga 2011, Munoz-Zuniga *et al.* 2012].

[8] http://www.openturns.org/

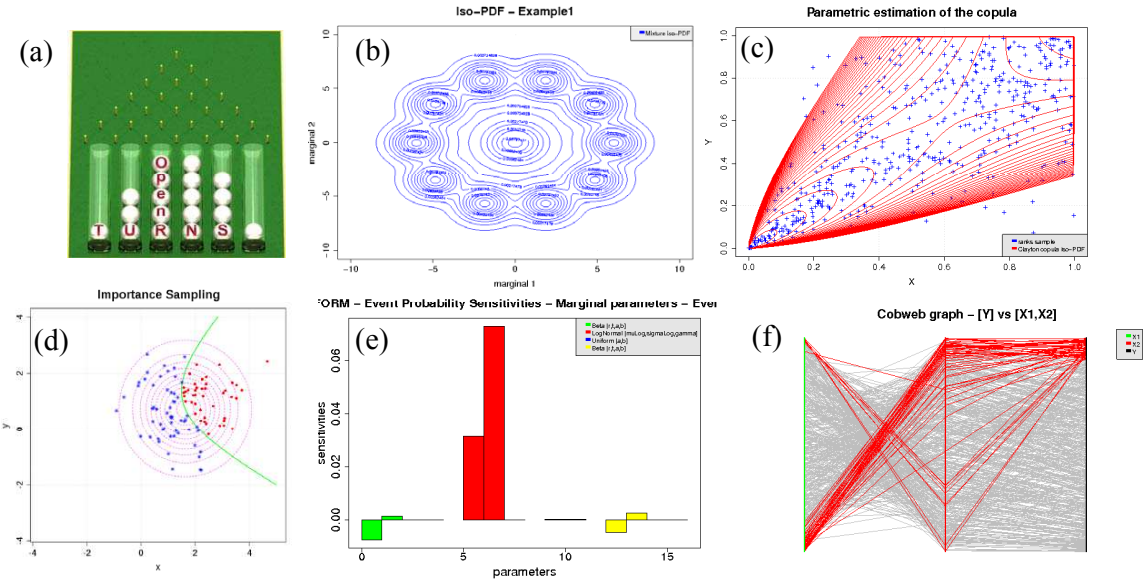


Figure IV.3 – Some Open TURNS snapshots: (a) the Open TURNS logo inspired by Galton's box experience, (b) modelling a multi-modal random vector of $\mathbb{R}^2$, (c) copula's fitting, (d) importance sampling in the Standard Space around the FORM design point, (e) FORM importance factors, (f) cobweb plots for visual and intuitive sensitivity analysis.

7 A very active research community

7.1 Working groups and research consortia

As already stated hereinbefore, the practice of uncertainty analysis and (more generally) probabilistic calculation is firmly fixed in the engineering studies of several industrial companies, in most cases within risk assessment contexts (e.g. hydraulic, nuclear or financial risk). The underlying mathematical tools are also well established theoretical and applied research topics since several years.

Indeed (cf. page 69), there has been a considerable rise of interest in many industries in the last years for these problems. Facing the questioning of their internal and/or external control authorities in an increasing number of different domains or businesses, large industrial companies have felt that domain-specific approaches were no more appropriate. In spite of the diversity of terminologies, most of these methods do share many common algorithms: a rather new need for a more adapted methodological support and mathematical tools arose quite simultaneously in several industrial companies and public establishments (with a major role, among others, of EDF, CEA, Airbus-Group), which decided to work more closely and, at the same time, turned to academic research as the growing complexity of the raised problems clearly put into evidence the limit of the standard mathematical tools belonging to engineers' background.

The first premises of these collaborative frameworks were put around 2003, and most of the dedicated working groups started their activity in 2005-2006, under different configurations: thematic group of scientific societies (IMdR, ESReDA, SFdS), industry-academic research (DICE) or software-development consortia (Open TURNS).

The project group *Uncertainty* (2005-2008) hosted by the ESReDA society (European Safety, Reliability & Data Association) and the working group *Incertitudes et industrie* (since 2006) of the IMdR (Institut pour la Maîtrise des Risques) were both driven by the same needs: defining a common methodological framework for uncertainty analysis and largely spreading the "uncertainty analysis viewpoint" in the scientific and industrial communities.

The main result of the ESReDA group (EDF, CEA, Airbus-Group, JRC, SAFRAN, Delft and Duisburg-Essen Universities) has been the collective book [de Rocquigny *et al.* 2008] which, starting from several industrial needs and examples, sketches the principles of the common methodological framework summarized in Section 3. The most significant achievement of the IMdR group, (with participants, among others, from EDF, CEA, EADS, IFP, IRSN, ONERA, INERIS) has been the definition of specialized training course about uncertainties in simulation (in three sessions), named "*Managez les incertitudes dans vos études*", to largely share and transfer the more common methods and tools.

In 2009 a new thematic group *Fiabilité et Incertitudes* was founded within the Société Française de Statistique (SFdS). The group contributes to spread and to promote uncertainty analysis in simulation within the French statistical community. Nowadays, special sessions about these topics are held within the annual *Journées de Statistique* or in congress supported by SFdS (e.g. ISI World Congress).

The DICE consortium (*Deep Inside Computer Experiments*, 2006-2009) was essentially devoted to research works about the design of computer experiments and the exploration of large and time consuming numerical code. Mainly driven by advanced research problems, under the lead of the École des Mines de St. Etienne (EMSE), the consortium gathered academic and industrial partners (Total, Renault, EDF, IRSN, ONERA) which funded the project. The main results of DICE are collected in open source packages of the R software [Roustant *et al.* 2012]. The activities of the DICE consortium continue in the framework of the ReDICE[9] project (2011-2015), also particularly focused on metamodelling techniques, under the lead of the University of Bern with almost the same hard core of academic and industrial partners as DICE.

Finally, concerning more specifically advanced research works, the most important French structuring framework is nowadays the GdR MASCOT-NUM[10] (*Méthodes d'Analyse Stochastique pour les Codes et Traitements Numériques*). Created in 2008, under the supervision of CNRS, it aims at coordinating research efforts in the scientific area of design, modelling and analysis of computer experiments. Mainly positioned on advanced research methods, it provides a framework for discussing and presenting research works, by organizing events like the *Annual GdR meeting*, as well as seminars and workshops. It is worth noting that, to the best of our knowledge, no academic Department exists specifically intended to uncertainty analysis and computer experiments, that is, in practice, the GdR MASCOT-NUM acts as a sort of "Informal Laboratory", gathering researchers dispersed in Applied Mathematics Departments of several French institutions and provide a paramount role in structuring the community.

A similar role is played by the MUCM Community (Managing Uncertainties in Computer Model) in UK (and hopefully at European scale in a next future) and the SIAM (Society for Industrial and Applied Mathematics) Activity Group on Uncertainty Quantification in USA.

7.2 The ANR OPUS project

This Section provides a quick overview of the ANR OPUS project (acronym of *Open source Platform for Uncertainty treatment in Simulation*) that I coordinated between January 2009 and September 2011. More details can be found in the final report of the project [OPUS 2011] and the references therein.

7.2.1 Context and background

The idea of the OPUS project raised within the French uncertainty analysis community in late 2006. Under the basis of a consolidated common framework, OPUS partners had the ambition to work together in a more structured way by building

[9] http://www.redice-project.org/
[10] http://www.gdr-mascotnum.fr

a funded research program.

The project is the merging of two different projects, both proposed to the *Agence Nationale de la Recherche* (ANR) in 2007: OPUS (mainly oriented towards software issues) submitted to the Program *Software Technologies* and COPRIN (*COnception en PRésence d'INcertitudes de systèmes complexes multi-physiques*) submitted to the Program *Intensive Computing and Simulation* (mainly focused on research works).

Taking into account the technical proximity and the interesting complementarity between the two projects, the ANR asked to the project teams to merge their technical programs in an unique project. That gives raise to the OPUS project, in its actual form (duration: from April 2008 to September 2011, global budget: 2.24 M€, ANR's fundings: 0.94 M€).

The OPUS project team reflected an interesting variety of points of view, scientific backgrounds and possible utilizations of project's results. The OPUS consortium comprised ten partners:

- four industrial partners: CEA, Airbus-Group, EDF R&D (coordinator), Dassault-Aviation (DA),
- five academic partners: École Centrale Paris (ECP), SUPÉLEC, Université Diderot Paris 7 (UP7), Université Joseph Fourier Grenoble 1 (UJF), INRIA, which took part in OPUS through the Scilab team (Foundation Digitéo),
- one SME working in the computer science service business: Softia.

The interesting "biodiversity" of the OPUS partners allowed an actual interaction between industry and research. That has been, finally, together with the "classical" deliverables of the project (reports, software codes, scientific papers etc.), the major result of the project. The generic aim of OPUS was "*to create and sustain an activity around generic uncertainty treatments by building and maintaining an integrated open source platform*"; more specifically:

- creating and disseminating open source tools for uncertainty, treatment, using cutting-edge algorithms provided by the scientific community,
- capitalizing the French know-how in uncertainty analysis and let a reference community arises,
- creating a lasting dynamics between different academic, industrial and business partners.

According to these principles, the OPUS collaborative works were based on a "loop" (cf. Figure IV.4) that starts from real industrial use-cases, arising from different business areas (energy production, nuclear safety, aerospace). The complex mathematical and numerical treatments to be performed within these use-cases let rise the need for advanced methods, coming from academic and scientific research. Once these methods have been developed and tested, they are integrated and perpetuated inside lasting software platforms, to be reused in other similar industrial studies. It is worth noting that in this figure the "OPUS world" (represented by the light-blue ellipse) has been widely open to contributions, represented by blue arrows in Figure IV.4, coming from the whole "uncertainty analysis" community.

7.2.2 Project Structure

The OPUS works were organized in the following work-packages:

- **WP0 and WP0': Coordination, Dissemination & Communication (WP0) and Expert College (WP0').** Besides the common charges of technical and administrative follow-up, animation and reporting, a great place has been left within OPUS to dissemination and communication activities. That is strictly related to the particular context of the activities around Uncertainty Analysis and Computer Experiments. As several thematic working groups exist in France and abroad, it was essential to involve a larger community than the OPUS team. For better defining the

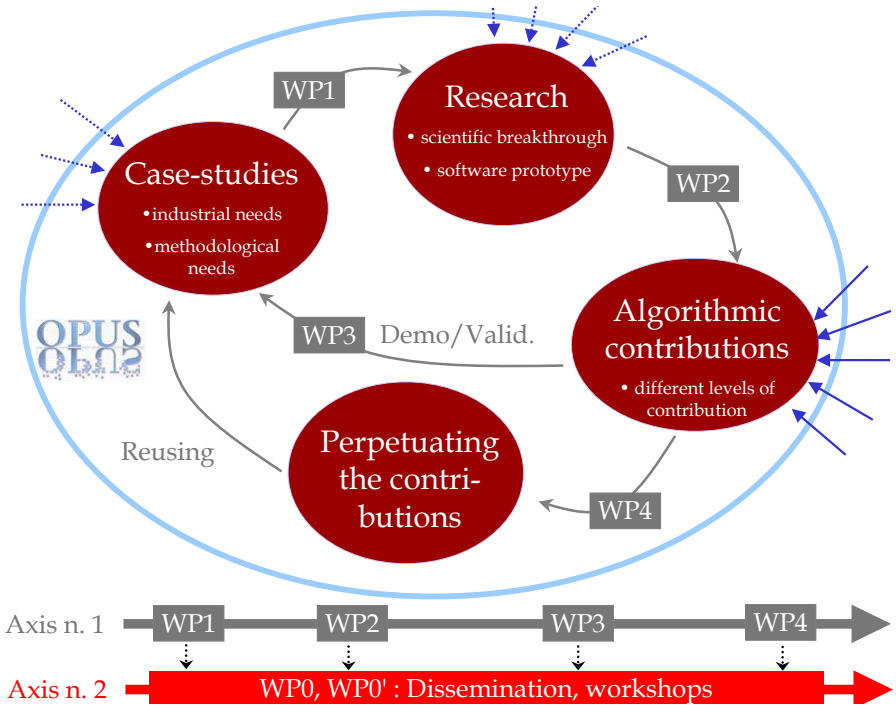


Figure IV.4 – The OPUS loop: from the industrial needs to reusable software contributions.

content and the targets of these works, a specific work-package (WP0') was created to host the permanent activity of project team specialists within an Expert College, lead by Anestis Antoniadis (UJF). WP0's mission also included the reviewing and endorsement of project deliverables. Besides scientific communications and publications (main outcomes of the WP2), the dissemination activities have been: six thematic workshops, the editing of a special issue of the scientific review *Statistics and Computing*, a special session of the 42th *Journées de Statistique* (2010), as well as several presentations during the periodical meetings of ANR and *Systématic Paris-Région* Business Cluster.

- **WP1: User requirements and specifications.** The goal of this work-package was to identify, throughout the first year of the project, a number of methodological and software needs and to fix the bases of partners' common work (in particular, a set of industrial use-cases, cf. Figure IV.5).

- **WP2: Scientific developments.** This was the "Research" work-package of OPUS, the major part of its content coming from the COPRIN project. The activities turned around several classes of problems/methods with different degrees of maturity:

  - "pre-industrial research" (WP2.1): problems and methods already relatively mature and, in some way, ready to be spread within the industrial community: metamodels based approximation, sensitivity analysis, probabilistic inverse modelling,

  - "upstream research" (WP2.2): more prospective methods and problems, to be treated in a rather academical framework: robust low probability quantiles estimation, uncertainty quantification in heterogeneous models couplings, robust and real-time implementation of parametric partial differential equations (intrusive methods).

- **WP3: Validation / Demonstration.** This work-package hosted the activity of testing and validating the methods and the algorithms (at least, the ones the WP2.1 dealt with) by putting into practice exemplary studies on the use-cases proposed by the industrial partners (WP1).
- **WP4: Industrialization / Product durability.** Here the word "Industrialization" is not to be interpreted in the strict sense of software engineering. The works carried on this work-package aimed at: (i) defining and structuring the types of contribution, (ii) making the software contributions interoperable with other commonly used software components and environments, (iii) ensuring a certain durability to the results of the project.

It is worth noting that the OPUS works followed two mainstreams (see Figure IV.4). The first one followed the logical chain of the work-packages from WP1 to WP4. The second mainstream (work-packages WP0 and WP0') concerned the transverse activities of communication and dissemination which took place all along the project duration and were constantly fed by the works of the other work-packages.

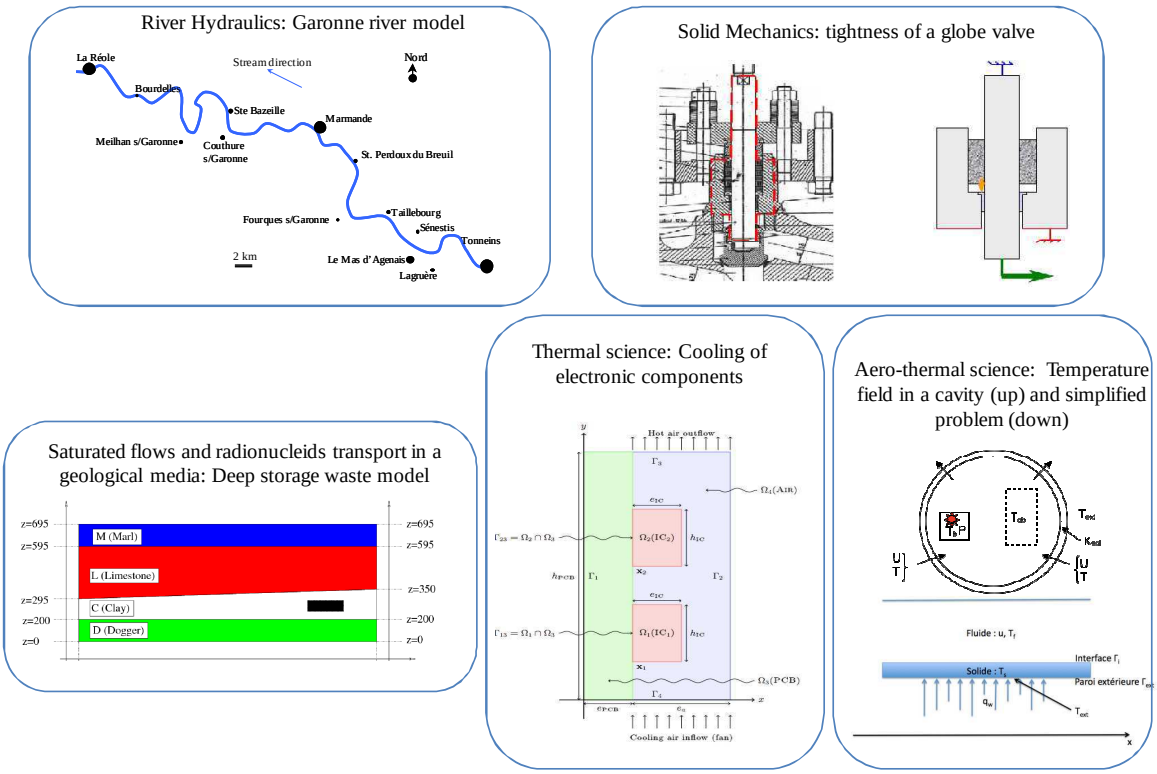


Figure IV.5 – The five industrial use cases proposed by the industrial partners of the OPUS project.

7.2.3 Some achievements

Research and dissemination. The research works, mainly carried within the WP2, provided many interesting methodological and algorithmic results.

An innovative adaptive strategy for building "specialized" kriging metamodels (i.e. specifically intended to the effective estimation of a given quantity of interest, as a probability of failure, by improving the approximation properties in proximity of the failure threshold) was formalized [Bect *et al.* 2012] and applied to a flood risk assessment prob-

lem [Arnaud *et al.* 2010]. The strategy, named SUR (stepwise uncertainty reduction) is actually based on the Bayesian decision-theory formulation of the estimation problem

A method, based on multi-element generalized polynomial chaos (gPC) metamodel, for estimating low-probabilities α-quantiles z_α of the output of a CPU time consuming numerical code $G(\boldsymbol{x})$, has been proposed by [Ko & Garnier 2013]. The ordinary non-intrusive strategy is improved by combining a "global" metamodel in the standard Gaussian space of the inputs with auxiliary "local" metamodels constructed in bounded domains about the design points (i.e. values of the input $\boldsymbol{X}$ likely to be mapped into values close to z_α).

The reduced bases method - a discretization approach consisting in approximating the solutions of parametrized partial differential equations (PDE's) with a linear combination of the elements of a basis formed by previously known (offline) solutions of the equation for given and ad hoc chosen parameters - has been also investigated in the framework of the OPUS project. This work lead to significant improvements in terms of methodology (analysis of the convergence properties of the so-called *greedy algorithm*, the mainly used strategy to choose the elements of the basis [Buffa *et al.* 2012]) and diffusion (application to use-cases [Vallaghé *et al.* 2011] and contribution to the development of the software library Feel++ [Prud'Homme *et al.* 2012])

The already cited works concerning the effective building of sparse non-intrusive polynomial chaos expansions and the adaptive directional sampling for estimating probabilities of failure were also partially hosted by OPUS.

The scientific production of the project has been significant: 10 papers published in peer-reviewed journals, as well as 15 communications to national and international conferences and several overall presentations of the project during institutional workshops organized by ANR and the Business Cluster Systématic Paris - Région.

The OPUS project organized six workshops including the closing one, covering both scientific/methodological and software implementation issues (it is worth highlighting that some of them have been jointly held with the GdR MASCOT-NUM and the working group *Fiabilité et Incertitudes* of the SFdS):

- Metamodelling and free software, EDF R&D, Clamart, October 2008,
- Learning and model selection, CEA, Saclay, April 2009,
- Spectral methods and polynomial chaos, Airbus-Group, Suresnes, November 2009,
- Uncertainty propagation, estimation of rare quantiles and low probabilities of failure, Institut Henri Poincaré, Paris, June 2010,
- Uncertainty quantification, high performance computing, calculation environments and software, University Joseph Fourier, Grenoble, March 2011,
- Numerical simulation and uncertainty analysis, OPUS closing workshop, Institut Henri Poincaré, Paris, October 2011.

Largely opened also to public and contributors not officially involved in the project, these biannual workshops have been a regular meeting place for the French technical and scientific community all along the period 2008-2011.

Algorithms and computer codes. The algorithmic contributions to OPUS project have been grouped into three categories, named OPUS: OPUS-Lib, OPUS-Contrib and OPUS-Forum respectively. These categories correspond to different levels of compliance to more and more strict coding rules. OpenTURNS (cf. Section 6) is considered as a prerequisite of the platform. The main requirements the code must fulfil (depending on the category) are sketched below.

- **OPUS-Lib.** The code fulfils a high level of quality, following established OPUS programming rules. This means a high level of quality and integration. In a nutshell, it must present the following features: (i) functionality demonstrated in an academic and/or R&D environment, possibly with a preliminary code proposed to the community for reviewing purpose, (ii) system software architecture development initiated to include interoperability, reliability, maintainability, extensibility, scalability, and security issues, (iii) a high level of documentation, (iv) basic software components are integrated to establish that they will work together.
- **OPUS-Contrib.** The code has a fair level of quality but most of restricting rules proper to to the -Lib level are released. Contributions at this level can be written in any of the following supported languages: C++, Python, R, Scilab, Matlab/Octave, and can have different levels of integration with respect to the other ones (compatibility or interoperability), for instance, a C++ code (as well as a Python or Scilab script) using -Lib and -Contrib features.
- **OPUS-Forum.** The contributor uses the forum space available in the OPUS web site, to discuss and propose features or ideas in a completely free way, from a simple algorithm to a piece of code or even a complete module (albeit incompatible with other OPUS contributions).

Table IV.1 presents a schematic view list of the contributions to the OPUS project.

Name	Language	Description	Level
RPyWrapper	Python	Wrap R in Python/Op.TURNS	Lib
PC OpenTURNS	C++, Python	Polynomial Chaos	Lib
NISP	C++, Scilab, Python	Polynomial Chaos	Lib
Feel++/Opus	C++, Python, Octave	Reduced Basis metamodel	Lib
Kriging STK	Matlab, Octave	Kriging-based metamodels	Contrib
MLE Inverse	R	Maximum likelihood estimations for inverse problems	Contrib
MCMC Inverse	R	Monte Carlo Markov chain for inverse problems	Contrib
Quantile		Quantile estimations	Forum
Funct. SA	R	Functional sensitivity analysis	Forum

Tableau IV.1 – Production of algorithms and computer codes of the OPUS project.

Part of the development of the polynomial chaos expansion in the OpenTURNS software was done within the OPUS framework[11] It is worth noting also the development of the library NISP (Non Intrusive Spectral Projection) under the form of a Scilab package [Baudin & Martinez 2010]. Definitely, the OPUS project gave a strong contribution to the diffusion of the non-intrusive polynomial chaos towards the technical community.

Feel++/OPUS is a framework for the reduced basis approximation of PDE's. It essentially provides a C++ interface for finite element codes and an implementation of some specific methods. The software is provided with several test-cases, mainly concerned with heat transfer. For each test-case, the user can manipulate various features of the approximation method (e.g. the reduced basis of functions) using different software environment as Python (using Open TURNS wrapping system) or Octave. Python and Octave scripts are also provided as examples

Among the OPUS-Contrib contributions, it is worth noting the development of a toolbox for kriging metamodelling in Matlab® language (and fully compatible with the free software Octave), named STK ("Small" Toolbox for Kriging). The

[11] Notice that this is a joint work of Airbus-Group, EDF and Phiméca. The contribution of Airbus-Group to these developments has been mainly funded by OPUS, while the contribution of EDF and Phiméca has been funded by own resources and by the ANR project MIRADOR (*Modélisation interactive des risques associés au développement d'ouvrages robustes*).

main features of STK are the implementation of: (i) a number of covariance functions, and tools to compute covariance vectors and matrices, (ii) a specific procedure (named REMAP) for estimating the parameters of the covariance from available data, (iii) advanced prediction procedures.

Other significant OPUS-Contrib contributions are R scripts for solving probabilistic inverse problems using different numerical methods (likelihood maximization by means of ECME *Expectation Conditional Maximization Either* and S(A)EM *Stochastic Approximation of Expectation Maximization*) and Bayesian posterior sampling by means of Metropolis-Hastings algorithm. These methods have been applied to the hydraulic use case (Garonne river) for inferring the probability distribution of the Strickler roughness coefficient [Couplet *et al.* 2010].

7.3 Other collaborative funded projects

OPUS has been the first French multi-partners project specifically focused on generic methods and tools for assessing uncertainties in numerical simulation. Other related projects on this topic were launched in the years 2009-2011. The CSDL project (*Complex Systems Design Lab*, 2009-2012), gathering 28 partners (large, intermediate and small companies and research institutes) under the coordination of Dassault-Aviation, has been one of the biggest project of the Business Cluster Systématic - Paris Région. Several OPUS partners were involved in CSDL, as EDF R&D, Airbus-Group, SUPÉLEC. This project was mainly oriented to design issues with application to aerospace and automotive manufacturing, and had a specific workflow concerning uncertainty quantification and metamodelling. Indeed, taking into account uncertainties tainting the predicted performance of a system, since the very early stage of the design process, is considered an important and challenging issue by manufacturers.

The Costa-Brava ANR project (*Complex spatio-temporal dynamics analysis by model reduction and sensitivity analysis*) [Gamboa 2013] aims to provide novel mathematical tools combining stochastic and deterministic approaches to sensitivity analysis for particularly complex computer models (complex physics involved, high CPU-time consuming, large dimension of both inputs and outputs). Computer models representing complex spatio-temporal dynamics are particularly targeted (e.g. large scale meteorological models).

Costa-Brava started in January 2010 for a duration of four years, involving the Institut de Mathématiques de Toulouse (coordinator), the University Joseph Fourier, CEA and IFP Énergies Nouvelles.

The HAMM ANR project (*Hybrid Architectures and Multiscale Methods*, 2010-2014) deals with more specific software features concerning computer experiments. In particular, it aims at the development, analysis and software implementation of mathematical models for multiscale applications on hybrid architectures. Large scale multiscale applications are indeed within reach thanks to the emerging computing infrastructures, but they require accurate and robust multiscale numerical methods that take into account these new architectures. This is very challenging as current software tools were not designed for these methods and architectures. Algorithmic and software developments of reduced bases methods initiated during OPUS continue within the HAMM framework. The project gathers four partners: University Joseph Fourier (coordinator), IFP Énergies Nouvelles, Bull and CEA.

The same will of getting closer future software requirements and future HPC hardware and middleware solutions inspires the European projects EESI (*European Exascale Software Initiative*, 2010-2011) and its sequel EESI2 (2012-2015). It is worth noting that EESI2 hosts a specific transverse task "Verification, Validation and Uncertainty Quantification".

Finally, on November 2013, the CHORUS (*Common Horizons of Research on Uncertainties in Simulation*) ANR project (actually, the sequel of OPUS) officially started for a duration of four years. Built around a hard core constituted by former OPUS partners, the project will be concerned from a methodological viewpoint with statistical validation of numerical codes, complexity reduction using structured approximations for non intrusive metamodelling, weakly intru-

sive and goal-oriented model reduction for parametrized partial derivatives equations, goal-oriented sampling and multi-fidelity models. From a software viewpoint a great attention will be given to high performance computing issues as well as the industrialization of the developed codes within the OpenTURNS platform. For this reason, the project team has been reinforced, with respect to OPUS, with additional software engineering skilled partners.

Chapter V

The extra-probabilistic temptation

Lo que pretendo decir es que ese hombre serìa capaz, a su manera, de calcular con bastante exactitud el marco de probabilidades. Imagine una màquina donde metiera todos sos datos de los que hemos hablado y diese como resultado un lugar exacto y una hora aproximada [...]
Una bocanada de humo vela las facciones del policìa. Apoya los codos en la mesa, interesado.
- Probabilidades, dice... ¿Eso es calculable?
- Hasta cierto punto.
Arturo Pérez-Reverte, El Asedio (2010)

What I was trying to say is that this man is somehow capable of calculating the range of probabilities with considerable accuracy. Imagine you could feed all the data we discussed earlier into a machine that would give you an exact location and an approximate time [...]
A cloud of smoke briefly veils the policeman's face. He leans his elbows on the table.
- Probabilities, you say, and this can be calculated?
- Up to a point.
Arturo Pérez-Reverte, The Siege (translated by Frank Wynne, 2013)

Reading notes

Technical context. Most of the works concerned with the problem of uncertainties in simulation are made within the probabilistic framework; that is, the uncertainties tainting all quantities than act in the problem are described by probability distributions. Other mathematical settings exist for describing and quantifying uncertainties. These settings are often proposed by scientists and practitioners as an alternative to the mainly used probabilistic methods and tools. Sometimes, the proposal is also accompanied with a (more or less harsh) criticism of probabilities, which could prove possibly inappropriate for dealing with some kinds of problems, alternately for theoretical and practical reasons.

I won't take part in this technical quarrel. Adopting a pragmatic viewpoint, I have been interested in studying and using these alternative settings for solving specific problems. In particular, in the remainder some details will be given on works that I carried on (i) fuzzy rule-based metamodelling and (ii) the propagation of "hybrid" uncertainties, i.e. a framework in which some uncertainties are described by probability distributions and other by possibility distributions.

The study concerned with fuzzy metamodelling was carried in 1999-2000 when I worked for the Dept. of Civil Engineering of the "Seconda Università di Napoli", within the framework of a research project focused on the simulation of pollutants dispersion processes occurring in rivers and estuaries.

The works about hybrid possibilistic-probabilistic uncertainty propagation were part of my activity at EDF R&D since 2008-2009 about uncertainties in numerical simulation. They have been carried in cooperation with the Politecnico

di Milano (2010-2012), then with the the Chair on *Systems Science and the Energy Challenge* (École Centrale Paris / SUPÉLEC). Extra-probabilistic frameworks are topics of great interest within the safety and risk analysis communities and their investigation proved useful to have an engineer's viewpoint of the state of the art methods and tools.

Contributions. As far as my communications and publications related with these topics and activities, are concerned:

- the works on fuzzy metamodelling of pollutants dispersion have been presented at the *Belgium Fuzzy 2000* conference [Di Natale *et al.* 2000] and later published in the *Belgian Journal of Operations Research, Statistics and Computer Science*[1] [Di Natale *et al.* 2001];
- the works on hybrid possibilistic-probabilistic uncertainty quantification have been the subject of three communications [Baraldi *et al.* 2011, Pedroni *et al.* 2012, Pedroni *et al.* 2013b] at the ESREL (*European Safety and Reliability Conference*) conferences, as well as an article [Pedroni *et al.* 2013a] published in the journal *Computers and Structures* and a second article [Pedroni *et al.* 2014], submitted and currently under revision.

Structure of the chapter and credits. Most of Section 1 is excerpted from the article [Di Natale *et al.* 2001] while Section 2 is a summary of a a number of papers and presentations co-authored with colleagues of the Politecnico di Milano and École Centrale Paris / SUPÉLEC (especially, [Pedroni *et al.* 2013a] and [Pedroni *et al.* 2014]).

1 A very early study: fuzzy rule-based metamodel of river pollution

Fuzzy sets logic is essentially a mathematical tool imagined for coping with uncertainty. Formalized in the 60's, first applied to control problems, then as a modelling technique, fuzzy rules prove to be valuable tools to model complex systems in presence of uncertainties, as well as to build approximate models of deterministic "well defined" systems.

The monograph of [Bárdossy & Duckstein 1995] provides a synthetic but complete overview of the fundamentals of fuzzy rule-based models and present some applications in different engineering domains (e.g. soil water movement, reservoir operation). Basically, one can distinguish two different approaches in fuzzy metamodelling, that could be somehow qualified (using currently popular terms in the framework of computer experiments) as "intrusive" and "non-intrusive" respectively. In the first case, fuzzy rules are built (more or less directly) from the structural equations of the phenomenon under investigation. For instance, [Özelkan & Duckstein 2001] developed a fuzzy rainfall-runoff model, designed from physically based equations aiming at modelling the hydrologic cycle. As another example, [Tran *et al.* 2002] revisited in a fuzzy fashion the popular "universal soil loss equation", improving its predictive properties.

In the second case, fuzzy rules are determined from data coming from actual or computer experiments, e.g. rule-based modelling of the link between large-scale atmospheric circulation patterns (and possibly climate anomalies as *El Niño - Southern Oscillation*) and regional-scale precipitations [Pesti *et al.* 1996, Galambosi *et al.* 1999], link between raw material properties and manufacturing process variables and the hardness of steel sintered components [Chatterjee *et al.* 2008], link between regional-scale environmental and human health indicators [Canavese & Ortega 2013] etc.

This second use (metamodelling) is the one we are mainly interested in, within the context of this document.

Fuzzy metamodels are often advocated in optimization problems, possibly demanding a high number of runs of a numerical code. For instance, [Kamali *et al.* 2005] makes use of a fuzzy rule-based metamodelling for finding out

[1] This journal, also known as *JORBEL*, has been published, up to 2003, by the Belgian Operation Research Society. In 2003 it has been "merged" with its French and Italian counterparts, giving rise to the *4OR* Journal (*Quarterly Journal of the Belgian, French and Italian Operations Research Societies*).

the best set of parameters of hydrologic models. Here, metamodels allow a quite exhaustive exploration of the input space. [Haberlandt *et al.* 2002] used a fuzzy rule-based response surface to model nitrogen leaching from arable lands. The issue is to provide forecasts at a regional level by means of several leaching assessments made at the patch scale, depending on very local parameters as climate, nature of soils or management conditions (crop rotations, fertilization). The *upscaling* of the patch scale models is computationally heavy and the CPU time needed makes the use of the actual models inappropriate for being integrated in a decision support system at regional scale.

As an interesting example of joint use of fuzzy and probabilistic methods, [Wang & Simpson 2004] propose to use fuzzy clustering techniques to identify interesting zones in the input space, on which refining kriging metamodels in design-optimization problems.

The motivation for the work presented hereinafter is slightly different: in some cases for different practical reasons, one could need very fast forecasts of variables of interest for bringing some light about decisions to be made in a very short period. For instance, in case of accidental release of a pollutant in a river, a fast evaluation of the arrival time of the pollutant front to a given point of the river or the estuary can bring more light to the decision of fast risk mitigating measures (e.g. bathing prohibition). The usefulness of fuzzy rule-based fast predictive models in environmental management and control is also advocated by [Woldt *et al.* 1997] or [Theisen & Glesner 1998].

1.1 The problem

The goal of water quality models is the simulation of changes in pollutants concentration as they move through the liquid environment [James 1993]. Some pollutants are practically inert and so the concentration variations are due only to advective-diffusive transport phenomena. Such a behaviour is typical of heavy metals (Cr, Cu, Ni, Hg) and many other substances dangerous for human health. The problem becomes more complex for the pollutants whose concentration changes depend also on chemical and biological processes which are superimposed to transport phenomena.

We do not deal here with the specific problem of reactive pollutants.

Introducing a polluting substance to environment gives rise to a chain of interactions between several environmental components, the effects of which can not always be fully taken in account in modelling. The main mechanisms which intervene in the pollutant transport in a water body are: (i) advection (transport due to the bulk movement of the water in which solute is contained), (ii) diffusion (transport due to the migration of particles essentially under the effects of turbulent eddies) and (iii) dispersion (migration of particles due to velocity shears).

I cannot resist the temptation of quoting the following description, full of imagery, of diffusion and dispersion by [Fischer 1968]: "*suppose that randomly walking drunks are getting on and off of busses in a random way, but that the busses operate on a fixed schedule. [...] Dispersion in a river is very similar to the 'drunk on a bus' problem. [...] The bus schedule is analogous to the variation of longitudinal velocities within the cross section. Dispersion is caused primarily by 'bussing' of the particles, that is, convection at the different velocities of different stream lines. The primary effect of the turbulence is to cause the particles to change busses.*"

Let us consider the problem of the propagation of a pollutant front in a river or a channel, due to the release of a pollutant at the abscissa $x = 0$. Depending on the distance from the release point, the influence of the diffusion and dispersion is different. According to standard practice, three zones are distinguished (named near, mid and far zones of mixing respectively, cf. Figure V.1) in which the dominant phenomenon (besides advection) is vertical diffusion, transversal diffusion and longitudinal dispersion respectively. This distinction, based on a deep theoretical understanding of mixing phenomena [Fischer *et al.* 1979, Rutherford 1994] allows simplifying models but may be tricky to transpose in practice to real large rivers for evaluating actual water quality indicators for regulatory purposes (cf. the interesting

discussion in [Jirka *et al.* 2004]). If one is mainly interested in what happens quite far from the release point, the pollutant transport can be well enough described by the 1D advection-dispersion equation:

$$\frac{\partial C}{\partial t} + v_x \frac{\partial C}{\partial x} - D_x \frac{\partial^2 C}{dx^2} = 0, \tag{V.1}$$

in which C and v_x are the cross sectional averaged values of the concentration and the velocity, and D_x is an effective longitudinal dispersion coefficient which takes into account the effects on C of variations of velocity across the channel cross-section [Rutherford 1994]. In practice, D_x depends on mean velocity v_x as well as geometrical and physical river features, i.e. depth, width, shear velocity [Fischer *et al.* 1979, Kashefipour & Falconer 2002]. Even in its simplified form (Equation V.1) the advective-diffusive transport equation has no general analytical solution and it has to be solved numerically [Hirsch 1988, Fletcher 1990].

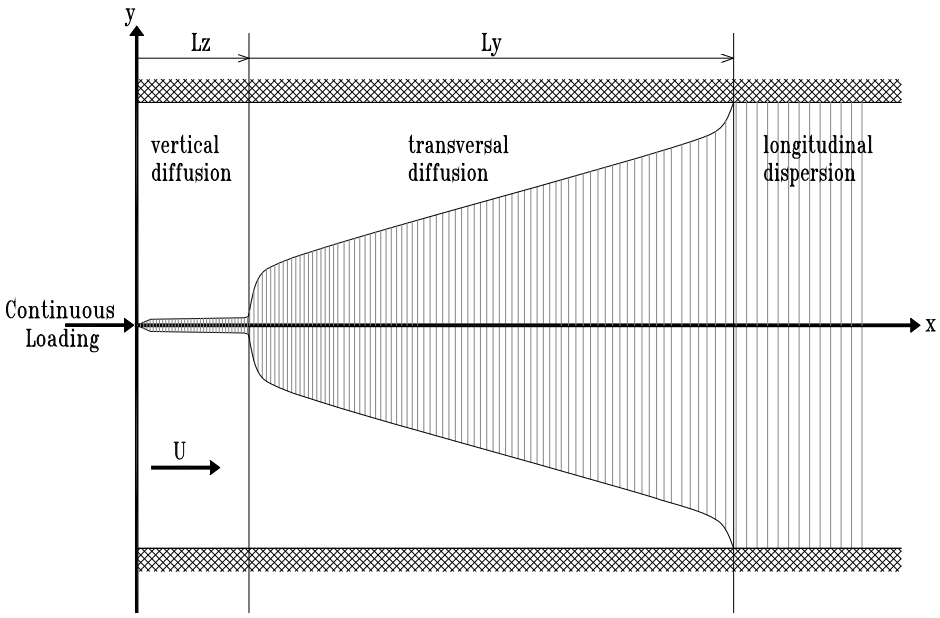


Figure V.1 – Schematic view of the different zones ("zones of mixing") concerned with the different phenomena involved in the transport of a pollutant in a river or a channel with respect to the distance from the point of release.

In the remainder, a fuzzy rule-based methodology is described for studying dispersion phenomena of a non-reactive pollutant in a water course. A numerical example of this methodology is given, related to a particular case. The methodology consists in the construction of a rule system by means of a "calibration data set", that is a number of previously run computer experiments. The rules are then combined to generate approximate forecasts of the concentration for any values of the inputs.

The goal of this study was to demonstrate the feasibility of fuzzy metamodelling for fast prediction of river pollution starting from a very limited amount of information, easily available from the viewpoint of a decision maker. In particular, the study concerns the case of an accidental continuous release at the point $x = 0$, starting at time $t = 0$. As the case study is the environmental management of a well defined river (and also for sake of simplicity), in this demonstration the considered input variables were only abscissa, time and mean velocity. The dispersion coefficient was not explicitly taken as an input of the metamodel, because (at least theoretically) it completely depends (for a determined water course) on

the velocity. It is worth noting that, strictly speaking, this assertion may not be true for natural rivers, the morphological conditions of which can be highly variable (due for instance to seasonal effects) but can be retained, for the exemplary purposes of the study, for an artificial channel.

1.2 The fuzzy model

In a nutshell (cf. [Bárdossy & Duckstein 1995] for an exhaustive introduction), in a fuzzy rule-based model physical equations linking inputs $X_1, X_2, \ldots$ (also called "premises") to an output Z are replaced by propositions:

$$\text{If } X_1 \text{ is } S_{1,i} \text{ and } X_2 \text{ is } S_{2,j} \ \ldots \text{ then } Z \text{ is } S_{z,k}, \tag{V.2}$$

in which $S_{1,i}, S_{2,j} \ldots$ and $S_{z,k}$ are fuzzy numbers defined in the input and the output space respectively. A fuzzy number (a particular case of fuzzy set) is defined by a support S and a "membership function" $m_S(x)$ which maps $x \in S$ to $[0,1]$, with the additional assumptions of "normality" (i.e. $\exists x \in S;\ m_S(x) = 1$) and quasi-concavity. A real value x of the variable X (also qualified as a "crisp" value) belongs to different fuzzy input sets (with different "degrees of fulfilment", i.e. the values of the membership functions for $X = x$) and thus activates, or "fires", a certain number of fuzzy rules. The fuzzy results (right-hand side of rules) of each activated rule are first composed into a single fuzzy set, then "defuzzified" to get, as the output of the metamodel, a real value z.

An interesting property of the fuzzy metamodels is that they are "universal approximators", i.e. under some additional assumptions of the function $G(X)$ (in particular, the continuity over a product of intervals) there exists a rule system able to approximate the actual function with a fixed precision. That is, for any $\varepsilon > 0$, for any method of combination of rules and defuzzification, there exists a rule system leading to a metamodel $\tilde{G}(X)$ such as $\mid \tilde{G}(x) - G(x) \mid < \varepsilon, \forall x$.

In the remainder a particularly intuitive way for building a metamodel from a set of numerical (or physical) experiments is presented and applied.

1.2.1 Learning

The learning phase consists in submitting to a rule system a given data set which simulates the process under known conditions. The better these data represent the phenomenon, the more precisely the fuzzy model is able provide realistic responses. First of all, a two input variables (x and t) model has been developed. Both variables have been normalized in the $[0,1]$ interval and later on fuzzified, by dividing the interval $[0,1]$ into 20 parts. Learning has been implemented with the so called "*counting algorithm*" [Bárdossy & Duckstein 1995] that operates in two different phases. First, the couples input-output are transformed into logical propositions (If ... Then ...). In general a single input (x, t) belongs to a certain number of fuzzy sets in the input space. So, a number of different propositions are obtained in which the left-hand side is fuzzy and right-hand side is crisp. Each proposition has a given degree of fulfilment (DOF), intuitively defined as the fulfilment of its conditions, or antecedents and calculated as the product of the membership functions of variables x and t. A filter provides for eliminating propositions which present a DOF smaller then a given value in order to improve precision and speed of calculations.

In the second step of learning the output variable C is fuzzified and the hybrid propositions, previously obtained, are transformed into propositions in which both left-hand side and right-hand side are fuzzy, that is, the rule system. The methodology consists in combining all propositions which have the same left-hand side and calculate minimum, maximum and mean of crisp right-hand sides. The corresponding fuzzy set (Figure V.2) is represented as a so-called triangular fuzzy number $\text{TFN}(C_{\min}, C_{\text{mean}}, C_{\max})$. Subsequently, using the same principles, a three-premise model has

been developed in which also the input variable v_x is fuzzified. The number of triangular membership functions for the variable v_x has been selected as 11, so that the normalized universe of v_x is divided into 10 equal parts.

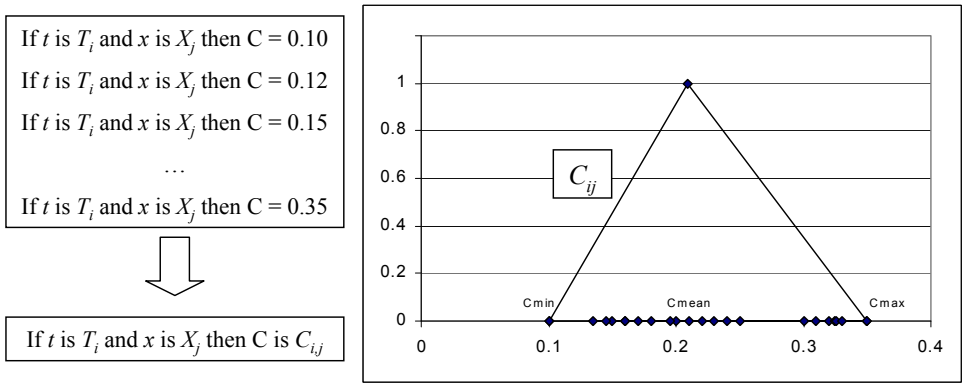


Figure V.2 – Schematic representation of the way rules are built according to the simple "counting algorithm".

1.2.2 Fuzzy rule-based computing

Once the learning phase is achieved and the rule system has been generated, the fuzzy algorithm allows, for a given vector (x, t), or (x, v_x, t) in the three-premise model, to calculate the corresponding value of the output variable C. Inputs fuzzification is made by singling out fuzzy sets to which x and t belong and their DOF. Then, for each rule whose premise is fired, DOF is calculated as product of membership functions of input variables. For the defuzzification a particular method has been chosen in order to take in account a peculiarity of the fuzzy responses. As it is clear from the description of the learning step, fuzzy responses are very different from one another. In fact, the bases, or supports, of the triangular membership functions of the output depend on the dispersion of the model response when the left-hand side of each rule is fixed. So, some of them are practically singletons, i.e. $C_{\min} \approx C_{\text{mean}} \approx C_{\max}$, while other ones have a very large base. The use of the classical centroid methods has not been able to give satisfactory results and little changes have been made to weights of the responses provided by each of the fired rules.

So, the fuzzy inference system is of the following form:

$$\tilde{G}(x, v_x, t) = \frac{\sum_{j \in \mathcal{J}(x,v_x,t)} \text{COG}_j \, \text{DOF}_j \, (1/A_j)^{\omega}}{\sum_{j \in \mathcal{J}(x,v_x,t)} \text{DOF}_j \, (1/A_j)^{\omega}} \tag{V.3}$$

in which: (i) $\tilde{G}(x, v_x, t)$ is the output of the metamodel, i.e. the predicted concentration, given (x, v_x, t), (ii) $\mathcal{J}(x, v_x, t)$ is the set of the indexes of the rules fired by the input vector (x, v_x, t), (iii) COG_j is the "center of gravity" of the fuzzy response of the rule j, (iv) A_j is the area subtended by the triangular membership function corresponding to the right-hand side of the rule j (if the triangular membership function degenerates in a singleton, then an arbitrary value A_s is assigned) and (iv) ω is a calibration coefficient.

Thus, each fuzzy response is weighted by a coefficient which increases with the DOF of the rule and decreases with the area of the fuzzy number in the right-hand side of the rule. For the choice of the parameters A_s and ω, several trials have been made in order to obtain the best fit between fuzzy results and validation data. In particular it has been empirically found that A_s has to be, obviously, small (basically smaller than any other value of A_i) ... but not too much. In fact the use of value of A_s which are several orders of magnitude smaller then the other terms in Equation V.3 risks to give too much importance to responses which are generated by only one point of the calibration test and which are not representative

enough of the phenomenon. The introduction of the term ω has improved the results. After a number of trials, it has been empirically found that the best results can be obtained for values between 0.4 and 0.6.

1.2.3 Numerical example

The (extremely simplified) exemplary case study is related to a 1D flow in a 1000 m long channel during an observation time of 900 s (15 minutes). The value of the turbulent diffusion coefficient has been taken as constant and equal to 0.35 m^2/s. Using a specific CFD finite-difference code (based on an Explicit-Upwind scheme) concentration profiles, for different value of time and velocity, have been obtained. The values of velocity which have been taken into account belong to the interval $[0,2$ m/s$]$. Part of the data have been set aside for the validation phase. Concentration is provided in a non-dimensional form with values between 0 and 1. Both two-premise model and three-premise model have been tested. In the two-premise model all data refer to a particular value of mean velocity v_x. In particular the following values have been taken into account: 1.0 m/s, 1.5 m/s and 2.0 m/s. Different simulations, with different value of A_s and ω, have been run in order to improve fitting the actual model and the metamodel. Depending on the value of v_x the mean absolute prediction error is between $8 \cdot 10^{-3}$ and $9 \cdot 10^{-3}$. Some results are graphically shown in Figure V.3.

The same validation data set has been used for testing the three-premise model. For $A_s = 0.005$ and $\omega = 0.5$, the average error on the whole validation test is equal to $1.1 \cdot 10^{-2}$.

1.2.4 Discussion

Both 2 and 3 premise fuzzy rule-based models seem to reproduce the phenomenon with acceptably small errors. Errors are due to the high non linearity of the functions which have to be identified by the fuzzy model. In fact, it has to be noticed how fuzzy profiles tend to be smoother, softer, than numerical ones, which is typical of fuzzy models. However, in practical applications, these errors may not be relevant (and less important than ones which occur because of a poor estimation of the dispersion coefficient).

The fuzzy model in comparison to numerical models has the advantage that it is able to calculate directly the value of concentration for given inputs without calculating the value of the function in the whole space-time grid. So the evaluation of particular results of interest (as for instance the arrival time of the pollution at a given x) is easier and faster; that can be valuable in environmental control. Although the example refers to a particular situation, because of its generality, the same procedure can be applied to other cases of pollution of water course. The possibility to train the fuzzy model with a "calibration data set" including both experimental and simulated data has also been considered.

2 Propagation of "hybrid" uncertainties

2.1 Motivations

Quantitative uncertainty assessment in engineering is commonly probabilistic, i.e. uncertain quantities are assumed to be random variables, described by probability distributions. This is also the main viewpoint of this manuscript.

In practice, in engineering (and simply in common life) uncertainty can be related to the inherent randomness of a phenomenon or to the imprecise knowledge of quantities which are definitely not-random but conveniently represented by means of random variables.

According to an usual terminology, the first are referred as *aleatory* uncertainties, the latter as *epistemic* uncertainties. A particularly interesting case of epistemic uncertainty (largely discussed in Chapter VI) is represented by the the one

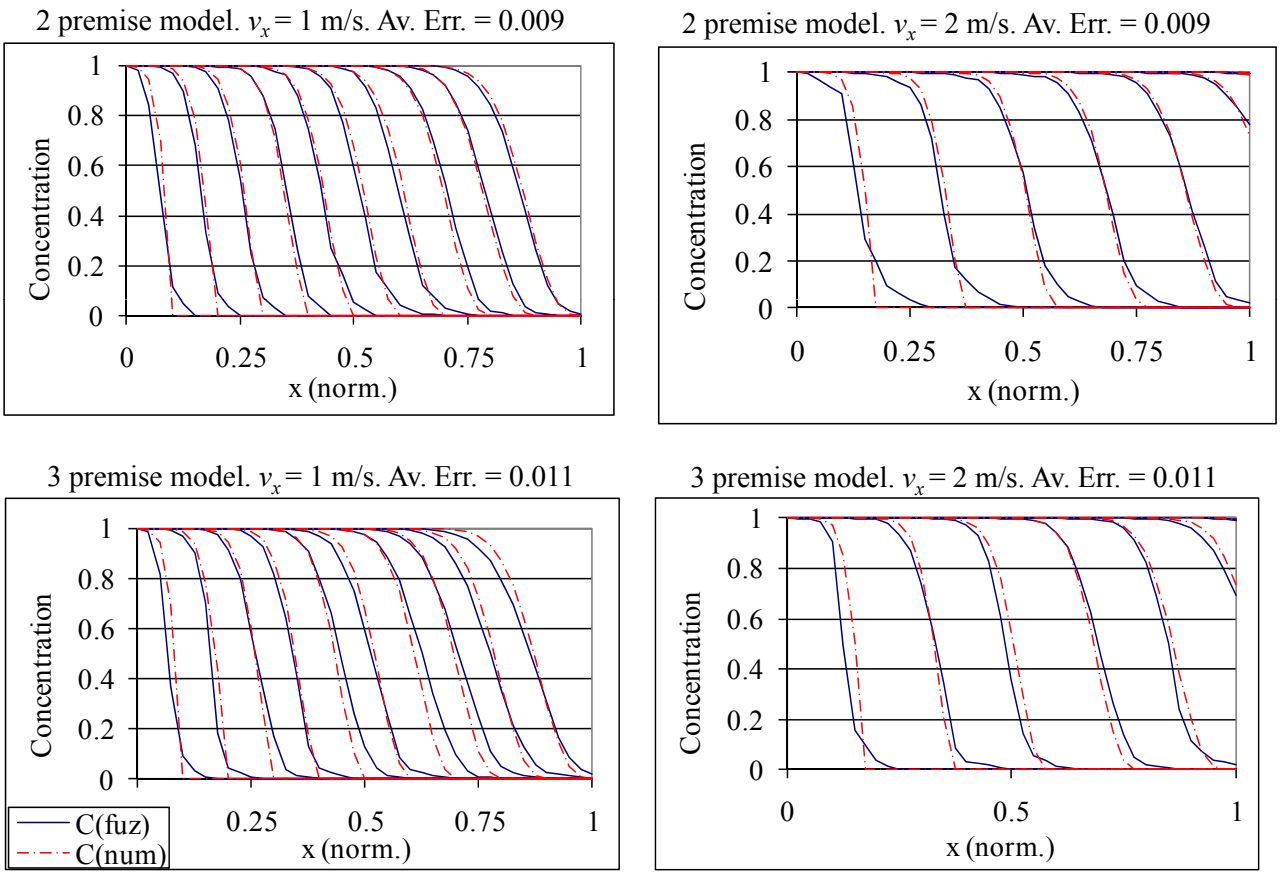


Figure V.3 – Comparison between the results of the fuzzy metamodel (blue solid line) and numerical finite difference code (red dotted line) for the two-premise model (inputs: distance from the source and time) and the three-premise model (inputs: distance from the source, time and mean velocity). The results are pollutant concentration longitudinal profiles (i.e. C vs. x) for different values of the time t. As one can see the arrival of the "pollutant front" is fairly well predicted by the metamodel.

tainting the estimation of the parameter $\boldsymbol{\theta}$ of a probabilistic model $p(\boldsymbol{X}|\boldsymbol{\theta})$ from an observed sample of the variable $\boldsymbol{X}$.

In spite of their very different nature, in common practice, all sources of uncertainty are assessed in a probabilistic framework: actually, this means that the analyst makes the implicit assumption of the existence of a joint probability distribution for the vector of all uncertain variables.

The justification of this fully probabilistic framework is definitely rooted in the rationale of Bayesian theory and namely in the "de Finetti's representation theorem", first proposed for binary sequences [de Finetti 1930a], then generalized by [Hewitt & Savage 1955] (cf. [Bernardo & Smith 1994] for a full introduction). Taking inspiration from the pedagogical presentation of [Bernardo 1996], roughly speaking, this theorem (in its parametric form) states that if one considers a sequence of *exchangeable* random quantities $(\boldsymbol{x}_1, ..., \boldsymbol{x}_n)$, i.e. such as the joint probability $p(\boldsymbol{x}_1, ..., \boldsymbol{x}_n)$ of the sequence does not depend on the order in which the $\boldsymbol{x}_i$'s have been observed, then there exists a parametric model $p(\cdot|\boldsymbol{\theta})$ and a (*prior*) distribution $\pi(\boldsymbol{\theta})$ such as:

$$p(\boldsymbol{x}_1, ..., \boldsymbol{x}_n) = \int \prod_{i=1}^{n} p(\boldsymbol{x}_i|\boldsymbol{\theta})\, \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta}.$$

As highlighted by [Bernardo 1996], "*if the observations are conditionally independent - as it is implicitly assumed*

when they are considered to be a random sample from some model - then they are necessarily exchangeable". Hence, exchangeability appears as a rather not-restrictive condition. The existence of $\pi(\boldsymbol{\theta})$ is an extremely powerful result. Besides all the important methodological and practical consequences, it can provide a justification to the probabilistic treatment of non-random albeit unknown quantities (the so called "*state of Nature*"), conditional on which random samples of a variable of interest (e.g. the output of a computer code) can be generated.

Prior distribution represents the initial knowledge of the analyst about $\boldsymbol{\theta}$ and it is updated by means of the Bayes formula. Probability is seen as a numerical quantification of a state of knowledge. This "translation" is not arbitrary but obeys some rationality principles. This "subjective" probability is associated to the idea of odd: the probability of an event depends on the amount that a rational individual is ready to bet on it[2].

Moreover, the full-probabilistic (Bayesian) setting proves adapted to the statistical practice in industry, business, biomedical, environmental applications, for both theoretical (it is rooted in a decision-making framework) and practical reasons (the prior distribution can be used to add expert knowledge to the statistical analysis).

We will come back to these points on the next chapter. Here, we simply highlight some arguments which are often raised against the probabilistic assessment of epistemic uncertainties and motivate the exploration of alternative settings, seemingly less informative than probability theory.

The exchangeability can be challenged in risk analysis, especially when dealing with extremely unlikely (and often also extremely costly) events: "*a probability model presumes some sort of model stability, populations of similar units need to be constructed (in the Bayesian context, formally an infinite set of exchangeable random variables). But this stability is often not fulfilled. [...] In a risk assessment context the situations are often unique and the establishment of chances means the construction of fictional populations of non-existing similar situations*" [Aven & Zio 2011]. Let us think, for instance, to the probability of a terrorist attack. In this case, "*one should need to define a large set of identical (exchangeable) attack situations, where some aspects (for example related to the potential attackers and the political context) are fixed and others (for example the attackers motivation) are subject to variation*" [Aven & Steen 2010].

In other terms (but definitely raising the same issue) other authors challenge the relevance of probabilistic assessments in presence of very poor knowledge and/or scarce data [Dubois 2006], [Baudrit *et al.* 2008], [Roy & Oberkampf 2011].

Another interesting point, raised, for instance, by [Regan *et al.* 2004], [Baudrit *et al.* 2008] or [Helton *et al.* 2011, Helton & Sallaberry 2012], is that epistemic and aleatory uncertainties must be differently treated in engineering studies: the common probabilistic practice of the *nested Monte Carlo* propagation, in practice the use of the predictive distribution of the quantity of interest as a summary of the overall uncertainty (cf. Section 2.2 of the Chapter VI), must be avoided. Although we fully agree on this point, we think that Bayesian theory can bring an appropriate answer to this question, which can properly fit theoretical and practical requirements of a risk analysis.

However, without taking any dogmatic position, we think that investigating of the use of alternative approaches (unfortunately, often introduced "in opposition" to probabilistic methods) for uncertainty quantification and propagation is an interesting field of research in engineering. These methods are appealing as the restitution of results under the form of "simple" bounds (not associated to a probability) is seemingly easier to understand and interpret for the practitioner. In addition, institutions concerned with regulation issues in different business areas seem to be more and more interested in non-probabilistic methods.

The works sketched in the remainder have been mostly carried within the framework of a three-year partnership with the Politecnico di Milano. The methods have been explored in this exploratory study from the practitioner's viewpoint.

[2] "*Let us suppose that an individual is obliged to evaluate the rate p at which he would be ready to exchange the possession of an arbitrary sum S (positive or negative) dependent on the occurrence of a given event E, for the possession of the sum pS; we will say by definition that this number p is the measure of the degree of probability attributed by the individual considered to the event E, or, more simply, that p is the probability of E*" [de Finetti 1930b], translated into English in [Kyburg & Smokler 1980].

Therefore, emphasis has been more placed on the understanding of the algorithmic aspects rather than on the underlying mathematical foundations (which are definitely quite complex).

2.2 Hybrid possibilistic-probabilistic framework

As usual, let us consider a system, the behaviour of which is described by a deterministic black-box function $Z = G(X_1, ..., X_n)$, mapping n input variables to $\mathbb{R}$.

The inputs are ordered in such a way that the first k form a random vector, represented by a probability distribution $p(\cdot)$ and the last $n-k$ are represented by "*possibility distributions*": $\varphi_{k+1}(\cdot), ..., \varphi_n(\cdot)$.[3]

The possibility distribution assigns to each value (let us say x) of a variable $X \in \mathcal{A} \subseteq \mathbb{R}$ a *degree of possibility* $\varphi(x) \in [0,1]$, such that $\varphi(x) = 0$ means that x is an impossible value for X, whereas $\varphi(x) = 1$ means that $\{X = x\}$ "*is just unsurprising, normal, usual, a much weaker statement than when probability is 1*" [Dubois 2006]. The possibility distribution is "*normalized*" in the sense that $\exists\, x, \varphi(x) = 1$.

Possibility distributions are linked to fuzzy intervals: under the (mild) condition of quasi-concavity, a possibility distribution is the membership function of an interval of $\mathcal{A}$. For sake of simplicity, we will admit this condition fulfilled, in the remainder. Another fashion to interpret this, is that a quasi-concave possibility distribution defines a set of nested interval with various credibility levels $\alpha \in [0,1]$.

From the possibility distribution $\varphi(\cdot)$, one defines for any subset $A \subseteq \mathcal{A}$ the so-called *possibility* and *necessity measures*, noted $\Pi(A)$ and $\mathrm{N}(A)$ respectively:

$$\begin{cases} \Pi(A) = \sup\limits_{x \in A} \varphi(x) \\ \\ \mathrm{N}(A) = 1 - \Pi(A^c) = \inf\limits_{x \notin A} \varphi(x). \end{cases} \tag{V.4}$$

The link between possibility and probability distributions is made by considering that from the pair $[\mathrm{N}(\cdot), \Pi(\cdot)]$ one can defines a family of probability distributions $\mathcal{P}_\varphi$, completely determined by the function $\varphi(\cdot)$:

$$\mathcal{P}_\varphi = \{p \text{ such as } \forall A, \mathrm{N}(A) \leq p(A)\} = \{p \text{ such as } \forall A, p(A) \leq \Pi(A)\},$$

such that:

$$\begin{cases} \sup\limits_{p \in \mathcal{P}_\varphi} p(A) = \Pi(A) \\ \\ \inf\limits_{p \in \mathcal{P}_\varphi} p(A) = \mathrm{N}(A), \end{cases} \tag{V.5}$$

which means that $\mathrm{N}(A)$ and $\Pi(A)$ are lower and upper bounds, respectively, for the probability $p(A)$. For further details concerning the link between possibility and probability theories, one can refer to [Baudrit 2005], [Dubois 2006] or [Dubois & Prade 2011].

The propagation of hybrid possibilistic-probabilistic uncertainties can be made by means of the algorithm described hereinafter, proposed by [Baudrit 2005, Baudrit *et al.* 2006], based on two steps: Monte Carlo simulation for propagating

[3] Notice that the notation usually dedicated to possibility distribution is $\pi(\cdot)$. Here, we preferred the symbol $\varphi(\cdot)$ to avoid any confusion with Bayesian prior and posterior distribution, noted $\pi(\cdot)$ and $\pi(\cdot|\cdot)$ in this manuscript. Nevertheless, according to usual notations, we keep noting $\Pi(\cdot)$ the possibility measure (cf. Equation V.4).

random (probabilistic) inputs and Fuzzy Interval Analysis for propagating possibilistic inputs.

More precisely, one first generates a m-random sample of the random inputs $\{x_1^{(i)},...,x_k^{(i)}\}$, with $i=1,...,m$. Then, the interval $[0,1]$ is subdivided in m_α intervals of length $\Delta\alpha$. For $\alpha=0$, $\Delta\alpha$, $2\Delta\alpha$ etc., the so-called α-cuts of the possibility distributions $\varphi_{k+1}(\cdot),...,\varphi_n(\cdot)$ are to be found, i.e. the $n-k$ sets $A_j^\alpha=\{x_j,\varphi_j(x_j)=\alpha\}$, with $j=k+1,...,n$ (we note "$j>k$" this collection of indexes in the remainder).

Then, for each $i=1,...,m$ the following procedure is applied:

(a) First, set $\alpha=0$.

(b) If $\alpha\leq 1$: find the α-cut intervals of Z: $[z_{\min}^{(i,\alpha)},z_{\max}^{(i,\alpha)}]$ with:

$$z_{\min}^{(i,\alpha)}=\min_{\{x_j\in A_j^\alpha\}_{j>k}} G\left(x_1^{(i)},...,x_k^{(i)},x_{k+1},...,x_n\right) \quad \text{and} \quad z_{\max}^{(i,\alpha)}=\max_{\{x_j\in A_j^\alpha\}_{j>k}} G\left(x_1^{(i)},...,x_k^{(i)},x_{k+1},...,x_n\right) \tag{V.6}$$

(c) if $\alpha\geq 1$: stop, otherwise set $\alpha=\alpha+\Delta\alpha$ and go to step (b).

Thus, in the end, one has m random realizations of m_α α-cuts, i.e. for any random sample $\{x_1^{(i)},...,x_k^{(i)}\}$, m_α (nested) intervals of values of the output Z, corresponding to the credibility levels $\Delta\alpha, 2\Delta\alpha$.... These intervals defines m possibility distributions $\varphi^{(i)}(z)$ of the output Z.

Let us now consider sets of possible values for Z, in particular intervals $]-\infty,z^\star]$. From the m possibility distributions built according to the procedure described hereinbefore, one can build m necessity and possibility measures (Equation V.4):

$$\begin{cases} \Pi^{(i)}(]-\infty,z^\star]) = \sup_{z\in]-\infty,z^\star]} \varphi^{(i)}(z) \\ \\ \mathrm{N}^{(i)}(]-\infty,z^\star]) = \inf_{z\notin]-\infty,z^\star]} \varphi^{(i)}(z). \end{cases}$$

That allows proposing m bounds for the probability distribution of Z, each pair of bounds depending on the value of the "probabilistic" inputs.

Let us consider the following way for combining these pairs of necessity-possibility measures, that is evaluating the averages:

$$\begin{cases} \mathrm{Bel}(]-\infty,z^\star]) = \sum_{i=1}^m \frac{1}{m}\mathrm{N}^{(i)}(]-\infty,z^\star]) \\ \\ \mathrm{Pl}(]-\infty,z^\star]) = \sum_{i=1}^m \frac{1}{m}\Pi^{(i)}(]-\infty,z^\star]). \end{cases} \tag{V.7}$$

It can be shown (cf. [Baudrit 2005, Baudrit *et al.* 2006, Baudrit *et al.* 2008]) that these quantities define lower and upper bounds for the probability $\mathbb{P}[Z\leq z^\star]$ (and thus for the cumulative distribution function of Z):

$$\mathrm{Bel}(]-\infty,z^\star]) \leq \mathbb{P}[Z\leq z^\star] \leq \mathrm{Pl}(]-\infty,z^\star]),$$

and, namely, they can be interpreted as the "*Belief*" and the "*Plausibility*" functions in the sense of the so-called "*Dempster-Shafer Theory*" (cf. Appendix, page 157).

Example. We consider here an oversimplified case of flood risk analysis excerpted from [Baraldi *et al.* 2011]. One wants to evaluate the probability that a water level in a given section of a water course exceed a fixed threshold. The

quantity of interest (here, the water level, noted Z_c) depends on some uncertain inputs, by means of the following analytic formula:

$$Z_c = Z_v + \left\{ Q / \left(BK_s \sqrt{(Z_m - Z_v)/L} \right) \right\}^{0.6}, \tag{V.8}$$

where:

- Z_m, Z_v are the riverbed levels (in m a.s.l.) in the upstream and downstream section respectively of the river portion under investigation;
- Q is the yearly maximal water discharge (in m^3/s);
- B are L the width and the length respectively of the river section (in m);
- K_s is the Strickler friction coefficient.

This exemplary case is also used in the next chapter (cf. page 112) with different data. It has also been used (possibly with slight modifications) for exemplary purposes in several recent papers (e.g. [Limbourg & de Rocquigny 2010, Munoz-Zuniga *et al.* 2012, Ko & Garnier 2013])

We consider B and L as constant parameters (equal to 300 and 5000 m respectively), while Q and (Z_m, Z_v) are random variables, modelled by a Gumbel and a bi-variate Gaussian distribution respectively. Actually, Q and (Z_m, Z_v) are supposed to be tainted with aleatory uncertainties, in the sense they are random in themselves.

As far as Strickler friction coefficient is concerned, it is not known with certainty by the analyst, but it can be imagined that it is not random in "itself" and its uncertainty is rather related to a lack of knowledge (of course, this interpretation can be challenged in some practical situations [Fu 2012], in which K_s can be supposed to be actually random in the common sense of the term). It is supposed that the analyst has a vague initial knowledge, under the form of a reference value ($K_s = 30$) and (large) bounds $[5, 60]$ outside which it is extremely unlikely to find any value of K_s.

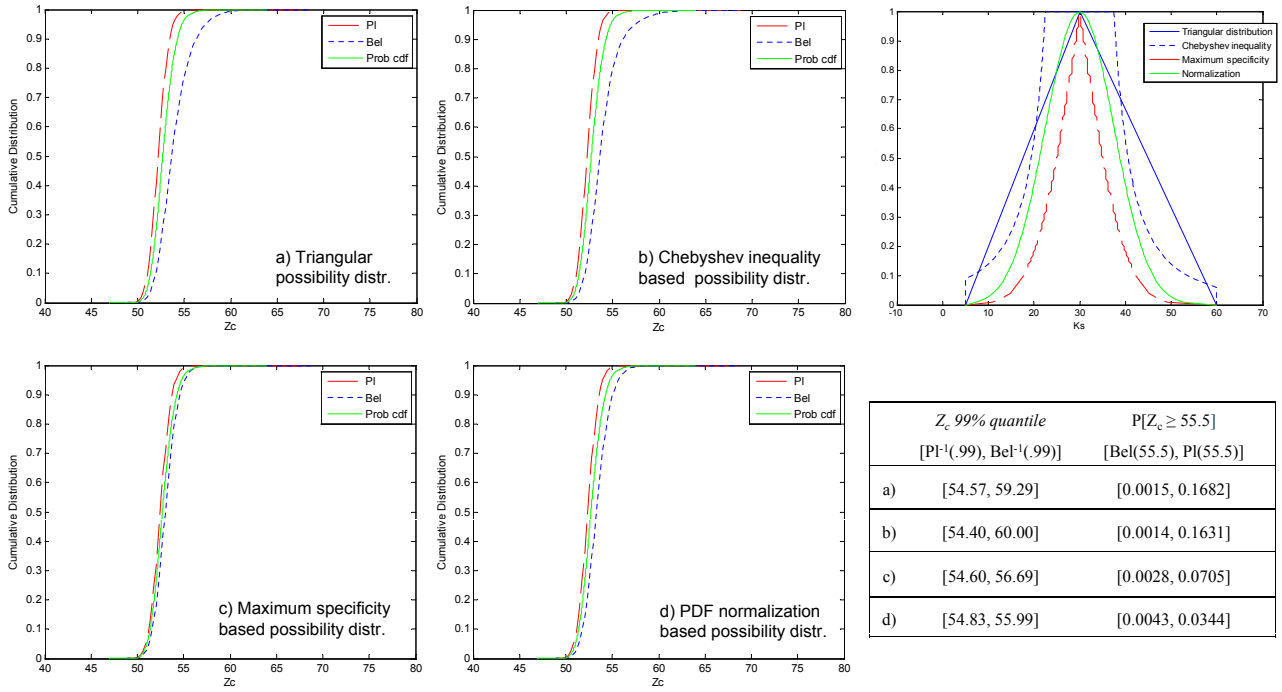
Four different elicitation strategies have been performed, leading to the possibility distributions $\varphi(K_s)$ shown in the upper right corner of Figure V.4. For more details about these techniques, and particularly the ones based on probabilistic-possibilistic transformation or probabilistic inegalities (Chebyschev and Camp-Meidel), cf. [Dubois *et al.* 2004], as well as [Baudrit 2005] and references therein.

Some results are shown in Figure V.4: bounds for the cumulative distribution function of the variable of interest Z_c (namely, Belief and Plausibility functions, cf. Equation V.7) as well as intervals for the 99% quantile and the flooding probability $\mathbb{P}[Z_c \geq 55.5 \text{ m}]$, i.e. the probability for the water level to exceed a fixed threshold (the height of an existing dike). The cumulative distribution function of Z_c (green curve between the red and the blue one) is obtained as the *predictive* distribution (cf. page 108) of Z_c ,under the assumption that the Strickler's coefficient is normally distributed with mean and standard deviation equal to 30 and 7.5 respectively.

2.3 Hybrid possibilistic-probabilistic "hierarchical" framework

The hybrid framework sketched hereinbefore can be "extended" by considering that purely epistemic uncertainties taint also the parameters of the probability distribution assigned to random variables $X_1, ..., X_k$ [Baudrit *et al.* 2008]. Thus, in this framework, the variables $X_{k+1}, ..., X_n$ are possibilistic, while parametric probability distributions are given to $X_1, ..., X_k$:

$$X_j \sim p(\cdot|\boldsymbol{\theta}_j) \quad \text{with} \quad j = 1, ..., k,$$



	Z_c 99% quantile [$Pl^{-1}(.99)$, $Bel^{-1}(.99)$]	$P[Z_c \geq 55.5]$ [Bel(55.5), Pl(55.5)]
a)	[54.57, 59.29]	[0.0015, 0.1682]
b)	[54.40, 60.00]	[0.0014, 0.1631]
c)	[54.60, 56.69]	[0.0028, 0.0705]
d)	[54.83, 55.99]	[0.0043, 0.0344]

Figure V.4 – Some examples of results of hybrid uncertainty propagation excerpted from [Baraldi *et al.* 2011]. Left and center graphs (numbered a, b, c, d): bounds for the probability distribution of Z_c according to different possibility distributions of K_s (shown in the upper-right graph) built from the same preliminary information. The values shown in the lower-right table are directly obtained by the left and center graphs (cf. [Baudrit 2005] for the interpretation of this method).

the parameters of which, θ_j's, are uncertain and modelled by possibility distributions $\varphi_j(\cdot)$. The propagation algorithm is very similar to the one depicted in the previous section (cf. page 99). The only difference is that random samplings are not made independently on possibilistic variables. For instance, let us focus on the random sampling of the variable X_j. For simplifying the presentation, let us consider the case where θ_j is scalar (noted θ_j). In practice, first α-cuts A_j^α are determined for θ_j; then the random interval of X_j corresponding to A_j^α and to the sample number i is bounded by:

$$\inf_{\theta_j \in A_j^\alpha} F^{-1}(u_j^{(i)}|\theta_j) \quad \text{and} \quad \sup_{\theta_j \in A_j^\alpha} F^{-1}(u_j^{(i)}|\theta_j),$$

in which $u_j^{(i)}$ is randomly sampled from the uniform distribution over $[0,1]$ and $F^{-1}(\cdot|\theta_j)$ is the inverse of the cumulative distribution function of X_j.

In the end, this algorithm produces, as the one described in the previous section, a number of random realizations of α-cuts of the output variables, which are interpreted as random realizations of the possibility distribution of the output and processed by formulas in Equation V.7 for obtaining probability bounds.

Example. The case study sketched hereby, fully described in [Pedroni *et al.* 2012] and [Pedroni *et al.* 2013a], is concerned with the same problem of flood risk assessment, sketched in the previous Section 2.2. The variable Q, K_s, Z_m and Z_v are supposed random and distributed according to Gumbel (Q) and Gaussian densities. In Table V.1 some details are given on the way possibility distributions on the parameter of such densities have been built from the available knowledge. The situation is, of course, exemplary and the example provided below has more methodological than practical relevance.

Variable and probability distribution	Available information and elicitation technique used
Discharge: $Q \sim \text{Gu}(\eta, \beta)$ (Gumbel)	Likely values for the mean and the standard dev. of η and β from previous studies are available [Pasanisi *et al.* 2009a, Limbourg & de Rocquigny 2010]. "Normalization" of the (truncated) corresponding Gaussian distributions, i.e. dividing the expression of the density by its max, such as: $\exists \eta, \beta, \varphi(\eta) = 1, \varphi(\beta) = 1$.
Riverbed up- and down-stream levels: $Z_m \sim \text{Norm}(\mu_{Z_m}, \sigma_{Z_m})$ $Z_v \sim \text{Norm}(\mu_{Z_v}, \sigma_{Z_v})$	Likely values for the mean and the standard dev. of $(\mu_{Z_m}, \sigma_{Z_m})$ and $(\mu_{Z_v}, \sigma_{Z_v})$ are available (cf. same references as previous line). They are used as inputs (μ, σ) for the procedure based on the Chebyshev probabilistic inequality.
Strickler's coefficient: $K_s \sim \text{Norm}(\mu_{K_s}, \sigma_{K_s})$	Physical bounds are available, as well as a five values of K_s, obtained by calibration of the code $G(\cdot)$ with an estimated error of $\pm 15\%$. For sake of simplicity, no possibility distribution is put over σ_{K_s}. Concerning μ_{K_s}, a trapezoidal distribution is chosen (cf. [Pedroni *et al.* 2012, Pedroni *et al.* 2013a] for further details).

Tableau V.1 – Possibilistic modelling of the parameters of the probability distributions of the uncertain parameters in a simplified flood risk assessment example.

It is worth noting that the proposed hybrid propagation algorithm, because α-cuts are determined for all possibilistic values at the same time for the same values of α (i.e. $\Delta\alpha$, $2\Delta\alpha$, $3\Delta\alpha$ etc.), artificially induces a kind of dependence between those variables. To overcome this issue, another propagation algorithm has been tested, entirely based on the formalism of Dempster-Shafer Theory. Actually (cf. Appendix, page 158), by a simple discretization, a possibility distribution can generate a set of intervals to which different credibility are given, defining thus a "Belief function" for that variable. The advantage of this algorithm (that we will not discuss here) is that random intervals are independently sampled. The same remark can be done for the algorithm proposed in Section 2.2.

However, as shown in Figure V.5 in this particular case, the results are very similar, in terms of Belief and Plausibility of the output Z_c.

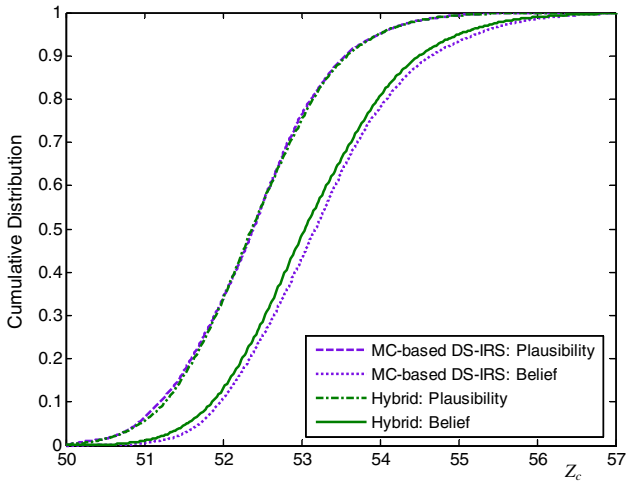


Figure V.5 – Comparison of the results given by hybrid propagation method and Dempster-Shafer Monte Carlo propagation (assuming independence) for the exemplary case presented in Section 2.3 (here the acronym "MC based DS-IRS" stands for "*Monte Carlo based Dempster-Shafer with Independent Random Sets*", cf. [Baudrit 2005, Helton *et al.* 2006a, Limbourg 2007] for further details). The Belief and Plausibility functions of the output Z_c obtained by means of the two methods are very close to one another.

2.4 Further works and comments

Further works made in this field, in cooperation with the colleagues of Politecnico di Milano and the Chair on *Systems Science and the Energy Challenge* (École Centrale Paris / SUPÉLEC) concerned the important issue in engineering of updating preliminary knowledge when new pieces of information are available.

In [Pedroni *et al.* 2013b, Pedroni *et al.* 2014], examples of application of a "possibilistic version" of the Bayes formula are shown. Let us consider again the case of Section 2.3 of uncertainties described by probability distribution the parameters of which are tainted with epistemic uncertainty, modelled by means of possibility distributions $\varphi(\cdot)$. According to [Lapointe & Bobée 2000], when some data $\mathcal{D}$ are observed, the updated possibility distribution may be written:

$$\varphi(\theta|\mathcal{D}) = \frac{\varphi_{\mathcal{L}}(\mathcal{D}|\theta)\cdot\varphi(\theta)}{\sup_{\theta}\{\varphi_{\mathcal{L}}(\mathcal{D}|\theta)\cdot\varphi(\theta)\}}, \tag{V.9}$$

in which $\varphi_{\mathcal{L}}(\mathcal{D}|\theta)$ is the "possibilistic likelihood", obtained by transforming the usual (probabilistic) likelihood $\mathcal{L}(\mathcal{D}|\theta)$ through the following "possibilistic normalization":

$$\varphi_{\mathcal{L}}(\mathcal{D}|\theta) = \frac{\mathcal{L}(\mathcal{D}|\theta)}{\sup_{\theta}\mathcal{L}(\mathcal{D}|\theta)}. \tag{V.10}$$

Notice that other forms of the possibilistic Bayes theorem, alternative to Equation V.9, can be defined as a result of other definitions of the operation of *conditioning* with possibility distributions (here, the product of the possibilistic prior and likelihood is used, but other alternative manners can be proposed). Cf. [Lapointe & Bobée 2000] for technical details, as well as [Arefi *et al.* 2010] for an example concerning lifetime data. The formula of Equation V.9 is often used because of its similarity with the classical Bayes formula and also because (together with other ones), "*it satisfies desirable properties of the revision process and lead to continuous posterior distributions*" [Lapointe & Bobée 2000].

Figure V.6 shows an example of results results excerpted by [Pedroni *et al.* 2013b], concerning the simplified flood risk assessment exercise sketched at page 100.

Some other refinements of this hybrid framework and recommendations are given in [Pedroni *et al.* 2014].

The studies and the considerations presented in Sections 2.1 and following do not pretend to be exhaustive. The problem is vast and many different theoretical and methodological tools are involved. The goal was rather to acquire a certain knowledge of the bases of this hybrid framework and apply these concepts to problems similar to the ones engineers can encounter in uncertainty and risk analysis.

As final remarks, it is worth reminding again the growing interest of the technical community for these extra-probabilistic frameworks of uncertainty analysis. Besides methodological considerations, the greatest practical difficulty in applying these methods is (possibly) the computational cost. Actually, the step (b) of the hybrid propagation algorithm (Equations V.6) requires solving two optimization problems that can be burdensome if the deterministic function $G(\cdot)$ is not monotonic. Hence, many runs of the computer code $G(\cdot)$ are necessary to solve these problems, which can be prohibitive if the code is CPU time consuming. In these cases, the only way for coping with this issue seems to be the use of metamodelling techniques [Helton *et al.* 2006a].

In spite of their intuitive and appealing first approach (intervals can be perceived as more easily understandable mathematical objects than probability distributions or confidence intervals), it must be said that the formalism (as well as the coexistence of multiple and strictly related extra-probabilistic frameworks) can definitely turn complicated for engineers which can have the feeling that different mathematical tools and algorithms are available for solving the same problem

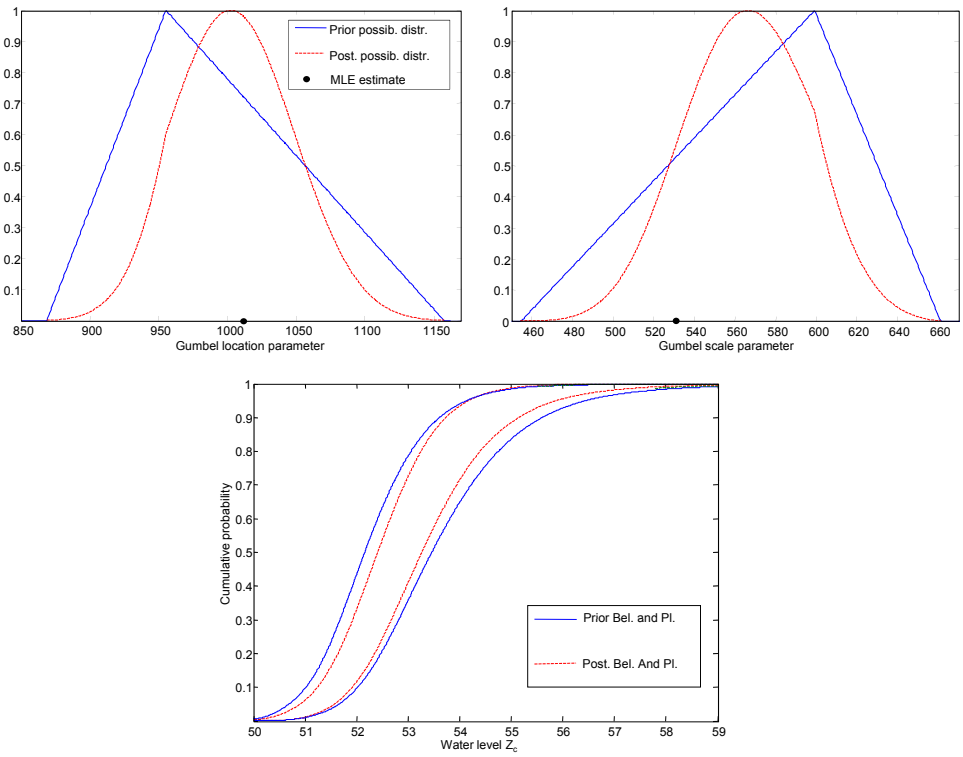


Figure V.6 – Updating and propagating possibility distributions. Upper part: prior and possibility possibility distributions of parameters of the Gumbel probability distribution of the discharge Q. Lower part: Belief and Plausibility functions of the water level Z_c obtained by propagating the prior and posterior possibility distributions of parameters through the deterministic model. Adding data, updating information, induces a reduction of the uncertainty tainting Z_c as the bounds for the CDF of Z_c get closer.

starting from the same hypotheses.

From the risk analyst's viewpoint, these methods can be attractive because of the reasons already evoked in Section 2.1: in presence of very scarce information the relevance of probabilistic uncertainty assessment can be challenged and the use of "less-informative" frameworks is imaginable.

An interesting perspective for future work is the definition of methodologies for the estimation of point risk indicators (e.g. a probability of failure). That can be done quite easily, at least from a conceptual viewpoint, in the fully-probabilistic framework (as it is largely discussed in Chapter VI). Actually, in some studies, it should be interesting to have at disposal a formal theory for choosing a point value of the quantity of interest to be estimated, especially when the interval resulting from the "hybrid" methodology shown hereinbefore is very large. Depending on the context of the study, the counterpart of the "conservatism" of the intervals produced by extra-probabilistic methods is the difficulty in making decision under their bases. Cf. [Le Duy *et al.* 2013] for an example of solution of this issue in the context of nuclear probabilistic risk assessment.

Finally, it is worth noting that in most cases one is also interesting in assessing a sensitivity analysis of the system's model output and/or risk indicators with respect to the input of the system's model. The state-of-the art methods of sensitivity analysis (essentially based on regression models or variance decomposition) rely on probabilistic assumptions and, to the best of our knowledge, whether some preliminary works have been concerned with this specific problem outside the probabilistic setting, e.g. [Ferson & Tucker 2006, Datta 2011, Chutia *et al.* 2013], much has to be done for establishing a consolidated and mature framework to be used by engineers.

Chapter VI

Bayesian point estimation of probabilistic risk criteria

On the flight over here, stuffed into a narrow seat, I had done some quick calculations on the back of a cocktail napkin. At the time I had just over five thousand bucks. At the current, legal rate of exchange, it would net me just over four thousand euros. If I lived very carefully, I estimated I could eke out three or four months [...] But forty-eight hours after landing in Paris, I had already spent over four hundred dollars.

Douglas Kennedy, The Woman in the Fifth (2007)

Reading notes

Technical context. The works presented in this chapter (as well as the ones presented in Chapters IV and V) are also concerned with uncertainties in computer simulation. The point here is the estimation of the so-called "quantities of interest" in uncertainty analysis studies (in the sense of the scheme of Figure IV.2) when it is reasonable to admit that "epistemic" uncertainties, due to lack of knowledge, taint the probabilistic model of the input variables.

This kind of problems is frequent in engineering and R&D studies and the question is of particular interest in risk and safety analyses. Actually, regulation bodies may question operators on the way they account for epistemic uncertainties, under a more or less prescriptive viewpoint (depending on the study and its context), and how they provide "conservative" values of safety criteria taking into account this lack of knowledge.

The works sketched hereinafter investigate this problem in the light of the Bayesian decision theory, which actually provides a mathematically consolidated framework for point estimation that can conveniently answer the question posed. This framework allows also to reinterpret the common heuristic consisting in using the predictive distribution, supposed to account "for all sources of uncertainties."

Contributions. This problem is fundamentally rooted in the rationale of Bayesian theory and personally the work carried allowed me to consolidate the basis of my practice of Bayesian analysis.

Some contributions are methodological. The book chapter [Parent *et al.* 2014] introduces the methodological decision problem of providing a design value for the height of a flooding protection dike, from a strictly statistical viewpoint. Cf. also [Keller *et al.* 2010a] for statistical considerations specifically focused on quantities of interest in lifetime analysis using the Weibull distribution.

The articles [Pasanisi *et al.* 2012c] and [Keller *et al.* 2011c] (published in *Reliability Engineering and System Safety*

and the *Journal de la Société Française de Statistique*, respectively) are more specifically focused on the case of computer experiments: the main result is that the common heuristics based on the predictive distribution actually lead to Bayesian point estimators, the underlying cost function of which is completely out of the analyst's control and defined by the nature of the quantity to be estimated (e.g. probability of failure, quantiles). The framework has also been the subject of several oral presentations: [Keller *et al.* 2011b], [Keller *et al.* 2011a], [Keller *et al.* 2011d], [Pasanisi *et al.* 2012b], [Pasanisi 2012b].

These considerations have been also applied to a particular "functional" risk criterion, namely the fragility curve, i.e. the function linking the probability of failure of a system or a structure to the intensity of the load it is submitted to (e.g. a seismic acceleration). This technical problem, of particular interest in the domain of nuclear safety, is described in an article [Damblin *et al.* 2014c] of the *Journal de la Société Française de Statistique*, and in a couple of presentations in national [Damblin *et al.* 2012a] and international congresses [Damblin *et al.* 2012b].

The study on seismic risk assessment has been completed by the investigation of the probability distribution function (pdf) of the seismic acceleration. The estimation of the underlying probabilistic model (following the "*peak over threshold*" formalism) is not trivial as historically observed magnitudes are usually not i.i.d. A full Bayesian methodology has been proposed and presented in a communication [Pasanisi *et al.* 2013b] and in an article [Keller *et al.* 2014b], currently under revision. The communications [Pasanisi 2014, Keller *et al.* 2014a] attempt to link the Bayesian frameworks for both fragility and seismic hazard in an unified decisional setting.

Finally, further studies have been concerned with the Bayesian estimation of the the so-called "*species sensitivity distribution (SSD)*" with the purpose to forecast the no-effect concentration of a chemical substance in a natural environment[1]. These works, which have not gone up to the definition of point estimators, are discussed in [Ciffroy *et al.* 2012] and [Ciffroy *et al.* 2013][2] .

Structure of the chapter and credits. Most of the text of this chapter (Section 1 to 4) is excerpted from the article [Pasanisi *et al.* 2012c], expressly intended to the engineers computer experiments community. Methodological considerations developed in [Keller *et al.* 2011c] are presented in Section 5 as an extension of the previously presented work; this may be surprising if one considers the years of publication of the articles, but the review process of [Pasanisi *et al.* 2012c] (first submitted in 2010 in a quite different form [Keller *et al.* 2010b] and finally published in 2012) was particularly long, so that this article has been finally published few months after its "natural" extension [Keller *et al.* 2011c].

Section 5 is adapted from [Keller *et al.* 2011c].

Section 6 is adapted from [Damblin *et al.* 2014c] (Sections 6.1 and 6.2) and [Keller *et al.* 2014b] (Section 6.3).

1 Coming back to the "quantity of interest"

According to the usual methodological scheme sketched in Chapter IV (cf. Figure IV.2), uncertainty analysis consists in assessing the pdf of the output Z by transferring the uncertainty from inputs $\boldsymbol{X}$ to Z through the function $G(\cdot)$ (notice that here the explicit dependence on fixed parameters is omitted for simplifying notations). In industrial practice, one is more often interested in a particular feature of the distribution of the uncertain variable Z: the so-called *quantity of interest* of the study. It can be a central value (mean, median), a dispersion parameter (standard deviation, coefficient of variation,

[1] SSD's are defined as "*statistical distributions describing the variation among a set of species in toxicity of a certain compound or mixture*" [Posthuma *et al.* 2001]. The cumulative distribution function of the SSD returns for each value of the concentration the potential affected fraction of natural species suffering from significant toxicity effects. According to a standard practice, the SSD is commonly assumed as log-normal.

[2] Article published in the journal *Environmental Toxicology and Chemistry*.

which is typical in measurement science problems) or, more often, a quantile corresponding to a low or high probability as well as the probability for Z to be greater than a given threshold value $z^\star$: $\mathbb{P}[Z > z^\star]$. This is typically the case in reliability analysis: the failure of the system is associated to the fact that some state variables take values outside a safety domain. The quantity of interest is here interpreted as a *probability of failure*. According to this scheme, the quantity of interest, i.e. the final result of the analysis, denoted ϕ in what follows, summarizes in itself the stakes that motivate the study. Indeed, it is very often the case in the industrial practice that the engineer must provide a result complying with a (more or less formalized) normative question: "what is the probability for the temperature to be greater than a given (admissible) value?", "what is the water level corresponding to an exceedance probability of 0.01?". [Aven 2010] pointed that, in practice, decision-making is much more complex and decisions cannot be justified by a simple probabilistic criterion. In practice, the simplified scheme of uncertainty analysis can be viewed as part of a more complex process [Aven & Zio 2011] of probabilistic risk analysis, which is focused on both consequences and the associated uncertainties. So, uncertainty analysis must be viewed as a (major) input of a wider analysis process.

Even if, in the common sense, this seems not to be a decision problem (rather, a statistical estimation problem), decision theory [Berger 1985, Parmigiani & Inoue 2009] is an useful tool to address it. And after all, following the teachings of [Bernardo & Smith 1994]: "*the supposed dichotomy between inference and decision is illusory, since any report or communication of beliefs following the receipt of information inevitably itself constitutes a form of action.*"

Within this framework, we notice that the random output Z has a pdf that depends on the parameter $\boldsymbol{\theta}$ of the inputs' pdf, although the initial question about ϕ was asked as if perfect knowledge was available about $\boldsymbol{\theta}$. As a matter of fact, the quantity of interest ϕ does depend on the possible value of parameter $\boldsymbol{\theta}$ and, in general, one should explicitly write $\phi = \phi(\boldsymbol{\theta})$ not to forget that knowing $\boldsymbol{\theta}$ is a prevailing condition before getting the ideal result. Of course, it has to be kept in mind (cf. Sections 5 and 6) that the evaluation of $\phi(\boldsymbol{\theta})$, even for a fixed value of $\boldsymbol{\theta}$, could be in practice a numerical challenge, e.g. when $\phi(\boldsymbol{\theta})$ is a low probability of failure and $G(\cdot)$ is a high CPU consuming computer model, but conceptually it is nothing but a deterministic function of $\boldsymbol{\theta}$. We will come back to this important point in the discussion of Section 4.

Sometimes, one deliberately neglects the uncertainty on $\boldsymbol{\theta}$ by simply considering that the method used for assessing $\boldsymbol{\theta}$ has provided a value, say $\widehat{\boldsymbol{\theta}}$, that is considered "good enough" for the purpose of uncertainty analysis. The estimation of ϕ is given by the value taken by the function $\phi(\boldsymbol{\theta})$ when $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ (plug-in approach):

$$\widehat{\phi} = \phi(\widehat{\boldsymbol{\theta}}). \tag{VI.1}$$

Some conceptual and practical problems arise when the analyst explicitly takes into account the uncertainty which affects the parameter $\boldsymbol{\theta}$. This setting, sometimes called *two-level approach* [Limbourg & de Rocquigny 2010], clearly shows two different layers: aleatory uncertainties over $\boldsymbol{X}$, ruled by the probability model $p(\boldsymbol{x}|\boldsymbol{\theta})$ (restricting the actual probability model to belong to a hypothesized parametric family whose elements are picked by tuning the possible value of $\boldsymbol{\theta}$), and epistemic uncertainties over $\boldsymbol{\theta}$ (i.e. the lack of knowledge of $\boldsymbol{\theta}$) [Aven 2011]. Indeed, as $\boldsymbol{\theta}$ is uncertain, and represented by a random variable, ϕ=$\phi(\boldsymbol{\theta})$ will be uncertain too. This means that, strictly speaking, the result of the study should be the pdf quantifying the uncertainty representation of the desired quantity of interest $\phi(\boldsymbol{\theta})$. Properly accounting for aleatory and epistemic uncertainties and separating the effects of these quite different sources of uncertainties is a fundamental issue, as highlighted, for instance in [Helton *et al.* 2011, Helton & Sallaberry 2012]. Here, several evaluations of the quantities of interest are made conditional to a given number of values of the epistemically uncertain variables: the result is typically shown as a set of curves (e.g. cumulative distributions functions of the variable of interest). But, sometimes, this kind of result is not satisfactory and the engineer is asked to provide a recommendation under the form of

a point evaluation of ϕ.

The following sections are focused on the point estimation of ϕ in a two-level probabilistic approach. In the light of the statistical decision theory, we reinterpret some standard practices and discuss their underlying, and most of the time hidden, hypotheses.

In the remainder (Section 2) we briefly recall the basic principles of statistical decision theory and show how the popular (and seemingly generic) two-level approach, consisting in propagating through the model $G(\cdot)$ the predictive distribution of $\boldsymbol{X}$:

$$\int_{\boldsymbol{\theta}} p(\boldsymbol{x}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta})\, d\boldsymbol{\theta}, \tag{VI.2}$$

where $p(\boldsymbol{\theta})$ is the pdf representing the epistemic uncertainty of parameter $\boldsymbol{\theta}$, induces implicitly the choice of a particular cost function that depends on the mathematical expression of the chosen quantity ϕ. Then, Section 3 compares plug-in and different two-level approaches on a dike reliability assessment toy example. These results and their implications for industrial uncertainty analysis are discussed in Section 4.

2 Decision theory and Bayesian estimation of the quantities of interest

2.1 Cost functions and Bayesian estimation

When adopting explicitly a two-level approach, the probabilistic assessment of $\boldsymbol{\theta}$ leads to consider, together with $p(\boldsymbol{x}|\boldsymbol{\theta})$ an additional pdf, say $p(\boldsymbol{\theta})$. Formally, in the Bayesian framework (see [Bernardo & Smith 1994, Bernardo 2011] for instance), $p(\boldsymbol{\theta})$ expresses the partial and imperfect knowledge the analyst may have about the fixed albeit unknown parameter $\boldsymbol{\theta}$, sometimes called the *state of Nature*. In other words, it is a probabilistic judgement encoding the various odds if one was committed to take bets with regards to the possible values of $\boldsymbol{\theta}$. This pdf could be interpreted, in a more practical way of thinking, as a measure [Aven & Kvaloy 2002] reflecting, for each possible values of $\boldsymbol{\theta}$, the confidence the analyst puts in the probabilistic model $p(\boldsymbol{x}|\boldsymbol{\theta})$ for predicting the random variable $\boldsymbol{X}$.

In practice, the analyst chooses for $p(\boldsymbol{\theta})$ one of the following:

- the prior distribution $\pi(\boldsymbol{\theta})$ (see [Helton 1996, Krzysztofowicz 1999], for instance) when inference on the output Z is purely based on expert knowledge concerning the physical process generating the data;
- the Bayesian posterior distribution $\pi(\boldsymbol{\theta}|\mathcal{D})$, i.e. the normalized product of the prior $\pi(\boldsymbol{\theta})$ and data likelihood $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$, when some data $\mathcal{D}$ are available, in addition to prior expertise;
- some approximation of the posterior above, the most common one being the Gaussian asymptotic distribution (cf. [Berger 1985], p. 224), obtained formally as the pdf of the *frequentist* estimator (the maximum of likelihood, MLE) in which the actual unknown value of $\boldsymbol{\theta}$ has been replaced by its estimate $\widehat{\boldsymbol{\theta}}$.

One can notice that, doing that, the analyst clearly stays in a Bayesian framework. According to the Bayesian viewpoint, estimation is a particular kind of decision problem: between all possible values of $\phi(\boldsymbol{\theta})$, one must be chosen, taking into account (i) the degree of belief one can put in each one of them, as measured by the distribution $p(\boldsymbol{\theta})$, and (ii) the stakes of the study. We will insist here on this second point.

Depending on the stakes of the study, or in an alternative formulation, the way the predicted value will be used in a more complex decision procedure, one can decide to provide a "central" value (e.g. mean, median) of $\phi(\boldsymbol{\theta})$ or a more conservative value, e.g. shifted to the right if the quantity of interest is a failure probability or to the left if it is a safety margin.

Bayesian analysis provides a general solution to this problem, based on decision theory. When choosing an estimator $\widehat{\phi}$, for the unknown quantity of interest, the analyst must minimize a *cost* (or *loss*) function which quantifies, at least conceptually, the costs, the consequences, of choosing a *bad* value for ϕ, i.e. a value which is different from the one he/she would provide in the theoretical case of *perfect* information (i.e. if the value of $\boldsymbol{\theta}$ were known exactly). More formally, let d be a possible value (a "guess") the analyst chooses for the quantity of interest ϕ. A loss function $C(\phi, d)$ measures the cost resulting from the non-optimal "decision" d when the "true" value of the quantity of interest is ϕ. For instance, a popular choice of cost function, meeting the natural requirement of attaining its minimum for $d = \phi$, is the quadratic loss, defined by:

$$C(\phi, d) \propto (\phi - d)^2. \tag{VI.3}$$

Such a loss function implies that under- and over-estimation of ϕ are indifferent.

Of course, as the actual value of ϕ is unknown, it is not possible to evaluate the cost function for a given value of d. Instead, given $p(\boldsymbol{\theta})$, one can evaluate its expectation (expected loss). We note it $\rho(d)$:

$$\rho(d) = \mathbb{E}[C(\phi, d)] = \int_{\boldsymbol{\theta}} C(\phi(\boldsymbol{\theta}), d) \cdot p(\boldsymbol{\theta})\, d\boldsymbol{\theta}. \tag{VI.4}$$

The Bayes estimator of ϕ is the value of d that minimizes the expected loss:

$$\widehat{\phi}_{\text{BAY}} = \underset{d}{\text{Argmin}} \rho(d). \tag{VI.5}$$

As stressed in [Parent & Bernier 2007], this procedure can be interpreted in terms of an *integrated sensitivity analysis*, where each possible cost resulting from decision d is weighted according to the probability $p(\boldsymbol{\theta})$ for such a cost to occur, evaluated conditionally on all the available knowledge.

From a more theoretical perspective, Bayesian estimators are known to have excellent frequentist properties, since they constitute the class of *admissible* decisions. Informally, this means that it is impossible to find an estimator that performs uniformly better than a Bayesian estimator (in terms of the *frequentist risk*, i.e. the expected loss with respect to the data density) [Berger 1985]. Two schools of thought stem from this result:

- frequentist statisticians, following [Wald 1939], consider integrating the cost function with respect to a certain distribution $p(\boldsymbol{\theta})$ as a simple way of obtaining interesting estimators, without giving any particular signification to $p(\boldsymbol{\theta})$;
- Bayesian statisticians go one step further in terms of interpretation, and consider $p(\boldsymbol{\theta})$ as a probabilistic bet on the unknown (but fixed) value of the parameter $\boldsymbol{\theta}$, based on both prior knowledge and the data at hand.

Thus in presence of a loss function, Bayes estimators always appear as a good choice. Note that, on the other hand, the uncertainty analysis may not always have such an explicit link with the decision making process, but may instead be limited to describing the uncertainty affecting an interest quantity. In this case, the advantage of choosing a Bayesian over a frequentist approach is less obvious; it is more a matter of opinion, as well as of practical convenience, and depends on the particular model at hand.

Choosing a cost function. Often used as a default choice, the quadratic loss can nevertheless be inappropriate if the cost depends on whether ϕ is over- or under-estimated. For instance, let us consider the case where ϕ is a failure probability

used as a criterion to set a condition-based maintenance policy. Here, the effective failure of an industrial component (more frequent when ϕ is underestimated) has much more costly consequences than those resulting from stopping or fixing the production chain (more frequent when ϕ is overestimated).

In this case an asymmetric cost function might be preferable, such as the weighted absolute loss:

$$C(\phi, d) = k_1(\phi - d) \cdot \mathbb{1}_{\{\phi \geq d\}} + k_2(d - \phi) \cdot \mathbb{1}_{\{d > \phi\}}. \tag{VI.6}$$

Its use implies that the additional cost of decision d is proportional to its absolute deviation from ϕ, multiplied by a different factor, k_1 or k_2 respectively, depending on whether d under- or over-estimates ϕ.

Though the above cost functions are most common, many others have been used in decision-oriented estimation problems, such as the $\alpha-$absolute loss [Ren *et al.* 2004], the LINEX [Varian 1974] or the entropy loss [Robert 2001]. Going back to the problem of tail probability estimation, in some cases, we are less interested in the probability itself than by its order of magnitude, that is, its logarithm. Thus, we might consider using the log-quadratic loss:

$$C(\phi, d) \propto (\log \phi - \log d)^2. \tag{VI.7}$$

All three losses are illustrated in Figure VI.1, in the case where ϕ is a tail probability. In this picture, the true value of the quantity of interest is supposed to be 0.05.

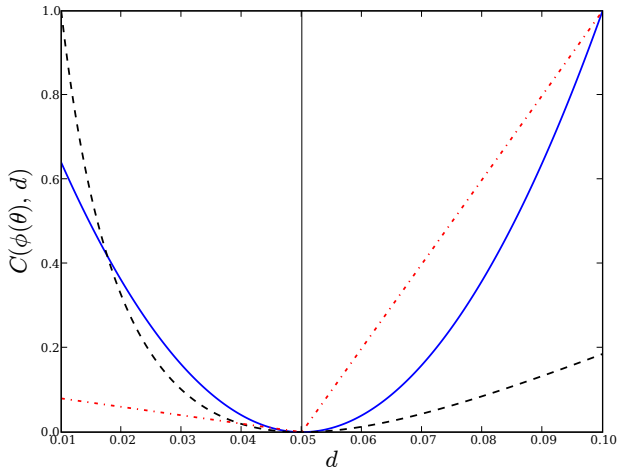


Figure VI.1 – Cost functions for the estimation of a tail probability. Continuous line: quadratic loss, dashed line: log-quadratic loss, dash-dotted line: weighted absolute loss, with $k_2 = 10 \times k_1$. The vertical line corresponds to the true value ϕ. Cost functions are normalized for viewing convenience.

Practical implementation. To implement the method described in Section 2.1, one needs the expression or, more often, a sample of the pdf of the quantity of interest. This is the main ingredient, and indeed the most difficult to obtain, for evaluating the integral in Equation VI.4. More precisely, the implementation proceeds through the following steps:

1. assessing epistemic uncertainty, i.e. the pdf $p(\boldsymbol{\theta})$;
2. generating a sample $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(m)}$ from $p(\boldsymbol{\theta})$;

3. for each $\boldsymbol{\theta}^{(i)}$ $(i = 1, \ldots, m)$, evaluating the corresponding quantity of interest $\phi(\boldsymbol{\theta}^{(i)})$. Then one can evaluate, for any possible "candidate" value of $\widehat{\phi}_{\text{BAY}}$, say d, the expected loss, i.e. the integral in Equation VI.4, by the Monte Carlo method: $\rho(d) \approx (1/m) \cdot \sum_{i=1}^{m} C\left(\phi(\boldsymbol{\theta}^{(i)}), d\right)$;

4. finding the value of d that minimizes $\rho(d)$. That will be the chosen estimate $\widehat{\phi}_{\text{BAY}}$.

In practice, Step 3 is the most tricky, as it demands several evaluations of ϕ. A single evaluation of ϕ is generally performed by Monte Carlo simulation, which demands, itself, several evaluations of the (possibly) time consuming numerical code $G(\cdot)$. The calculations are much easier when the expression of the function $G(\cdot)$ is explicit, as it will be the case in the following numerical example.

2.2 Reinterpreting the common predictive approach

The general formulation (Equation VI.5) of the Bayesian solution to the point estimation problem can bring some light to the current predictive approach. It consists in first propagating the predictive distribution (Equation VI.2) of $\boldsymbol{X}$ trough the deterministic model $G(\cdot)$ to obtain the predictive distribution $\tilde{p}(z)$ of Z (in practice a sample of it), then estimating any characteristic quantity of the (unknown) density $p(z|\boldsymbol{\theta})$, such as expected values, tail probabilities, quantiles, etc., by the corresponding characteristic quantity of $\tilde{p}(z)$ [Geisser 1971, Christensen & Huffman 1985].

Can these *predictive estimators* (PE) be interpreted as Bayes estimators? What is the associated cost function? The answer depends on the expression of the quantity of interest ϕ. We will consider the two cases hereinafter: (i) ϕ is a probability of failure, (ii) ϕ is a quantile.

In the first case, ϕ is, indeed, the mean of a particular deterministic function of the output Z:

$$\phi = \mathbb{P}[Z > z^{\star}] = \mathbb{E}\left[\mathbb{1}_{\{Z>z^{\star}\}}\right]. \tag{VI.8}$$

Using standard results from Bayesian analysis, one can prove that the predictive estimator:

$$\widehat{\phi}_{\text{PE}} = \int \mathbb{1}_{\{Z>z^{\star}\}} \cdot \tilde{p}(z)\, dz \tag{VI.9}$$

is the Bayes estimate of $\phi(\boldsymbol{\theta}) = \mathbb{P}[Z > z^{\star}|\boldsymbol{\theta}]$ associated to the quadratic loss function (Equation VI.3). This result remains valid for any quantity of interest which can be expressed as the mean of any deterministic function of the variable of interest Z. The proof is given in [Christensen & Huffman 1985].

Thus, in this case the predictive estimator is actually a Bayes estimator (in disguise), associated to the quadratic loss. This procedure implicitly forces the choice of a particular symmetric cost function, which has been chosen based on the expression of the quantity of interest, rather than on decisional aspects. In most cases, when estimating a probability of failure, under- and over-estimations have quite different consequences, so, common good sense would favour a asymmetric cost function rather than the quadratic loss.

Consider now the case of a quantity of interest ϕ which is a quantile of Z associated to a given probability α. In this case the predictive estimator is the α-quantile of the distribution $\tilde{p}(z)$, i.e. the quantity $\widehat{q}_{\alpha}^{\text{PE}}$, such that:

$$\mathbb{P}[Z \leq \widehat{q}_{\alpha}^{\text{PE}}] = \int_{-\infty}^{z=\widehat{q}_{\alpha}^{\text{PE}}} \tilde{p}(z)dz = \alpha. \tag{VI.10}$$

According to a classical result of Bayesian analysis (cf. [Berger 1985] or [Parent & Bernier 2007], for instance), $\widehat{q}_{\alpha}^{\text{PE}}$ is formally the Bayes estimate of Z relative to the weighted absolute loss (Equation VI.6), with $\alpha = k_1/(k_1 + k_2)$. Note

that, formally, this predictive estimator is not a Bayes estimator of the quantile of Z given the unknown $\boldsymbol{\theta}$, in the sense of Equation VI.5.

The above interpretation of $\widehat{q}_\alpha^{\mathrm{PE}}$ as a predictor of the future observation Z is reasonable when Z itself is central to the decision process. An example of such a situation is described in [Parent & Bernier 2007], in the context of dam design. Here, Z represents the yearly maximal water level of a river, and d the height of a dam to be constructed next to the same river. Using the absolute weighted loss, the optimal dam height is seen to be precisely $\widehat{q}_\alpha^{\mathrm{PE}}$, if α is the relative cost resulting from a flood (due to an undersized dam), when compared to the sum of the global costs (expected damages + sure investments).

However, if the decision-maker needs to consider costs that are directly related to the actual value of the quantile q_α, then the PE should not be used, since it addresses a decisional problem that has nothing to do with quantiles. A simple approach in this case would be for instance to adopt the quadratic cost function (Equation VI.3), in which case the Bayesian estimate of q_α is the posterior mean of $q_\alpha(\boldsymbol{\theta})$.

3 A simple numerical example

We now compare the MLE plug-in to the Bayes estimates on a case study, concerning the safety evaluation of a flood protection dike.

There exists a rich literature on this subject, an overview of which can be found in [Miquel 1984], with recommended methodological guidelines for hydraulic engineers. Following the terminology introduced in this work, the following example uses the *method of yearly maxima*, meaning that the flood probability is estimated from a record of yearly maximal river discharges.

The variable of interest Z is here the yearly maximal water level of the river at the location of the dike, noted Z_c in the following. Following the toy case-study also used in Chapter IV (cf. page 100), we assume that Z_c can be computed given a number of input variables, according to the analytical formula:

$$Z_c = Z_v + \left\{ Q \Big/ \left(BK_s \sqrt{(Z_m - Z_v)/L} \right) \right\}^{0.6}. \tag{VI.11}$$

Additionally, we note $Z_d = 56\,m$ the altitude of the protection dike, and consider the problem of estimating the probability for a flood to occur, that is, the probability:

$$P_f \quad = \quad \mathbb{P}(Z_c > Z_d)$$

that the maximal water level Z_c exceeds Z_d (probability of failure), in view of verifying that it remains below a given imposed threshold α, chosen here as $\alpha = 0.01$. This problem can equivalently be formulated as that of estimating the $(1-\alpha)-th$ percentile:

$$q_{(1-\alpha)} \quad ; \quad \mathbb{P}(Z_c > q_{(1-\alpha)}) = \alpha, \tag{VI.12}$$

in view of verifying that $q_{0.99}$ does not exceed the dike's altitude Z_d.

3.1 Probabilistic model

We assume that, among the inputs of the model, the width B and the length L of the river section are perfectly known and that, for this example, their values are set to 300 and 5 000 m respectively. Z_m, Z_v and K_s are considered as random variables with known distributions (i.e. without epistemic uncertainty), given by:

- $Z_m \sim \text{Tr}(49, 50, 51)$;
- $Z_v \sim \text{Tr}(54, 55, 56)$;
- $K_s \sim \text{Norm}(30, 7.5)$,

where $\text{Tr}(a,b,c)$ is the triangle distribution with mode b and extremal values a and c. The yearly maximal discharge Q is modelled as a random variable following the Gumbel distribution $\text{Gu}(\eta, \beta)$ with location parameter η and scale parameter β. The Gumbel distribution is a standard choice to depict variations of maxima (see [Coles 2001]) and its probability density function is given by:

$$\text{Gu}(q|\eta,\beta) = \frac{1}{\beta}\exp\left[-\exp\left(\frac{\eta-q}{\beta}\right)\right]\exp\left(\frac{\eta-q}{\beta}\right). \tag{VI.13}$$

As in most cases, the vector of model parameters $\boldsymbol{\theta} = (\eta, \beta)$ is of course unknown, and must be estimated using expert knowledge and/or the recorded data set. For the example developed hereafter, we consider a simulated sample of $n = 30$ annual maximal values, as described below. The two quantities of interest considered here, P_f and $q_{0.99}$, are functions of the distribution of the maximal water level Z_c, itself a function of $\boldsymbol{\theta}$, hence we note them $P_f(\boldsymbol{\theta})$ and $q_{0.99}(\boldsymbol{\theta})$ in the following.

Note that, for any given value of $\boldsymbol{\theta}$, neither of these quantities have analytical expressions. Instead, Monte Carlo estimations may be derived from a sample ${Z_c}^{(1)}, \ldots, {Z_c}^{(m)}$ of the maximal water level's distribution.

3.2 Synthetic dataset

We simulated 30 realizations of the Gumbel distribution with shape parameter $\eta_0 = 1000$ and scale parameter $\beta_0 = 600$ (see Table VI.1). We used this sample, symbolizing maximal annual discharge values, as input data for our numerical experiment, "forgetting" the real parameter value used to generate the data (that is, forgetting that $(\eta, \beta) = (\eta_0, \beta_0)$). Note that this is purely a toy example, meant to illustrate the advantages of Bayesian methodology for point estimation in uncertainty analysis. In particular, we do not address here the art of choosing an appropriate data model, which is indeed an issue when analyzing real data, combining parsimony concerns, the quest of realism as well as the analyst's skill and experience.

$$\begin{array}{rcccccc}
\mathcal{D} = & \{1\,306, & 1\,235, & 1\,830, & 2\,442, & 1\,128, & 3\,330, \\
& 1\,530, & 3\,192, & 2\,647, & 238, & 706, & 1\,903, \\
& 1\,594, & 935, & 1\,100, & 2\,204, & 1\,366, & 1\,629, \\
& 522, & 642, & 1\,173, & 424, & 1\,837, & 1\,391, \\
& 789, & 383, & 1\,858, & 917, & 1\,084, & 1\,026\}
\end{array}$$

Tableau VI.1 – 30 discharge values, simulated from the $\text{Gu}(1\,000, 600)$ distribution.

Plug-in approach. The simplest way of assessing whether P_f and $q_{0.99}$ are within the prescribed security bounds is via the plug-in approach (one-level approach). Accordingly, we computed the MLE of the Gumbel parameters from our simulated dataset, using standard optimization algorithms, and found:

$$\widehat{\eta}_{\mathrm{ML}} = 1\,059; \quad \widehat{\beta}_{\mathrm{ML}} = 607.$$

Computing the corresponding plug-in estimates of the considered quantities of interest, using a Monte Carlo approximation, yielded:

$$\widehat{P}_f = 0.0085; \quad \widehat{q}_{0.99} = 55.87\ m.$$

By a strict application of the probabilistic criterion, one could conclude from here that the dike under study is safe since $\widehat{P}_f < 0.01$, or, equivalently, $\widehat{q}_{0.99} < 56$ m. However, this answer is not fully satisfactory since the estimated values of the quantities of interest are very close to their safety bounds. It is therefore necessary to account for the uncertainty on the unknown parameter vector $\boldsymbol{\theta}$ when estimating P_f and $q_{0.99}$. A fashion for doing it is using a Bayesian approach, as described in Section 2, to derive optimal estimators of our quantities of interest under uncertainty, through the minimization of a posterior expected loss that encodes the costs resulting from estimation errors.

Prior and posterior distributions of model parameters. We used the following priors to describe the preliminary knowledge about the distribution of the yearly maximal discharge:

- Location parameter: $\pi(\eta) = \mathrm{Ga}(1, 500)$;
- Scale parameter $\pi(\beta) = \mathrm{Ga}(1, 200)$,

where $\mathrm{Ga}(a, b)$ is the Gamma distribution with shape parameter a and inverse scale b. Such priors are poorly informative since they are highly dispersed, with their best guesses (means) respectively located at $\eta = 500$ and $\beta = 200$.

The joint posterior $\pi(\eta, \beta|\mathcal{D})$ has no analytical expression, but many different methods exist to draw a sample from it. We chose the accept-reject algorithm, which consists in drawing candidate values from any proposal density $g(\eta, \beta)$ approximating the target posterior density, and accepting each candidate $(\eta^{(c)}, \beta^{(c)})$ with probability proportional to $\pi(\eta^{(c)}, \beta^{(c)}|\mathcal{D})/g(\eta^{(c)}, \beta^{(c)})$. This strategy has the advantage over the more generally applicable Monte Carlo Markov Chain (MCMC) techniques of producing independent draws from the exact posterior distribution.

We used as a proposal density the normal asymptotic approximation of the posterior:

$$\mathrm{Norm}\left(\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}}, \frac{1}{n}\mathcal{I}^{-1}(\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}})\right),$$

where $\mathcal{I}^{-1}(\cdot)$ is the inverse of the Fisher information matrix. In this fashion, we obtained $m = 10^4$ draws $(\boldsymbol{\theta}^{(i)})_{1 \leq i \leq m}$ from the posterior distribution, as illustrated in Figure VI.2. Thus, instead of considering a single possible value for $\boldsymbol{\theta}$ as in the previous plug-in approach, the Bayesian paradigm considers a full distribution of possible values, explicitly describing the epistemic uncertainty about $\boldsymbol{\theta}$ initially coded by $\pi(\boldsymbol{\theta})$ and updated using the information provided by the data.

Bayesian estimates. From this sample of the model parameters posterior distribution, we were able to derive samples from the posterior distributions of the desired quantities of interest, by computing for each draw $\boldsymbol{\theta}^{(i)}$ corresponding values $P_f(\boldsymbol{\theta}^{(i)})$ and $q_{0.99}(\boldsymbol{\theta}^{(i)})$. These distributions are shown in Figure VI.3 (middle). In particular, we can see that the posterior

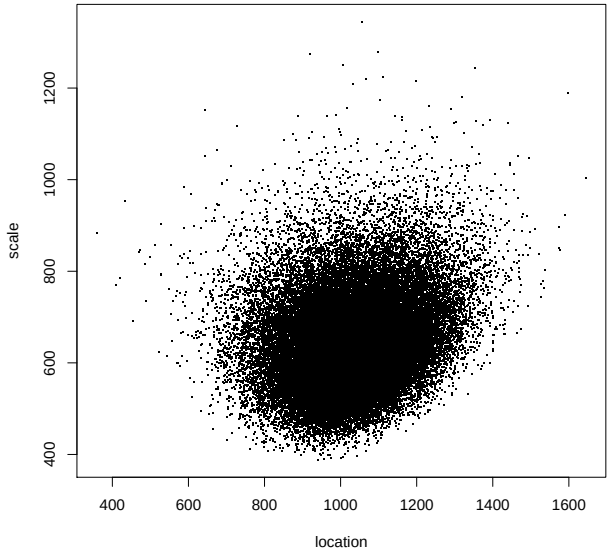


Figure VI.2 – 10^4 draws from the posterior distribution of the maximal discharge density parameters (η, β).

probability for the safety bound $P_f < 0.01$ (or, equivalently, $q_{0.99} < 56$) to be met is roughly "around 0.5", meaning that data alone cannot provide a clear answer to whether the safety condition is met.

With this in mind, Bayesian estimates for both quantities of interest can be derived for any loss function formalizing what we considered to be the costs of over- and under-estimation.

For instance, we can first imagine that no conservatisms have to be considered. This can be the case if we already know that the provided value will anyway be penalized (in some manner) by the end-user in a further risk analysis or decision process. In this case, the quadratic loss seems a reasonable choice. Accordingly, we estimated both quantities by their posterior means (see Figure VI.3, solid vertical line):

$$\widehat{P_f}^{\text{BAY}} = 0.0107; \quad \widehat{q_{0.99}}^{\text{BAY}} = 55.96\ m.$$

These values are very close to those obtained by ML plug-in approach, which is not very surprising given that the mean of the posterior distribution can be approximated by the ML estimator, as also recalled in section 2.1.

However, the quadratic loss equally penalizes over- and under-estimation, which may be inappropriate in contexts where a more conservative value of the reliability criterion is expected.

We can realistically imagine, in a design context, that the flood probability will be directly used to decide whether a goods warehouse will be located (if $\phi < 0.01$) or not ($\phi > 0.01$) near the riverside.

In this case, even if the "effective" costs of the consequences of the decision are not precisely known, at least at this (early) stage of uncertainty analysis, we may have some reasons to give a more conservative estimation of the quantity of interest, for instance, penalizing (using the weighted absolute loss function) under-estimation 9 or 99 times more than its

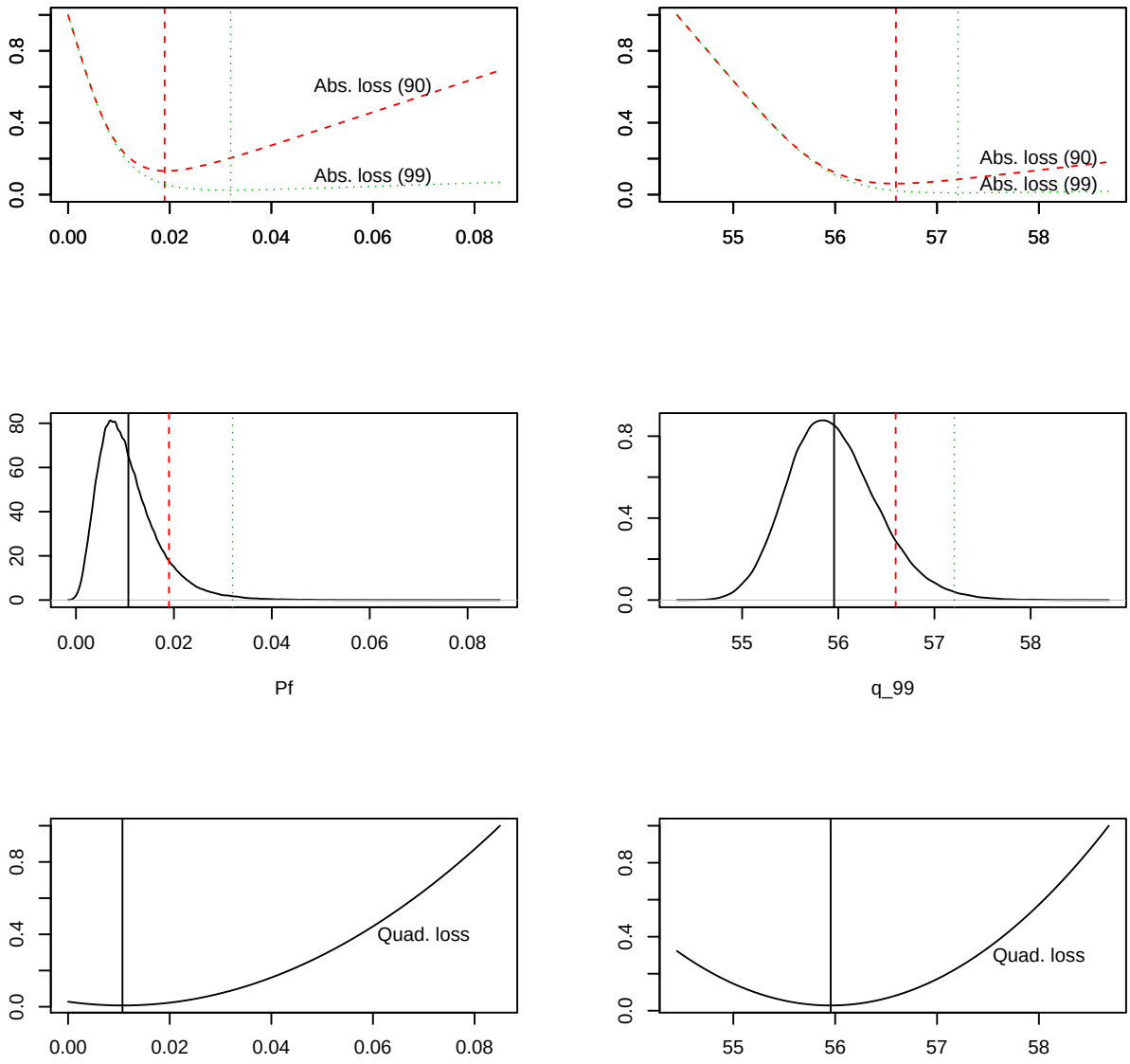


Figure VI.3 – Bayes estimates of dike failure probability P_f (left) and water level quantile $q_{0.99}$ (right). Top: Expected weighted absolute loss functions (dashed red line corresponds to the 90-th quantile, dotted green line to the 99-th quantile). Middle: posterior distributions of both quantities of interest. Bottom: Expected quadratic loss functions. Vertical lines indicate the Bayesian estimator values.

over-estimation. Hence, we estimated each quantity by its 90-th or 99-th posterior percentile (see Figure VI.3), yielding:

$$\begin{aligned} &\text{90-th percentile:} \quad \widehat{P}_f^{\text{BAY}} = 0.019; \quad \widehat{q}_{0.99}^{\text{BAY}} = 56.60 \text{ m}. \\ &\text{99-th percentile:} \quad \widehat{P}_f^{\text{BAY}} = 0.032; \quad \widehat{q}_{0.99}^{\text{BAY}} = 57.21 \text{ m}. \end{aligned}$$

In both cases, these conservatively biased estimates clearly indicate that the imposed bounds are not respected for the dike

under study.

Thus, in the present case the specification of a cost function, embodying the stakes motivating the uncertainty study, appears critical in choosing the value which will be retained as the final estimate of the interest quantity. A full description of the uncertainty affecting the quantities of interest (such as provided by the posterior distribution), is a valuable information that should be considered a result of the uncertainty analysis as much as the final point estimate. In all cases, even point estimates derived from different cost functions may also provide some useful pieces of information as an aid to the decision process to be guided by the analysis.

4 Comments

Decision theory and Bayesian estimation can help the analyst provide a point estimate of an uncertain quantity of interest, which takes into account both epistemic uncertainty tainting the predicted result and the consequences of its under- or over-estimation. Of course, one can argue that a cost function is a quite narrow view of the stakes that motivate a real industrial study. Nevertheless, it is better than nothing and it can be helpful to let some underlying hypotheses become clearer. However, Bayesian estimation is not a panacea and some theoretical and practical problems arise in deploying this methodology.

Exploiting all the available information. A key point in statistical practice is the need to exploit all the available information, in the best way possible. In real-life industrial applications, uncertainty and risk analysis problems are often ill-posed and information can be scarce. The viewpoint sketched here, clearly Bayesian, adds more ingredients to the current analysis. First, the prior information allows to account for available knowledge, which is particularly useful in presence of poorly informative data. Second, the cost function explicits the consequences of a "bad" prediction in a way that strongly depends on the use of this result for a further risk analysis.

Both ingredients encode, in a certain sense, the background knowledge the analyst brings into her/his study. Hence, the cost function is not a formal input provided "as is" by a hypothetical stakeholder: it is rather a tool to guide the analysis toward a result which is more appropriate to the framework of his/her work than if it were based on popular heuristics.

How can it work in practice? The formal procedures to elicit both loss functions and prior distributions share similar difficulties. The techniques to correctly elicit experts' prior beliefs (i.e. encode their knowledge as probabilistic judgements as in [Kadane *et al.* 1998]) and loss functions are just becoming a mature field of research at the interface between psychology and probability [O'Hagan *et al.* 2006]. Their influence on the outcome should be carefully checked, thus leading to decisional sensitivity analyses as proposed in [Eckert *et al.* 2009]. Finally, [Abraham & Cadre. 2004] studied the conditions under which optimal decisions do not change with regards to a relatively large class of loss functions.

In practice, at least, it does not seem too hard to start the elicitation with a basic question: "*are under- and over-estimation indifferent with respect to further use of the result in a risk analysis context?*". If the answer is yes, then the use of a quadratic cost function is appropriate and will indeed ease calculations. If not, an asymmetric cost function (e.g. the weighted absolute loss) can be used to justify the choice of another estimator, more conservative than the mean. Both quadratic and weighted absolute loss functions are straightforward to interpret and lead to explicit estimators, namely the mean and quantiles of $\phi(\boldsymbol{\theta})$, respectively.

The choice of a quadratic function in case where under- and over-estimation are equally accepted is motivated by practical computational issue. Any other symmetric cost function will comply with the statement above, but the quadratic function spares the analyst to explicitly solve the optimization problem of the Equation VI.5, and seems a pragmatic

choice if the only element for eliciting the stakeholder preferences is the answer to the basic question above. A similar argument could be advanced for motivating the choice of weighted absolute losses as asymmetric cost functions. Of course, if the analyst has at his/her disposal more information for eliciting a more proper cost function, the advantages of using a tailored loss function will justify the increased complexity of calculations.

Furthermore, the above question directly involves the quantity of interest, i.e. the key result of the study. Thus, even if the Bayesian mathematical setting may seem quite theoretical, the cost function refers to a quantity which is clearly understandable by both the analyst and the end-users of the study. Here, we agree with [Aven & Kvaloy 2002] that focusing on observable quantities (rather than on the *state of Nature*) is capital when putting into practice the Bayesian paradigm for uncertainty and risk analysis.

As a practical example of the application of this procedure, in the context of ecological risk assessment, the quantity of interest is typically the minimal concentration HC_α of a certain chemical hazardous to a given proportion α of the species in a given habitat. This is expressed as the α-percentile of the *species sensitivity distribution* (SSD), representing the distribution of tolerance values to the target chemical for a randomly sampled species within the studied habitat. It is argued in [Hickey *et al.* 2009] that the use of the weighted absolute loss or the LINEX loss allows to rationally choose an estimator that optimizes the costs associated with over- and under-estimations. Note that, in this context, the predictive approach would be totally useless to estimate HC_α since it would be unable to account for the costs resulting from the different types of estimation errors.

Theoretical considerations. From a strictly theoretical point of view, our reinterpretation of the predictive approach puts into evidence a certain lack of coherence in the common predictive approach.

First (and most important), as stressed in Section 2.2, the predictive approach can be viewed as a particular case of Bayesian estimation with a cost function which is completely imposed by the nature of the quantity of interest. Here, the major drawback is that this procedure automatically picks, without the involvement of the analyst nor the agreement of the decision maker, the cost function to be minimized. On the contrary, it should be chosen depending on the background and the purpose of the study.

Second, the result of the predictive approach is delicate to interpret since the predictive density has no phenomenological interpretation, and should only be interpreted in terms of probabilistic bets. Thus, the interpretation of the predictive pdf as "the density of a future observation given the data" must be understood strictly as the subjective belief concerning the values of the output Z, based on the data at hand, and not as the natural variability of Z seen as a physical quantity. Indeed, the latter is given by the (unknown) output density $p(z|\boldsymbol{\theta})$, which is independent of whatever data available once the unknown $\boldsymbol{\theta}$ is revealed.

Difficulties. Of course, in real-world studies, the Bayesian procedure is made difficult due to several practical problems which should not be overlooked.

First, one could point the classical arguments about the "difficulty" of the Bayesian analysis, namely the posterior inference when the probabilistic model is not conjugate. That is, in our opinion, a quite old-fashioned point of view. Nowadays, a wide palette of methods, tools and software for Bayesian calculation are available and they do not seem much more difficult than other mathematical tools more currently used by the engineers. See [Marin & Robert 2007], for instance, for an excellent overview of the current state of the art in Bayesian estimation techniques, and [Boreux *et al.* 2010] for their application to several case studies. When comprehensive Bayesian approaches are unpractical, it is also possible to adopt simplifying strategies, such as in [Krzysztofowicz 1999, Maranzano & Krzysztofowicz 2008]. These consist in focusing

the Bayesian treatment on the most important sources of uncertainty with respect to the decisional problem considered, treating other sources by more conventional methods, such as MLE.

Second (and most important), in many cases (e.g. in reliability analysis, when ϕ is a low probability of failure) one single evaluation of ϕ can already be a numerical challenge, demanding several runs of a CPU time consuming numerical code. Then, the resolution Equation VI.5, which involves several evaluation of ϕ (for several values of $\boldsymbol{\theta}$) can be infeasible. Indeed, this problem is common to all level-two approaches and no miracle solution exists.

In such cases, the classical tools to speed-up the evaluation of ϕ can be used, such as FORM/SORM approximation, accelerated Monte Carlo sampling, metamodelling, together with parallel or distributed computing. In reliability analysis, this problem has been raised, for instance, in [Der Kiureghian 2008], where the use of a Gaussian approximation of the density of ϕ is advocated. In particular, metamodelling (e.g. polynomial chaos, kriging) seems to be a very useful technique for coping with this issue. In this case, however, the approximation error should also be assessed and accounted for in the estimation of ϕ.

Of course these problems exist and we cannot ignore them. However, these practical difficulties do not seem a valid reason to elude the debate. Of course, if the Bayesian procedure described above should really be impossible to perform, the predictive approach remains a valid alternative, as it demands a single evaluation of ϕ. But the analyst has to keep cautious because this method may lead to a non-conservative estimation of ϕ. This can eventually lead the analyst to penalize the estimator provided by the predictive approach, when using it in a further risk analysis.

Another solution, as heuristically sketched in [de Rocquigny 2006], is to use a plug-in approach with a penalized value of $\boldsymbol{\theta}$, for instance a conservative quantile of $p(\boldsymbol{\theta})$ (the "conservatism" of which can be possibly difficult to justify if $\boldsymbol{\theta}$ is multidimensional). Even though these solutions seem to be rooted in a certain common sense, unfortunately they cannot be justified theoretically.

Avoiding to hide the dust under the carpet. The take-home message we wish to address here is that the Bayesian procedure described above, in spite of its (effective or apparent) difficulties, allows to make explicit hypotheses and modelling choices which could be hidden in common approaches and are indeed important.

Uncertainty analysis is meant to provide input elements to risk analysis and decision-making and it is capital to quantify the epistemic uncertainty tainting the final result and to provide an outcome which takes into account (in a formal and explicit way) the stakes motivating the study.

The Bayesian rationale provides a coherent solution to deal quantitatively with this problem. It consists in:

- encoding prior expertise as a probabilistic judgement quantifying epistemic uncertainty;
- updating this prior bet through the information conveyed by the data thanks to the model;
- encoding socio-economic consequences into a cost function expressing the discrepancy between the optimal decision that could only have been taken under perfect knowledge and the actual decision;
- recommending the decision that minimizes the expected regret due to imperfect assessment of the unknowns.

Avoiding any dogmatism, we acknowledge that it is neither the only solution to this problem nor the easiest one to put into practice, but it can nevertheless help to bring more light to the different steps of the analysis, and let the end-user of the result be more aware of the assumptions on which this analysis is based.

5 Further considerations

5.1 A more general result

The methodological considerations sketched hereinbefore have been extended in a further study [Keller *et al.* 2011c] (actually published few weeks before the initial study reported in the previous sections). Here, the state of Nature (noted Θ) includes, together with the joint cdf $F(X)$ of the input X, also the code $G(\cdot)$, as also evoked by [Bernier 1980]. Actually, $\Theta = (F, G)$ represent all the (functional) parameters, the uncertainty analysis and the estimation of the desired quantity of interest depend of. In other terms, every estimation is made conditional to a given Θ. For instance, one can imagine that the model $G(\cdot)$ depends on an uncertain parameter (e.g. the Strickler's roughness coefficient K_s in a hydraulic model). In this case, in a Bayesian setting the probability distribution of K_s can be seen as a probability measure of all possible hydraulic models.

The same theoretical approach is presented by [Aven 2010], who basically refers to Θ as the "background knowledge" the risk analyst has of the phenomenon under investigation.

If we note $H(\cdot|\Theta)$ the cumulative distribution function (cdf) of the output Z, the quantity of interest can be noted $\phi(\Theta) = \phi\left(H(\cdot|\Theta)\right)$ to highlight that ϕ depends on Θ via the cdf H of Z.

Again, the plug-in approach, consists simply in replacing Θ with its best estimate (cf. Equation VI.1).

The popular predictive approach (cf. Section 2.2) consists actually in replacing H by its predictive estimate (which turns out to be simply its expectation):

$$\hat{H}_{\text{pred}}(z) = \mathbb{E}_{\Theta}\left[H(z|\Theta)\right] = \int H(z|\Theta) \cdot p(\Theta) d\Theta, \tag{VI.14}$$

in which $p(\Theta)$ is a prior or a posterior distribution depending on whether data are available or not. Hence, the predictive estimator of ϕ is:

$$\hat{\phi}_{\text{pred}} = \phi\left(\hat{H}_{\text{pred}}(\cdot)\right). \tag{VI.15}$$

The result concerning the "Bayesian-theory" reinterpretation of popular predictive estimators in engineering risk assessment studies (cf. Section 2.2) can be generalized by the following theorem.

Theorem 5.1 *Under the assumptions that:*

1. *as function operating on the space of cumulative distribution function $H(\cdot)$, ϕ can be defined as the minimiser of a certain (cost) function $C(d, H(\cdot))$:*

$$\phi\left(H(\cdot)\right) = \underset{d}{Argmin}\, C\left(d, H(\cdot)\right)$$

2. *the function C respects the following property:*

$$\sum_j p_j C\left(d, H_j(\cdot)\right) = C\left(d, \sum_j p_j H_j(\cdot)\right),$$

for any cdf $H_j(\cdot)$ and for any set of $p_j \in [0,1]$, such that $\sum_j p_j = 1$,

then, the predictive estimator $\hat{\phi}_{pred}$ defined by Equation VI.15 is the Bayes estimator of ϕ related to the cost function C.

The proof is given in Appendix, page 159.

The interpretation of this theorem is straightforward: predictive estimators of quantities of interest are actually Bayesian estimators related to particular cost functions, completely determined by the expression of the quantity of interest. Following this methodology which seemingly "accounts for all possible uncertainties tainting the output Z", the risk analyst delegates the choice of the cost function to the heuristic he/she uses!

The assumptions above are not restrictive and this result applies for the main quantities of interest currently used in risk assessment and in particular for the three cases sketched hereinafter.

The case of a probability of failure. Let:

$$\phi(\Theta) = \mathbb{P}[Z \geq z^\star] = 1 - H(z^\star). \tag{VI.16}$$

One can imagine that here $z^\star$ is the height of a protection dike and Z a water level evaluated by means of a hydraulic computer model. The predictive estimator of this quantity of interest is:

$$\hat{\phi}_{\text{pred}} = 1 - \hat{H}_{\text{pred}}(z^\star).$$

As seen hereinbefore, this predictive probability of failure is nothing but the posterior mean $\mathbb{E}_H[\mathbb{P}[Z \geq z^\star]]$, i.e. the Bayes estimator related to the quadratic cost function $C(d, H(\cdot))$, which, neglecting the term not depending on d can be written as:

$$C(d, H(\cdot)) = d^2 - 2d(1 - H(z^\star)).$$

We are clearly in the hypothesis of Theorem 5.1 as $\mathbb{P}[Z \geq z^\star]$ is the solution of an expected-cost minimization problem and the cost function C is "linear" in the sense of Hypothesis 2 of Theorem 5.1.

The case of a quantile. Following the hydraulic risk assessment example already introduced in Section 3, let us consider the case where the quantity of interest is a quantile of the water level Z, e.g. the centennial level $q_{0.99}$, corresponding to an annual exceedance probability of 1-0.99:

$$\phi(\Theta) = H^{-1}(0.99|\Theta) \quad \text{and} \quad \hat{\phi}_{\text{pred}} = H^{-1}_{\text{pred}}(0.99). \tag{VI.17}$$

In this case, the expression of $\phi(\Theta)$ may be written as the minimizer of a piecewise linear cost function that can be written (after some algebra [Parent & Bernier 2007]):

$$C(d, H(\cdot)) = \int_{-\infty}^{d} H(z|\Theta)\, dz - 0.99\, d$$

which obeys to Hypothesis 2. The predictive quantile is thus the Bayes estimator related to this cost function.

The case of an optimal design value. Let us consider the quantity of interest representing the optimal design value for the height a flood protection dike, given by the minimiser of the following cost (studied in detail by [Bernier 2003]):

$$C(d, H(\cdot)) = I_0 \cdot d + C_0 \cdot \mathbb{E}_H\left[\mathbb{1}_{\{Z>d\}}(Z-d)^2\right], \tag{VI.18}$$

in which d is the height of the dike and I_0 and C_0 initial and damage marginal cost (e.g. in M€/m). This case again obeys to hypotheses of the aforementioned theorem and thus the same conclusion applies with respect to the predictive

estimators.

5.2 Descriptive approaches

The Bayesian point estimation of quantities of interest, requires the analyst has at his/her disposal a cost function (cf. discussion of Section 4) summarizing the decisional stakes underlying the study. Nevertheless, in some cases the analyst is simply asked to verify the compliance with a safety criterion, generally formulated as an inequality: $\phi \geq \phi^{\star}$ and it is extremely difficult for him/her to quantify a priori, by means of a simple cost function, the consequences of the under- or over-estimation of ϕ.

That can happen when the study is part of a much wider analysis aiming at assessing the risk related to catastrophic and extremely costly event as in [Helton 1996, Helton & Sallaberry 2012] or [Maranzano & Krzysztofowicz 2008].

In these cases, an extremely valuable result is "simply" the representation of the uncertainty taintig ϕ by means of an empirical distribution function or a credibility interval. In practice, the analyst stops his/her study at the distribution functions shown in the middle of Figure VI.3. That may seem somehow reductive, but it has the advantage to let the decision maker be aware of the uncertainty tainting the risk assessment and can motivate uncertainty mitigation actions, as for instance collecting more data for better characterizing the phenomena under investigation.

6 Extension to a functional criterion: the fragility curve

The decisional approach sketched in the previous section aims at providing Bayes point estimators of quantities of interest used in risk and safety assessment. The methodology shown hereinbefore is presented with respect to scalar quantities of interest. Of course, it can be extended to vector quantities as well as to functional risk criteria. In particular, this Bayesian setting has been applied to the estimation of seismic "*fragility curves*", a popular functional risk criterion, commonly used in many engineering fields and namely in nuclear safety studies.

According to the norm ISO13824 [ISO 2009] *"a fragility curve describes the probability that the actual damage to a structure exceeds a damage criterion, when the structure is subjected to a specified load intensity"*. In the case of seismic risk assessment the load is usually expressed as a scalar characteristic of a seismic signal, typically the horizontal peak ground acceleration (PGA, cf. Figure VI.4), common choice in civil engineering [Solomos *et al.* 2008, Huang *et al.* 2011].

For a given value of the PGA, the assumption is made that the occurrence of the "damage" event is random. First, at the same value of the PGA can correspond several seismic signals, more or less harmful for the structure. Second, the behaviour of the structure may be influenced by disturbing phenomena the quantitative characterization of which is possibly tricky and consequently tainted with uncertainties [Straub & Der Kiureghian 2008]. Hence, the fragility curve, noted Φ in the remainder of this chapter, may be interpreted as the cumulative distribution function of the "structural capability", i.e. the maximum load the structure under investigation can bear without damage.

Fragility curves are useful tools in structural analysis as they provide a more complete information than the usual "failure probability" (established for a reference value of the load only), as discussed by [Schultz *et al.* 2010]. Although their most common application is seismic risk assessment, it is worth noting that fragility curves are used in many other engineering framework, e.g. evaluating the reliability of hydraulic works submitted to extremely high water level (as in the risk assessment studies following the Katrina Hurricane [Ayyub *et al.* 2009]) or window glazings submitted to explosive blast loading [Stewart & Netherton 2008].

The assessment of fragility curves is commonly based on the assumption that the function $\Phi : \mathbb{R}_+ \mapsto [0,1]$ (assumed to be continuous) is the cdf of a log-normal random variable. In the following, we note X the logarithm of the capability.

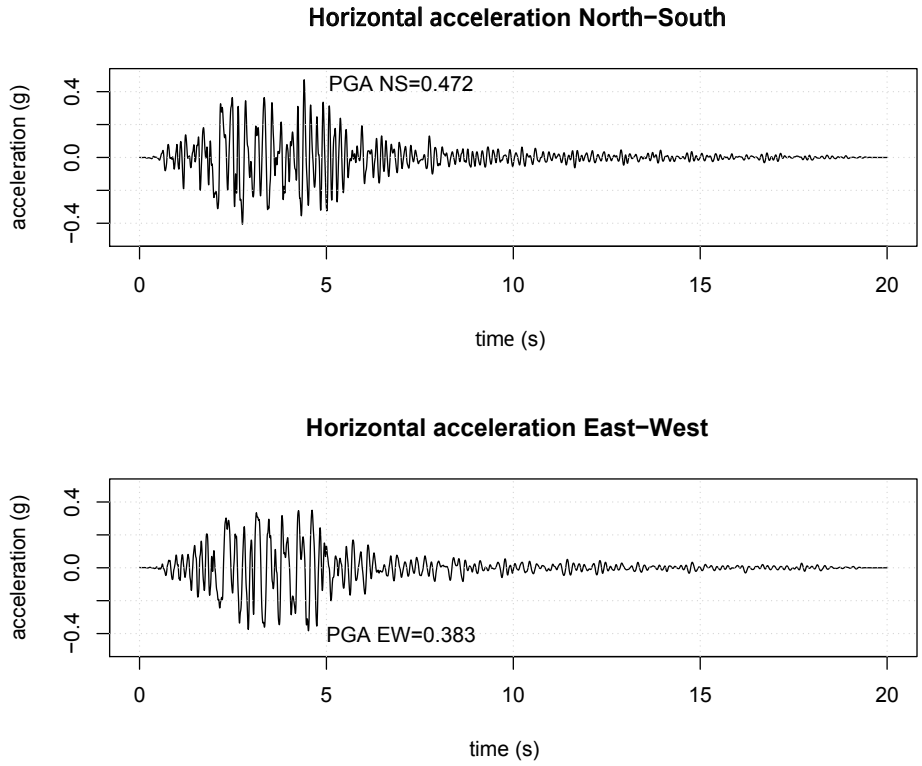


Figure VI.4 – Example of seismic signal and evaluation of the peak ground acceleration (PGA). The present accelerogram was recorded by the measurement station of Nocera Umbra (Italy) during the earthquake of 1997/09/26, 9:40 a.m. (mean PGA 0.427). Horizontal accelerations are measured along the two directions North-South and Est-West and are expressed in g unit (1 g = 9.80 m/s^2). Data source: SISMA (Site of Italian Strong-Motion Accelerograms).

In the standard practice the assessment is made either following an approach entirely based on the expertise, defined by the Electric Power Research Institute (cf. [EPRI 1994] for further details), or by the statistical analysis of actually observed (e.g. [Hoshi *et al.* 2011]) or simulated data. As actual damage data may be scarce due to the rarity of severe earthquakes liable to generate damages on highly safe structures, observations are generated by mock-up or (most often) numerical experiments. A structure is virtually excited by a number of random seismic signals and, a binary indicator is returned as the output of each experiment: 1 if the damage conditions has been reached, 0 otherwise (cf. Figure VI.5).

6.1 Accounting for estimation uncertainties

Under the log-normal assumption, assessing a fragility curve consists in estimating the mean μ and the standard deviation σ, or equivalently the precision $\tau = \sigma^{-2}$, of the Gaussian distribution of the log-capability X from pairwise data $\mathcal{D} = (x_i, y_i)_{i=1,\dots,n}$, where x_i denotes the observed logarithm of the seismic load (here, the PGA) and y_i is the binary indicator of damage (Figure VI.5).

Of course the estimation of the parameters of the fragility curve, noted $\boldsymbol{\theta} = (\mu, \tau)$ in the remainder, is tainted with statistical (epistemic) uncertainties.

The model proposed by [EPRI 1994] separates the aleatory uncertainty tainting the (Gaussian) random variable "log-capability" $X|\mu$ (i.e. with known μ) and the epistemic uncertainty tainting μ, modelled by a log-normal pdf (with log-

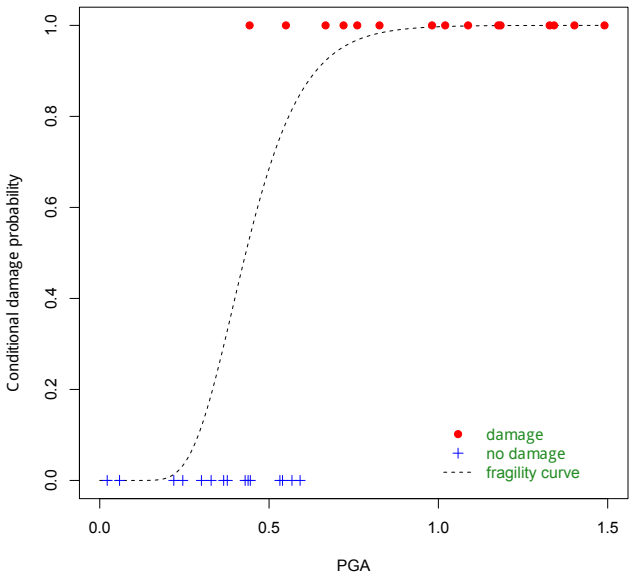


Figure VI.5 – Example of data used for assessing seismic fragility curve, consisting in coupled values of the PGA and of the binary variable indicating if the structure has been damaged or not.

normal mean equal to 1), such as the pdf of X is the product of log-normal distributions.

Rooted in the underlying assumptions of the EPRI model, which roughly speaking let μ bear the epistemic uncertainty of the parameter $\boldsymbol{\theta}$, [Shinozuka *et al.* 2000] proposed to represent the uncertainty tainting the ML estimators of $\boldsymbol{\theta}$ only focusing on the pdf of the estimator $\hat{\mu}_{\mathrm{ML}}$ and suggested the following 90% envelop for the fragility curve, bounded by:

$$\Phi(\cdot|\hat{\mu}_{5\%}, \hat{\tau}_{\mathrm{ML}}) \quad \text{and} \quad \Phi(\cdot|\hat{\mu}_{95\%}, \hat{\tau}_{\mathrm{ML}}),$$

where $\hat{\mu}_{5\%}$ and $\hat{\mu}_{95\%}$ are the 5% and 95% empirical quantiles of a bootstrap sample of the estimator of μ and $\hat{\tau}_{\mathrm{ML}}$ is the ML estimator of τ. The same authors also propose to use, in engineering studies, the curve given by the following expression:

$$\int \Phi(\cdot|\hat{\mu}, \hat{\tau}_{\mathrm{ML}})\, \pi(\hat{\mu})\, d\hat{\mu},$$

where $\pi(\hat{\mu})$ is a normal distribution fitted on a bootstrap sample of $\hat{\mu}$. Of course this modelling, although in common practice proves to be intuitive and easy to handle (because of the multiplicative properties of the log-normal distribution), is questionable and an approach accounting for the uncertainties of both mean and precision of the fragility curve seems more correct from a mathematical viewpoint.

The Bayesian setting is a natural framework for coping with this problem. [Straub & Der Kiureghian 2008] proposed to use the predictive curve:

$$\int\int \Phi(\cdot|\mu, \tau)\, \pi(\mu, \tau|\mathcal{D})\, d\mu\, d\tau, \tag{VI.19}$$

in which $\pi(\mu, \tau|\mathcal{D})$ is the posterior pdf of the parameters.

In the remainder, we go a step forward and apply the fully Bayesian methodology introduced before in this chapter for obtaining "point" estimators of fragility curves. The proposed use case concerns the numerical simulation of a mock-up of a three-storey asymmetric structure (SMART[3] benchmark). The numerical model evaluates, for a given seismic signal,

[3] Acronym of "Seismic design and best-estimate Methods Assessment for Reinforced concrete buildings subjected to Torsion and non-linear effects",

the differential drift between two chosen points of the structure, which is considered as damaged if the drift exceeds a fixed threshold value. 50 evaluations of the deterministic model have been made, under randomly sampling signals, by means of the SALOME-Meca software [Delmas *et al.* 2011], coupling Code_Aster and OpenTURNS (cf. Chapter IV, Section 6) for mechanical and probabilistic simulations respectively within the SALOME platform (cf. Chapter IV, Section 2).

6.2 Bayesian estimation of fragility curves

The methodology follows the four following steps: (i) choosing a prior distribution for Φ i.e., in practice, for $\boldsymbol{\theta} = (\mu, \tau)$ as Φ belongs to the log-normal family, (ii) obtaining a sample of the posterior distribution, (iii) choosing a cost function $C(\tilde{\Phi}, \Phi) : \Xi^2 \mapsto \mathbb{R}_+$ (here Ξ denotes the space of the fragility curves), (iv) choosing as estimator of Φ the curve minimizing the expected cost:

$$\hat{\Phi} = \underset{\tilde{\Phi}}{\text{Argmin}} \int_{\Phi \in \Xi} C\left(\tilde{\Phi}, \Phi\right) \pi(\Phi|\mathcal{D})\, d\Phi \tag{VI.20}$$

Prior distributions of (μ, τ). The following prior modelling, usual when dealing with Gaussian distributions has been chosen:

$$\begin{cases} \pi(\mu|\tau) = \text{Norm}(u, v\tau) \\ \\ \pi(\tau) = \text{Ga}(a, b), \end{cases} \tag{VI.21}$$

that is the so-called Gamma-Normal prior. The advantage of this distribution is that it is conjugated in the ideal case of data consisting in an i.i.d. sample of the normal log-capacities. Of course, it is not the case in the present study, but the problem can be tackled by using a data augmentation technique [Marin & Robert 2007], as in the study presented in Chapter III (page 35).

Posterior sampling. A sample of the posterior pdf can be obtained by means of a Gibbs sampler including a data augmentation step for simulating the latent log-capacities (cf. Appendix, page 159). Figure VI.6 shows the results obtained (for $u = 0$ and $a = b = v = 0.01$) for the SMART test-case.

Cost functions. The log-normal parametric assumption concerning Φ simplifies the mathematical problem as both spaces of the possible decisions and of the states of Nature Ξ are actually subsets of real vectors. They both consist in the space of possible values for $\boldsymbol{\theta} = (\mu, \tau)$, i.e. the subset $\mathbb{R} \times \mathbb{R}_+$. In the remainder we note $\Phi = \Phi_{\boldsymbol{\theta}}$ to highlight the fact that choosing a fragility curve actually consists in choosing a value of $\boldsymbol{\theta}$.

The optimization problem formulated in Equation VI.20 can be rewritten as:

$$\hat{\boldsymbol{\theta}} = \underset{d \in \mathbb{R} \times \mathbb{R}_+}{\text{Argmin}} \int_0^{+\infty} \int_{-\infty}^{+\infty} C\left(\Phi_d, \Phi_{\boldsymbol{\theta}}\right) \pi(\boldsymbol{\theta}|\mathcal{D})\, d\boldsymbol{\theta}. \tag{VI.22}$$

benchmark launched in 2008 by EDF and CEA.

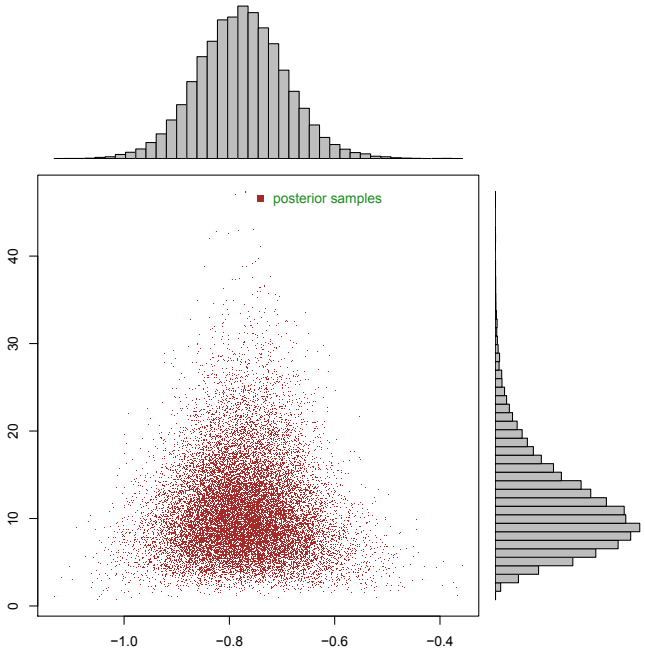


Figure VI.6 – Scatterplot of 20 000 values randomly sampled from the posterior pdf of the parameters μ (x-axys) and τ (y-axis) estimated from 50 evaluations of the SMART model. The histograms of the marginal samples are also represented.

The following cost functions are considered:

$$\text{Gener. quadr. loss: } C_2(\Phi_d,\Phi_{\boldsymbol{\theta}}) = \int_0^{+\infty} \left(\Phi_d(x)-\Phi_{\boldsymbol{\theta}}(x)\right)^2 dp(x), \tag{VI.23}$$

$$\begin{aligned}\text{Gener. abs. loss: } C_{k_1,k_2}(\Phi_d,\Phi_{\boldsymbol{\theta}}) &= \int_0^{+\infty} k_1 \left(\mathbb{1}_{\{\Phi_d(x)\geq\Phi_{\boldsymbol{\theta}}(x)\}}\right)\left|\Phi_d(x)-\Phi_{\boldsymbol{\theta}}(x)\right| dp(x)\\ &+ \int_0^{+\infty} k_2 \left(\mathbb{1}_{\{\Phi_d(x)<\Phi_{\boldsymbol{\theta}}(x)\}}\right)\left|\Phi_d(x)-\Phi_{\boldsymbol{\theta}}(x)\right| dp(x),\end{aligned} \tag{VI.24}$$

in which $p(x)$ is a probability measure of the seismic solicitation. For the exemplary calculations shown hereinafter a uniform probability measure has been chosen. Depending on the context of the study, the analyst can decide to use different measures giving more or less weight to different sets of values of the seismic load (not necessarily the highest values).

The functions VI.23 and VI.24 can be inutitively interpreted as generalizations of the quadratic and piece-wise linear losses previously defined for scalar risk criteria (Equations VI.3 and VI.6 respectively). The positive parameters k_1 and k_2 govern the symmetry of the function C_{k_1,k_2}. It has to be noticed that, as for the scalar case, only the ratio k_1/k_2 is actually significant with respect to the optimization problem to be solved. In the example of results shown hereby different functions have been considered respecting the condition $k_1+k_2=1$.

"Point" estimation of the fragility curve. The Bayes estimator of the fragility curve is the solution of the optimization problem formulated in Equation VI.22. In practice the integrals in the Equations VI.22, VI.23 and VI.24 have been evaluated by means of Monte Carlo simulation. As an example, some results, related to more and more asymmetric

functions are shown in Figure VI.7.

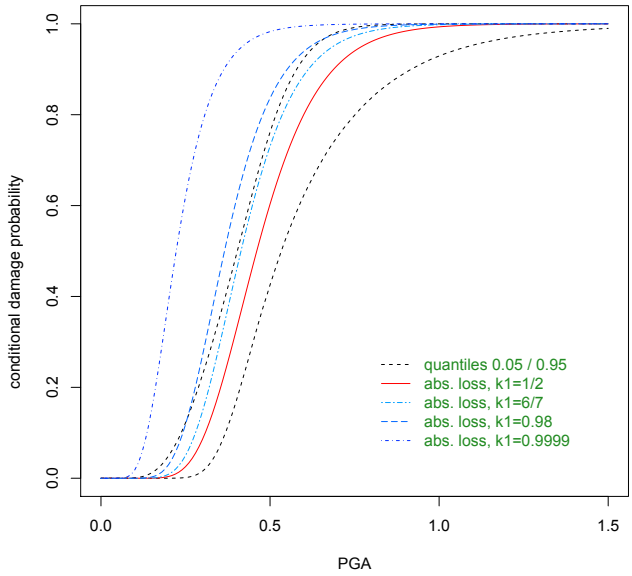


Figure VI.7 – Examples of fragility curves obtained by means of Equation VI.22 from the posterior distribution of (μ, τ) shown in Figure VI.6. Different generalized absolute losses (Equation VI.24) have been used. Notice that $k_1 + k_2 = 1$, that is $k_1 = 1/2$ defines a symmetric cost function and values greater than 1/2 define more and more asymmetric functions, leading to more and more conservative values of the fragility curves (i.e. greater and greater values of the damage probability for the same PGA). In the same graph, 95% credibility bounds for damage probability are also shown.

6.3 On-going work on seismic hazard assessment

Seismic fragility curve is one of the two major inputs of seismic risk assessment. Rather intuitively, the second one is the pdf of the seismic load. These two ingredients allows to evaluate the (unconditional) damage probability of the structure submitted to a random earthquake:

$$P_f = \int_0^{+\infty} \Phi(x)\, p(x)\, dx, \tag{VI.25}$$

where x is the feature of the seismic load, the fragility curve Φ refers to (typically the PGA).

The assessment of the pdf of the PGA (or other relevant quantities related to the seismic load) for a given geographical zone is commonly referred as "seismic hazard assessment". In the remainder, we give some basic ideas and results concerning an on-going work on Bayesian assessment of seismic hazard, carried within the framework of the SIGMA international project[4].

(Very) roughly speaking, seismic hazard assessment involves the following steps:

- the subdivision of the geographical region under investigation (e.g. the South-Eastern quarter of France) in homogeneous seismotectonic zones by means of geological considerations: each zone is characterized by its (constant) "seismicity parameters" ruling the occurrence frequency and the magnitude of the earthquakes,

[4] Acronym of "Seismic Ground Motion Assessment", on-going project gathering EDF, CEA, Areva, ENEL and several academic partners. Cf. http://projet-sigma.com for further information.

- the estimation of seismicity parameters for each zone from historical data available (dates of events and recorded or estimated magnitudes),

- the evaluation, in every point of the geographical region under investigation (which generally stretches out on several zones) of the pdf of the feature of interest of the seismic load (e.g. the PGA). This calculation involves a functional relation, called "attenuation law", between the location and the magnitude of an earthquake on the one hand and the effect on a distant point of interest (where for instance a civil engineering structure is located) on the other hand.

Statistical model for seismicity parameters. The probabilistic and statistical models governing seismic hazard are deeply rooted into seismologist's practice [Gutenberg & Richter 1944, Cornell 1968, McGuire 1976]. A commonly accepted hypothesis within the technical community is that, inside any given zone, the number K of occurrences of earthquakes, the magnitude of which is greater than a given magnitude $M_{\min}$, over a period of length t, is Poisson-distributed. Generally t is taken equal to one year, so that the Poisson parameter λ is to be interpreted as the mean annual number of occurrences (notice that it depends on $M_{\min}$).

The other common assumption is that the magnitude follows a truncated exponential distribution. In other terms, the number of occurrences K and magnitudes M of earthquakes in a given zone are ruled by the following model:

$$\begin{cases} p(k|\lambda) = \dfrac{\lambda^k e^{-\lambda}}{k!} \\ \\ p(m|\beta, M_{\min}, M_{\max}) = \dfrac{\beta e^{-\beta m}}{e^{-\beta M_{\min}} - e^{-\beta M_{\max}}} \cdot \mathbb{1}_{\{m \in [M_{\min}, M_{\max}]\}}. \end{cases} \tag{VI.26}$$

It is worth noting that $M_{\min}$ and $M_{\max}$ are assumed to be known, that is the above model is completely determined by the value of λ and β. The estimation of such a model is, in theory, straightforward if one has at his/her disposal i.i.d. records of earthquakes (dates and magnitudes). In particular, within the Bayesian setting, if one chooses Gamma prior distribution for both λ and β, the posterior distribution of λ is still Gamma distributed (Gamma-Poisson conjugated model). The posterior pdf of β is not conjugate (because of the truncation of the exponential pdf in Equation VI.26) but a posterior sample can be easily obtained (for instance) by Metropolis-Hastings or Importance Sampling.

In practice, inference turns much trickier. Actually, the i.i.d. assumption is met if the magnitudes of all the earthquakes occurred in a given observation period have been observed. But real data are much more heterogeneous: namely, the historical period over which magnitudes can be supposed to have been observed in their entirety differs according to the value of the magnitude itself. Severe earthquakes occurred centuries ago left traces in the regional history and their magnitudes can be a posteriori inferred on the basis of recorded damages, e.g. earthquakes of Lisbon in 1755 (magnitude 8.5-8.7) and Southern Calabria (1783, magnitude 6.9). On the other hand, the memory of ancient low-intensity earthquakes is normally lost. As a matter of fact, the complete observation period depends on the magnitude: the higher the magnitude, the longer the observation period: hence, available data are not i.i.d. with respect to the simple model described above.

To overcome this issue, [Weichert 1980] proposed to subdivide the interval $[M_{\min}, M_{\max}]$ in a number J of intervals of half-length δ and center m_j. Then, by means of expert considerations, to each interval j a time period t_j is associated in which one assumes that all earthquakes have been observed. The Weichert's model (nowadays a reference within the seismologists community) relies on the assumption that the number K_j of earthquakes observed in the period t_j, the

magnitude of which fall in the interval of center m_j is Poisson distributed:

$$K_j \sim \text{Pois}(\lambda\, t_j\, p_j), \quad \text{with} \quad p_j := \mathbb{P}[m_j - \delta \leq m < m_j + \delta] = \frac{e^{-\beta m_j}}{\sum_i e^{-\beta m_i}}. \tag{VI.27}$$

Under the additional assumption that the observed numbers k_j of earthquakes in each interval are independent, the likelihood of the data $\mathcal{D} = \{k_j, t_j\}_{1 \leq j \leq J}$ is given by the following expression:

$$\begin{aligned} \mathcal{L}(\mathcal{D}|\lambda,\beta) &= e^{-\lambda \sum_j t_j p_j} \lambda^{k_{\text{tot}}} \prod_j \frac{(t_j p_j)^{k_j}}{k_j!} \\ &= \exp\left(-\lambda \frac{\sum_j t_j e^{-\beta m_j}}{\sum_j e^{-\beta m_j}}\right) \left(\frac{\lambda}{\sum_j e^{-\beta m_j}}\right)^{k_{\text{tot}}} \prod_j \frac{(t_j e^{-\beta m_j})^{k_j}}{k_j!}, \end{aligned} \tag{VI.28}$$

in which the indices of all sums and products go from 1 to J, and $k_{\text{tot}} = \sum_j k_j$.

Under Gamma priors, posterior samples of (λ, β), according to the expression of the likelihood of Equation VI.28, can be obtained by means of a two-step algorithm. Actually, it can be shown that the conditional posterior distribution $\pi(\lambda|\beta, \mathcal{D})$ has a closed (Gamma) form; thus, the sampling strategy consists in first drawing β (e.g. by means of Importance Sampling), then sampling λ, conditionally on β.

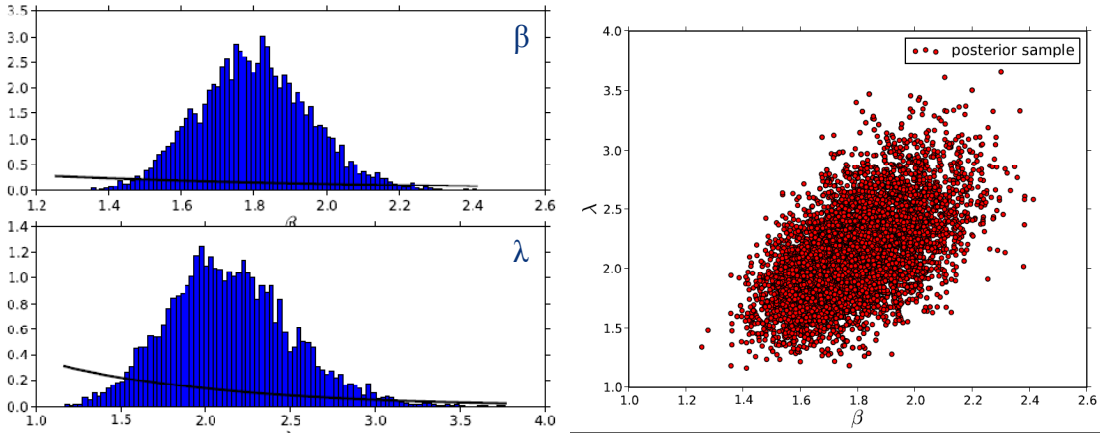


Figure VI.8 – Example of Bayesian estimation of seismicity parameters following the Weichert statistical model. Left: marginal prior (continuous black lines) and posterior distributions (histograms). Right: scatterplot of a Monte Carlo sample of the posterior joint distribution.

Probability distribution of the seismic acceleration. Once seismicity parameters have been assessed, the next (and final) step is to appraise the probability distribution function of the seismic acceleration (e.g. the PGA) in given points of interest of a geographical domain.

The evaluation relies on the probabilistic model sketched hereinbefore, as far as occurrences and magnitudes of earthquakes are concerned, and on a (semi-empirical) relation, called "attenuation law" [Berge-Thierry *et al.* 2003], between the magnitude of an earthquake occurring in a given source point, say A, and the resulting acceleration in the point of interest B (cf. Figure VI.9). Attenuation laws strongly depend on local features of the zone under investigation, and they constitute a topic of active research. Roughly speaking, they can be expressed in the generic form: $a = g(m, l, \upsilon)$, a being the acceleration at point B, m the magnitude of the earthquake occurring at point A, l a given distance (e.g. epicentral or

hypocentral) between A and B, and υ additional local parameters[5].

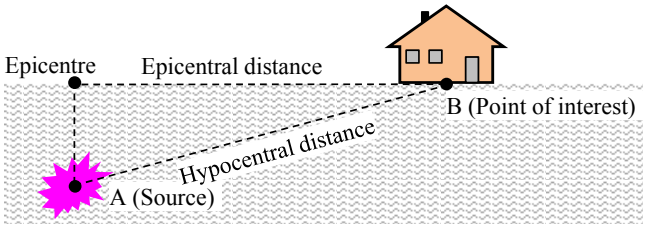


Figure VI.9 – Epicentral and hypocentral distances between a source point A (where the sesmic energy is released) and a point of interest B (where the evaluation of the acceleration is needed).

In practice, one is interested in assessing the probability distribution of the random variable "annual maximum acceleration". Once all the ingredients are available (seismicity parameters and attenuation law), it is straightforward to put in practice a Monte Carlo procedure for simulating, for a high number of (future) years (e.g. 10 000), the annual number of earthquakes, their magnitudes and the accelerations that come with. That allows evaluating the annual maxima of the acceleration for each of the simulated years and thus a sample of the random variable of interest. By means of this sample, the cdf and the complementary cdf of the acceleration (as well as the probability to exceed a fixed level) are estimated by their Monte Carlo estimators.

Figures VI.8 and VI.10 show some exemplary results of this on-going work [Pasanisi *et al.* 2013b]. It is worth specifying that these results have been established for demonstration purposes only and they are not to be extrapolated for drawing any general conclusion about seismic risk assessment carried by EDF and its partners of the SIGMA project.

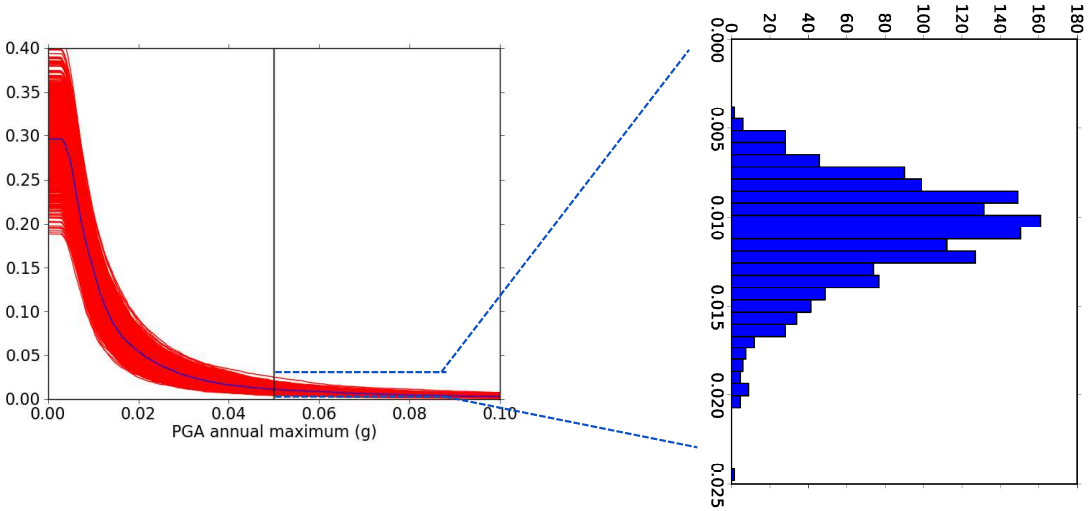


Figure VI.10 – Example of results concerning the assessment of the pdf of the annual maximum PGA. Each of the complementary cdf's shown in the left part correspond to a value of the seismicity parameters (λ, β) sampled in the posterior distribution represented in Figure VI.8. The blue line corresponds to the posterior mean. The acceleration have been evaluated using the Berge-Thierry attenuation law. In the right part of the figure, it is shown the posterior distribution of the probability for the annual maximum PGA to be greater than a threshold value, arbitrarily fixed at 0.05 g.

[5] In particular, for the exemplary calculations shown here, the popular relation: $\log_{10}(\text{PGA}) = a \cdot m - \log_{10}(b \cdot l) + c$, in which, l is the hypocentral distance (known as the Berge-Thierry law) has been used.

Chapter VII

An application of multi-criteria decision-aid

Que tu aies foutu le camp de chez nous en 1945, me dit-il, je te comprends [...] Mais tu sais bien que maintenant tout est changé et que rien n'est plus comme avant [..]
Bien sûr, la France est un beau pays et Paris une très belle ville, mais tout cela fait partie d'un monde qui s'ecroule, de l'Occident pourri [...] Et voilà qu'au lieu de respirer librement, comme citoyen libre en une société libre, tu choisis de faire le taxi dans les embouteillage de Paris ? Je ne te comprends pas.
François Suliny, Le piéton de Stalingrad (1975)

That you got out from us in 1945, he said me, I understand you [...] But you know well that everything has changed now and nothing is as before anymore [...]
Of course, France is a beautiful country and Paris a wonderful city, but it's all part of a world that is collapsing, the rotten West [...] And now, instead of breathing freely, as a free citizen in a free society, you choose to be taxi driver in Paris traffic jams? I do not understand you.
François Suliny, The infantryman of Stalingrad (personal translation)

Reading notes

Technical context. The works presented in this chapter have been made during my first period at EDF R&D (Energy in Buildings and Territories Dept.) from 2004 to 2008. The first decade of the 2000's has been a very strategic moment for the company (and more widely for all the French and European energy sector): passage of EDF to the status of stock company in 2004, opening of the capital in 2005, institution of the energy savings "*white certificates*", total opening of the market of the energy in 2007 ... In this context, EDF wished to diversify sale offers for its individual customers by more and more positioning itself as supplier of energies (with a final "s", in particular: electricity and natural gas) and services.

The technical problems I have been concerned with, within this framework, have been: (i) the evaluation of thermal systems using different sources of energy for covering heating and air-conditioning needs of buildings and (ii) the definition and the evaluation of different solutions for improving the energy efficiency of existing dwellings. In both problems, the appraisal and ranking of possible alternatives are to be made according to different (and possibly antagonistic) criteria. That motivates the use of specific multi-criteria decision-aid methods: the quite classical ranking according to the life-cycle cost as well as the methods of the ELECTRE family, the use of which has been rather innovative in that context.

Contribution. Most of the work carried in this period has been delivered in internal reports, most of which in restricted access, because of the underlying commercial stakes. That concerns particularly works on multi-energy systems. Nevertheless, in this summary it is worth mentioning the development of a MatLab® based tool (PADME, French acronym

of "Program for multi-energy decision-aid") [Pasanisi 2006], as well as studies involving building thermal dynamic simulation for the estimation of energy needs [Pasanisi & Sivadier 2004] and multi-criteria ranking of alternative thermal systems to cover these needs ([Pasanisi 2004b], [Pasanisi & Coince 2006], [Pasanisi & Évin 2007]), as well as a pedagogical summary report on multi-criteria decision-aid method [Pasanisi 2005].

The more perspective works concerning energy retrofitting of existing buildings have been more widely diffused within the technical and scientific community [Pasanisi & Ojalvo 2007, Pasanisi & Ojalvo 2008]. A software tool, named REFLEX (cf. Section 3), has been developed within a partnership with the University of Liège (cf. also [Pasanisi 2007]).

Finally, it is worth noting that these innovative ideas contributed to stimulate the reflection about new families of tools supporting more and more tailored energy services [Pasanisi & Bieret 2007], [Pasanisi & Heijmans 2007], [Pasanisi 2008a], [Le Mouel *et al.* 2007].

Structure of the chapter and credits. This chapter is basically made of two parts. The first one, mostly adapted from [Pasanisi 2005], is an introduction to multi-criteria decision-aid methods. The second one, focused on the problem of energy retrofitting is adapted (and partly excerpted) from the article [Pasanisi & Ojalvo 2008], published in the journal *Foundations of Computing and Decision Sciences*.

1 Decision making: a difficult process

Following the definition of [Roy 1996]: *"decision aiding is the activity of the person who, through the use of explicit but not necessarily formalized models, helps obtain element of responses to the questions posed by a stakeholder of a decision process. These elements work towards clarifying the decision [...]"* with the goal of increasing *"the consistency between the evolution of the process and this stakeholder's objectives and value system."*

Notice that, according to this definition, the analyst does not aim necessarily to recommend the best decision, but simply the one that fits, in the best way possible, the requirement and the value system of the decision maker.

Most of decisions are difficult to made: from decisions concerning national safety or economics to most common decisions as choosing between two job proposals or which car to buy, these problems share two common issues (cf. also Chapter I, page 4).

- The presence of uncertainties tainting the outcome of the decision to be made or the action to be taken. For instance [Parent *et al.* 2014], when designing a dike in order to mitigate flood risk, on the one hand the decision maker must compare (more or less) certain costs (the building costs of the protection work) to very uncertain possible consequences (costs generated by rare but possible extreme water level) and, on the other hand, whatever the choice of the protection work, he/she will never be 100% sure that any damage will never occur in the future.

- The fact that, normally, there exists no alternative that is better than all the others, according to all the criteria that are at play in the analysis: one wants to buy a car which is ideally comfortable, easy-to-drive, beautiful, fuel-efficient and cheap at the same time ... and which, of course, does not exist! The problem is qualified as "multi-criteria" because the decision must be made under the basis of multiple and antagonistic criteria and the choice necessarily implies to make compromises.

Statistical decision theory essentially addresses the problem of making decision in an uncertain context. It is one of the pillar of the Bayesian analysis, as shown in the examples of the Chapter VI. In the remainder, we deal with other families of methods, specifically intended to cope with the antagonistic nature of the criteria involved in the decision process.

Many taxonomies of multi-criteria decision-aid (MCDA) methodologies exist. We particularly insist here on two great families of methods: (i) the ones aiming at transforming a multi-criteria problem in a single-criterion one by means of the definition of a single synthesising criterion, (ii) the methods based on pairwise comparisons of different actions, using the so-called "outranking" concept.

For a more complete introduction, cf. for instance the books [Vincke 1992], [Roy & Bouyssou 1993, Roy 1996], [Pomerol & Barba-Romero 2000], as well as the monograph of [Dodgson *et al.* 2009] more specifically intended to general public.

Introducing some formalism. In what follows, we consider the analysis of a finite set of possible alternatives $\mathcal{A} = \{a_i\}_{i=1,\dots,m}$. We assume that the analyst has at his/her disposal a set of n criteria obeying to some mild conditions (exhaustive, coherent, non-redundant). Formally a criterion is a function $g_j(\cdot)$ mapping the space of possible alternatives $\mathcal{A}$ to $\mathbb{R}$, such that it is possible to establish a preference relationship between two different alternatives, say a_1 and a_2 by comparing the real number $g_j(a_1)$ and $g_j(a_2)$. The performances of the alternatives are summarized in the so-called "performances matrix":

$$\begin{pmatrix} g_1(a_1) & g_2(a_1) & \dots & g_n(a_1) \\ \dots & & & \\ g_1(a_m) & g_2(a_m) & \dots & g_n(a_m) \end{pmatrix}. \tag{VII.1}$$

The assumption is made that no alternative is "dominated" by any other which, in practice, means that no action is over-performed by any other according to all criteria. More formally, assuming without loss of generality that criteria are "increasing" measures of the performances of an action (the higher $g_j(a_i)$, the better a_i), there exist no (a_i, a_k) such that:

$$g_j(a_i) \leq g_j(a_k) \;\; \forall j \;\text{ and }\; \exists l; g_l(a_i) < g_l(a_k). \tag{VII.2}$$

The vector $\mathbf{g}(a_i) = (g_1(a_i), \dots, g_n(a_i))$ maps each alternative into a point of $\mathbb{R}^n$, which is its representation in the so-called "performances space".

Starting from these assumptions, the different multi-criteria methodologies basically differ in the way the performances are aggregated to establish a final recommendation, which, depending on the methods can be: (i) identifying the best alternative, (ii) assigning them to a number of predefined classes or (iii) establishing a ranking (these three problems are often referred α-, β- and γ-type, respectively).

A very brief introduction is given in the following sections.

1.1 Methods based on a single synthesising criterion

The keystone of this family of methods is the definition of a function $V : \mathbb{R}^n \mapsto \mathbb{R}$ who transforms the criteria vector of each alternative a_i: $\mathbf{g}(a_i) = (g_1(a_i), \dots, g_n(a_i))$ into a single criterion:

$$V(a_i) = V\left(g_1(a_i), \dots, g_n(a_i)\right).$$

These methods are firmly rooted into the so-called "Multi-Attribute Utility Theory" (MAUT). A rich literature involving axiomatic considerations, conditions the function V must obey to (as well as elicitation methods) is available (e.g. [Keeney & Raiffa 1993]).

The simplest form for the function V is the additive one. One can think, for instance, to the weighted sum of the

criteria:

$$V(a_i) = \sum_{j=1}^{n} w_j g_j(a_i).$$

We do not discuss here the several interesting theoretical and practical problems concerning this class of methods. We limit ourselves to two particular methods which prove useful in engineering practice, and have been used in particular in the technical domain (comparing energy efficiency solutions) the remainder of this chapter is concerned with.

Distance from an ideal solution. The idea underlying this family of methods is that the synthesising criterion summarizing the "value" of the alternative a_i is the distance (in the performances space) between the point of coordinates $(g_1(a_i),\ldots,g_n(a_i))$ and an ideal point, representative of a (not-existing) action $a^\star$ (dominating all the others), the coordinates of which are the components of the vector:

$$\mathbf{g}(a^\star) = \left(\max_i g_1(a_i),\ldots,\max_i g_n(a_i)\right), \quad i = 1,...,m.$$

Let us also define an anti-ideal point $a_\star$, represented by the vector:

$$\mathbf{g}(a_\star) = \left(\min_i g_1(a_i),\ldots,\min_i g_n(a_i)\right), \quad i = 1,...,m.$$

The most commonly used distance functions for this problem are derived by the expression of the Minkowski distance:

$$V(a_i) = \left(\sum_{j=1}^{n} w_j \left|\frac{g_j(a_i) - g_j(a^\star)}{g(a_\star) - g(a^\star)}\right|^{\omega}\right)^{1/\omega}. \tag{VII.3}$$

In the expression above, the w_j's act as weights and the value of ω rules the compensations between different criteria: for $\omega \to \infty$ the distance above tends to the Chebyshev distance and the best alternative is the one minimizing the maximum of the differences between actual and ideal actions (minimax problem).

This method, known as "*Compromise Programming*" [Zeleny 1973], is very popular among the engineers for its intuitive representation and its easy implementation. It is very commonly applied since four decades to many different technical fields, for instance (restricting ourselves to some recently published works among many others): the operation strategy of distributed energy resources (power and heating) [Ren *et al.* 2010], the expansion of drinking water urban infrastructures [Chang *et al.* 2012] the management of waste produced by livestock operation [Gebrezgabher *et al.* 2014].

Life-cycle cost. Life-cycle costing is a technique, especially used in civil engineering design, to evaluate whether a system meets the client's financial requirements or not. Life-cycle cost (LCC) is defined as the *"cost of an asset or its parts throughout its life-cycle, while fulfilling the performance requirements"* [ISO 2008].

It can be also employed as a global financial criterion to rank different alternatives, particularly in cases when options presenting lower initial capital costs have also higher running cost (and/or, possibly, higher end-of-life costs).

Basically, the idea underlying the LCC is to sum, for a given alternative a_i, the initial (acquisition, construction) and end-of-life cost, noted C_{0_i} and E_i respectively, and the discounted running costs $K_{l_i}(t)$, cumulated over a reference period of analysis T, e.g. the expected lifetime of the system:

$$V(a_i) = LCC(a_i) = C_{0_i} + \sum_{t=1}^{T}\sum_{l_i=0}^{r}\left[\frac{K_{l_i}(t)}{(1+d_{l_i})^t}\right] + \frac{E_i}{(1+d_e)^T}. \tag{VII.4}$$

Of course, the formula above should be adapted on a case-by-case basis, depending on the particular problem to be solved. However, it is worth noting the relevant role played by the value of T: the higher the period of analysis is, the more low running costs could balance high initial costs.

This methods is often used for comparing energy performances of buildings and/or HVAC (heating, ventilation, air-conditioning) systems.

Cf. [Uygunoğlu & Keçebaş 2011, Marszal *et al.* 2012, Zhu *et al.* 2012, Ng Cheng Hin & Zmeureanu 2014] as some recent examples of the numerous works in this specific application field. In these cases, running costs are maintenance and energy costs (e.g. electricity and natural gas bills) and the evaluation must account, in the best way possible, for the evolution of energy prices, which proved increasingly difficult as the considered period of observation increases.

Other criteria similar to the LCC, as the return-on-investment or the internal-rate-of-return can be proposed as alternatives of LCC. The main difficulty is that strictly financial criteria, undoubtedly useful as "normative" objective references, may fail to reproduce the value system of a decision maker when non-financial criteria (e.g. environmental, aesthetic, comfort etc.) are supposed to play a relevant role.

1.2 Methods based on the outranking

The criteria aggregation procedures of this family of methods are not based on the comparison of the different alternatives according to a single criterion, but instead on several criterion-by-criterion pairwise comparisons, which are finally processed, by means of different algorithms, to provide a global recommendation among the alternatives of the set $\mathcal{A}$. Intuitively, these techniques are akin to vote procedures: each criterion is entitled to one vote either for or against the considered alternative and the final judgement accounts for the number and the importance (criteria are weighted) of the expressed preferences.

This methods have been developed since the end of the 60's thanks to the seminal works of Bernard Roy [Roy 1968] and its colleagues of the so-called "French-speaking School". The basic principles of some of them are sketched below. For a particularly pedagogical presentation, cf. the monograph of [Maystre *et al.* 1994] from which the following short introduction is inspired.

ELECTRE. This popular method[1], first proposed by [Roy 1968] and renamed "ELECTRE I" to distinguish it from its variants, is based on a two-steps procedure. First, for each pair of alternatives (a_i, a_k), are determined the criteria for which the performance of a_i is better, equal or worse respectively than the one of a_k. Hence, three sets of criteria are determined, the weight of which (in the final recommendation) is evaluated as:

$$\begin{cases} w^{+}_{i,k} = \sum_{j=1}^{n} w_j \, \mathbb{1}_{\{g_j(a_i) > g_j(a_k)\}} \\ w^{=}_{i,k} = \sum_{j=1}^{n} w_j \, \mathbb{1}_{\{g_j(a_i) = g_j(a_k)\}} \\ w^{-}_{i,k} = \sum_{j=1}^{n} w_j \, \mathbb{1}_{\{g_j(a_i) < g_j(a_k)\}}. \end{cases} \tag{VII.5}$$

These weights are used for evaluating, for each pair (a_i, a_k), two indices expressing the "global agreement" and dis-

[1] ELECTRE is the acronym of "*Élimination et choix traduisant la réalité*" (Elimination and choices translating reality).

agreement about the assertion "a_i outranks a_k":

$$\begin{cases} \text{Concordance index: } c_{i,k} = \dfrac{w^+_{i,k} + w^=_{i,k}}{w^+_{i,k} + w^=_{i,k} + w^-_{i,k}} \\ \text{Discordance index: } d_{i,k} = \dfrac{1}{\max\limits_{i,k,j}\left(g_j(a_k) - g_j(a_i)\right)} \left[\max\limits_j \left(g_j(a_k) - g_j(a_i)\right)\right] \mathbb{1}_{\{w^-_{i,k}>0\}}. \end{cases} \tag{VII.6}$$

The alternative a_i is considered to outrank a_k if the concordance (and discordance) index is greater (lower, respectively) than a given threshold, i.e. $c_{i,k} \geq c^\star$ and, $d_{i,k} \leq d^\star$, that is, following the analogy with vote procedures, the criteria speaking in favour of a_i are the majority, but the most discordant opinion has not too strong arguments. The procedure has to to be repeated for each pair of alternative.

The second step of the method allows to obtain a final judgement about the overall set of alternatives, from the pairwise outranking relations. That can be done by representing the outranking relations as a graph in which the arrow directed from a_i to a_k means that a_i outranks a_k.

The set of the best solutions is represented by the set of nodes $\mathbb{B}$ such that (i) each node of the graph not belonging to $\mathbb{B}$ is outranked by at least one node of $\mathbb{B}$ and (ii) the nodes of $\mathbb{B}$ are not outranked by any other. In cases where the outranking graph shows cycles, the nodes of the cycles are aggregated into a single node, to let $\mathbb{B}$ exist.

As a conclusion, this method highlights either the best alternative or a number of "best alternatives" which are equivalent or difficultly compared between them.

ELECTRE III. First proposed at the end of the 70's [Roy 1978], this is one of the most popular between the several variants of the ELECTRE method. The principle remains the same: (i) first, establish outranking relations, (ii) then, using these relations to make a final recommendation which is, in this case, a ranking of the alternatives.

With respect to the ELECTRE methodology sketched above, the outranking relations are made under the basis of a finer comparison between the performances.

The concordance and discordance indices of the statement "a_i outranks a_k" are evaluated for each criterion by means of functions (graphically shown in Figure VII.1) involving the three thresholds q_j, p_j and v_j (with $q_j < p_j < v_j$). We note $c_j(a_i, a_k)$ and $d_j(a_i, a_k)$ respectively, the indices related to the criterion $g_j(\cdot)$.

As one can see, the credibility of the outranking statement depends on the "opposite" comparison a_k vs. a_i. Let us first consider the concordance index $c_j(a_i, a_k)$: it is 1 if $g_j(a_k) < g_j(a_i)$, it starts decreasing when the difference $g_j(a_k) - g_j(a_i)$ exceeds an "indifference threshold" q_j and it is null once this difference is significantly high, i.e. greater than a "preference threshold" p_j.

The reasoning concerning the discordance index is similar: it is null unless $g(a_k)$ is "well below" $g(a_i)$, and 1 once the difference $g_j(a_k) - g_j(a_i)$ exceeds a "veto" parameter v_j, beyond which the statement "a_i outranks a_k" becomes meaningless.

The concordance and discordance indices are used to evaluate a "degree of fulfilment" $\delta_{i,k}$ of the outranking relation. Let us first define the (global) concordance index $C_{i,k}$:

$$C_{i,k} = \frac{\sum\limits_{j=1}^{n} w_j c_j(a_i, a_k)}{\sum\limits_{j=1}^{n} w_j}.$$

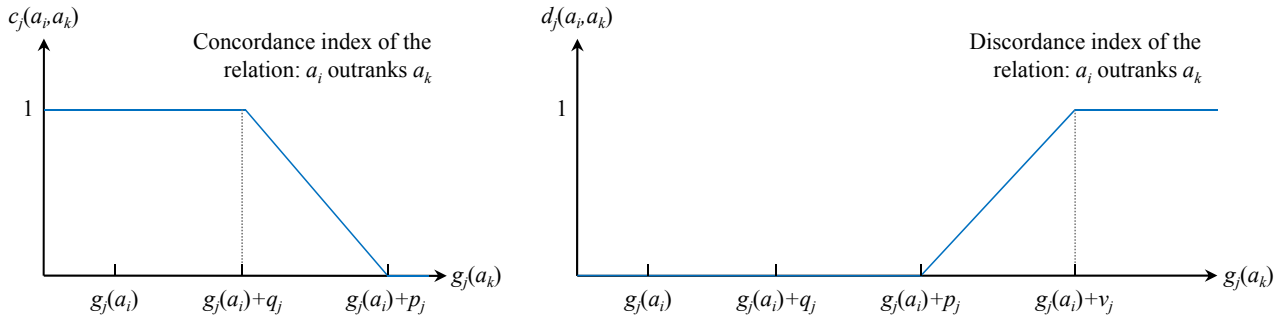


Figure VII.1 – ELECTRE III concordance and discordance indices of the relation "a_i outranks a_k", related to the criterion $g_j(\cdot)$. One can see the role played by the parameters q_j, p_j and v_j.

The expression of $\delta_{i,k}$ is:

$$\delta_{i,k} = \begin{cases} C_{i,k} \text{ if } C_{i,k} \geq d_j(a_i,a_k) \ \forall\, j \\ \\ C_{i,k} \prod\limits_{l \in \mathcal{F}_{i,k}} \dfrac{1-d_j(a_i,a_k)}{1-C_{i,k}} \text{ otherwise,} \end{cases} \tag{VII.7}$$

in which $\mathcal{F}_{i,k} = \{j; d_j(a_i,a_k) > C_{i,k}\}$ denotes the set of the "considerably discordant" criteria for which the discordance index is greater than the global concordance index.

It is worth noting that the condition $g_j(a_k) - g_j(a_i) > v_j$ for a single criterion automatically leads to the rejection of the outranking relation, no matter the weight of the criterion w_j. The result of these pairwise comparison is a set of m^2 degrees of fulfilment of the outranking relation ($m \times m$ fulfilment degrees matrix). The exploitation of these "fuzzy" outranking relations, in order to obtain a final ranking of the proposals, is made by means of a (non trivial) algorithm, called "distillation" algorithm that we do not detail here (cf. [Roy & Bouyssou 1993, Maystre *et al.* 1994] for a full description).

PROMETHEE. The PROMETHEE method[2] [Brans *et al.* 1984] shares with ELECTRE the main ides of exploiting pairwise outranking relation.

Let us consider the pair (a_i, a_k). In the PROMETHEE method to each criterion $g_j(\cdot)$ is assigned a "preference function" $f : \mathbb{R} \mapsto [0,1]$ the argument of which is the difference $\Delta_j(a_i,a_k) = g_j(a_i) - g_j(a_k)$ between the performances of a_i and a_k. Intuitively, the value of the function $f_j(\Delta_j(a_i,a_k))$ measures the credibility of the statement "a_i outranks a_k" over the usual scale from 0 (not credible) to 1 (extremely credible). The choice of the functions belongs to the analyst and strongly depends on the problem to be solved. They are normally defined by one or two parameters. Figure VII.2 shows the most common functions used by the PROMETHEE methods.

The value of the preference functions are aggregated into their weighted sum, called "preference index":

$$P_{i,k} = \frac{\sum\limits_{j=1}^{n} w_j f_j(\Delta_j(a_i,a_k))}{\sum\limits_{j=1}^{n} w_j}.$$

Finally, for obtaining a ranking of the alternatives, three quantities, named "flows", for each alternative a_i are evaluated:

[2] Again a name borrowed from Greek mythology, but actually PROMETHEE is the acronym of "*Preference ranking organization method for enrichment evaluations*."

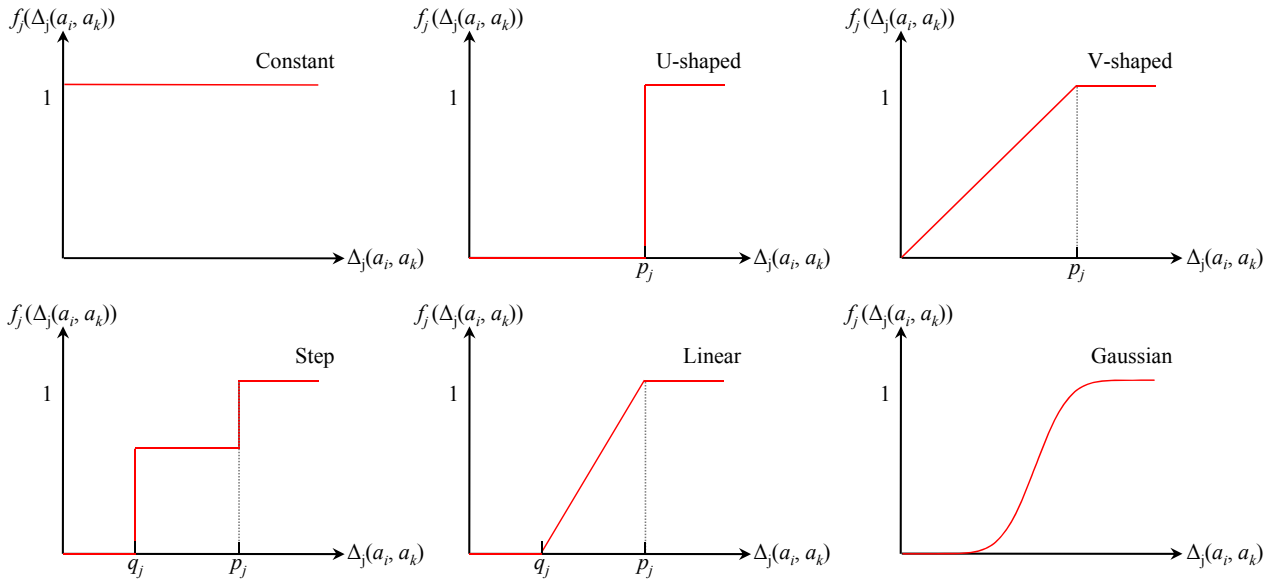


Figure VII.2 – PROMETHEE most common preference functions.

$$\begin{cases} \text{Positive outranking flow: } \Psi_i^+ = \sum_{k=1}^{m} P_{i,k} \\ \text{Negative outranking flow: } \Psi_i^- = \sum_{k=1}^{m} P_{k,i} \\ \text{Net outranking flow: } \Psi_i^= = \Psi_i^+ - \Psi_i^-. \end{cases} \tag{VII.8}$$

Two variants of the method exist, named PROMETHEE I and II. They differ in the way the flows are exploited.

In PROMETHEE I:

- a_i is preferred to a_k if: $\begin{cases} \Psi_i^+ > \Psi_k^+ \text{ and } \Psi_i^- \leq \Psi_k^- \text{ or} \\ \Psi_i^+ = \Psi_k^+ \text{ and } \Psi_i^- < \Psi_k^-. \end{cases}$
- a_i and a_k are indifferent if: $\Psi_i^+ = \Psi_k^+$ and $\Psi_i^- = \Psi_k^-$,
- a_i and a_k are incomparable, otherwise.

In PROMETHEE II:

- a_i is preferred to a_k if: $\Psi_i^= > \Psi_k^=$,
- a_i and a_k are indifferent if: $\Psi_i^= = \Psi_k^=$.

At this stage, the algorithm did not give a ranking yet ... but just binary preference relations. That could be seen a richer results than a simple "ranking", because it highlights incomparable alternatives, i.e. alternatives which are not indifferent (because of their great difference) but they cannot be ranked between them by the decision maker.

In the end, the final ranking of the alternative is established under the basis of the preferences relations determined hereinbefore: rank 1 is given to alternatives such that none is preferred to (even if they are not comparable with many others), rank 2 is given to alternatives such that only rank 1 alternatives are preferred to, and so on.

2 Energy retrofitting of existing building: a specific and complex problem

Existing buildings represent today significant and rather attainable sources of energy savings. Driven by the new framework of the European energy market and aware of its environmental and social stakes (growing scarcity of fossil energies, energy independence, global warming ...), EDF has become a major player in the domain of energy efficiency (EE) in buildings. In addition, the total opening of the market of energy (July 2007 in France) drove (and continues driving) energy suppliers to more and more diversify the services offered to their customers.

The REFLEX (Effective Retrofitting of an Existing Building) tool, sketched hereinafter, has been developed in this technical context. It allows to define, appraise and rank the best energy retrofitting scenarios for a given existing dwelling (in particular, it is focused on individual houses).

The energy retrofitting of an existing dwelling is a complex process resulting from the combination of several actions (walls, roof, floor insulation, setting up more efficient heating and ventilation systems, use of renewable energies etc.). Therefore, the number of alternatives to be examined is generally huge. The great number of solutions is not the only characteristic of the problem. The best way to define this decision-aid problem could be: "How to propose to a customer the best retrofitting scenarios taking into account his/her constraints (budget above all) and his/her requirements?". In other terms, the objective is here to help a customer to spend a given budget in the best way possible and to reach, at the same time, high energy efficiency targets.

The first thing to be noticed is the central role of the customer: finally, he/she decides what will be made in his/her house (that may seem trivial but it is not useless to remind it). Consequently, the proposed solutions have to be coherent with his/her preferences and objectives. The problem is complex because customer's choice is done on the basis of several criteria which are generally in conflict between them, for example the cost of the retrofitting action, the reduction of energy consumption, the inconvenience caused by the works, thermal comfort etc.

Thus, it is a multi-criteria problem for which there is not a strictly speaking optimal solution but a set of solutions which are more or less in conformity with a set of identified objectives. The ranking of the possible alternatives must be done by means of suitable methods which aim more at managing the conflicts between the different objectives (cf. Section 1), rather than seeking a mathematical optimum. Specific multi-criteria decision-aid (MCDA) procedures (Compromise Programming, ELECTRE, PROMETHEE) seem, in this context, more powerful tools than techniques based on a strictly financial unique criterion, like return-on-investment time for instance, which do not always fit the preference system of a residential customer .

If the central role of the customer is indisputable, he/she is not the only player in this problem. The proposed alternatives must fit also the interest of the energy supplier. From its point of view, the problem is also multi-criteria as for instance energy savings and energy sells are antagonistic criteria. And, of course, the supplier must take into account the preferences of the customer, as finally it's him/her to decide!

MCDA methods have often been used to define the best scenarios to improve the energy efficiency of new or existing buildings. Many case studies, methodology papers and tools descriptions can be find in the technical and scientific literature concerning either a given single component of the building (e.g. the heating system) or the whole building, as in our approach. As an example, [Blondeau *et al.* 2002] used a customized version of the ELECTRE methods to determine the most appropriate ventilation strategy in an university building between a given number of predetermined scenarios. In a slightly more theoretical framework, the optimal shape of a building which minimizes a multi-attribute utility function of building costs and heating consumptions is analysed by [Marks 1997]. Later on, [Jedrzejuk & Marks 2002] used a similar approach to optimize the external configuration and the internal partition of a block of flats.

[Azar & Hauglustaine 2001] developed a powerful tool named ACME (French acronym of "Aid for the multiple cri-

teria conception of the building envelope") which finds the best option for the envelope of a new building using a multi-objective genetic algorithm. The optimized envelope gives at the same time the best possible performances and the most equal satisfaction for a given number of players involved in the decision process (e.g. the client, the architect, the public authority). Each decision maker is described by a given set of criteria weights and PROMETHEE-like preference functions (Figure VII.2). Another application in buildings design is given by [Wang *et al.* 2004] (life-cycle cost vs. life-cycle environmental impact).

The MCDM-23 tool [Balcomb & Curtner 2000], developed in the framework of the Solar Heating and Cooling Program of the International Energy Agency, helps decision makers to specify and prioritize criteria in large buildings design and to compare between them a given number of design schemes using scoring techniques and star-diagrams. Nevertheless, it does not generate nor evaluate the solutions to be compared.

Some examples can be found in the specific framework of energy retrofitting of existing buildings. A case study of application of the ELECTRE III method to three real-life retrofitting projects of office buildings in Switzerland is presented by [Rey 2004]. The BIDS (Building Investment Decision Support) case-based tool [Loftness & Hartkopf 2002] estimates the benefits of a given number of retrofitting actions for improving the performances (and between them energy efficiency) of a commercial building using aggregated financial criteria: the economic-value-added (EVA) and the return-on-investment (ROI). The estimations are made by comparing the examined building to the results of retrofitting actions previously made on similar buildings. Within the ORME (Office Rating Methodology) framework [Roulet *et al.* 2002], a multi-criteria method for rating existing office buildings (according to their costs, energy consumptions, environmental impacts and indoor environment quality) and for ranking retrofit scenarios has been proposed.

The "twin" EU projects EPIQR (for large residential buildings) [Jaggs & Palmer 2000] and TOBUS (for office buildings) [Caccavelli & Gügerli 2002] are both based on a three-stages methodology to generate retrofit scenarios. First, a diagnosis of the general state of the structural and functional components of the building (around 50-60) is made with respect to deterioration, functional obsolescence, energy consumption and indoor environmental quality. Then, a list of actions on each elements is proposed and evaluated, and finally the effects of the proposed actions are visualized (but not ranked) in synthetic graphs in order to help the architect to define on his/her own the "best" retrofit scenario by aggregating the retained actions. An interesting analysis of the utilization of these tools in the professional practice (with a positive feedback) is given by [Flourentzou & Roulet 2002].

[Flory-Celini 2008] investigated the problem of the bioclimatic retrofitting of existing buildings. The principle of bioclimatic architecture is to take advantage in the best way possible of natural phenomena involving the building environment and materials (thermal inertia, permeability, micro climate, natural ventilation, hours of sunshine, orientation, greenhouse effect) to enhance the energy efficiency. The developed methodology involves the use of techniques for the design of numerical experiments (namely, Hadamard, Rechtschaffner or factorial design) in order to preselect the efficient bioclimatic actions to be retained and an ad hoc variant of the REFLEX tool (REFLEX-Bio) to build, appraise and rank the solutions.

Among the existing tools, the main original feature of REFLEX is to completely link all the different steps of a retrofitting project in an unified framework: (i) the description of the existing building, (ii) the generation of possible retrofitting scenarios, (iii) the fine evaluation of the performances of the alternatives and finally (iv) their multi-criteria analysis by taking into account the point of view of the two players involved in the decision process: the customer and the energy supplier. An overview of REFLEX is given in the next sections.

3 REFLEX Overview

In a glance, the REFLEX tool is composed of the following modules, described hereby:

- a solutions generation module which, starting from the description of the existing house, the wishes and the constraints expressed by the customer identifies the possible retrofitting solutions;
- a solutions evaluation module which evaluates the performances of the retrofitting solutions according to the criteria considered in the decision-making process;
- an interactive multi-criteria filter which allows the decision-makers to rule out the solutions whose performances according to one or more criteria are considered unsatisfactory;
- a multi-criteria analysis module which ranks the retained solution according to the points of view of the customer and the energy supplier.

3.1 Retrofitting solutions generation

The solutions generation module is used at the same time to collect data concerning the dwelling and the customer and to design the retrofitting solutions to be evaluated and analysed. This module has been developed within the framework of a scientific partnership between EDF R&D and the University of Liège (LAPT, Department of Architecture).

To avoid any confusion, we first precise the terminology used hereby. A dwelling is divided into a given number of *components* whose performances can be improved by a retrofitting intervention. In particular, the eight components sketched in Figure VII.3 are distinguished. An *intervention* is a technical action on a given component, for example: "Replacement of the existing boiler by a new one". Each intervention has several *variants*. For instance, in the case of the boiler, the variant can be defined by its type (condensing or not) and the fuel used (natural gas, oil, LPG, wood-pellets etc.). Finally, a *solution* is a combination of variants of interventions, for example: "outside walls insulation by 12 cm fibreglass blanket and replacement of the old boiler by a new natural gas condensing boiler".

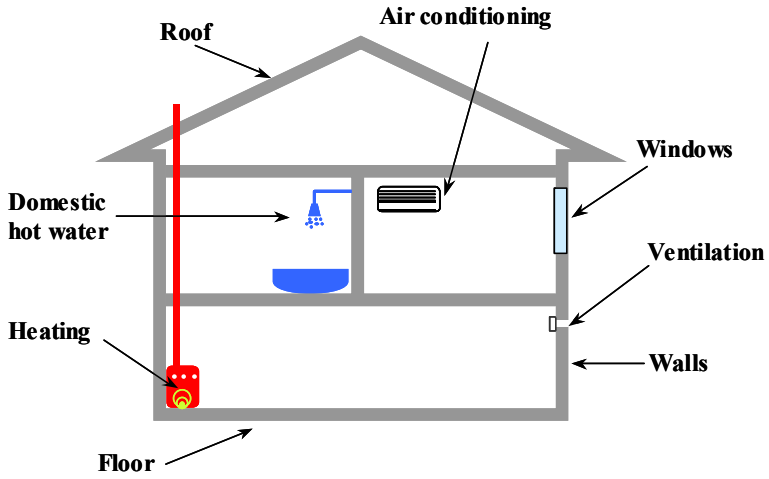


Figure VII.3 – Components of a building the energy performance of which can improved by means of retrofitting interventions.

The generation of the solutions is made on the basis of a detailed description of the existing house. Two different data capture modes are available in a graphical user interface (GUI): a "simplified form" and an "expert form." In the simplified

mode, the user enters the requirements and constraints of the customer (expected reduction of the energy consumption, maximal budget, unwished systems and materials, preserving the appearance of the front side etc.) and a minimal number of pieces of information concerning the house under investigation: between them, the dwelling type, age, number of floors, floor area, energy used for the heating and for domestic hot water (DHW). By switching to the expert mode, the user is able to manage a greater amount of data, concerning the general environment of the dwelling (urban or rural location, weather data ...), the occupants (occupancy and heating/cooling scenarios ...) and the different components to be retrofitted. Data inputting is done directly on decision trees: each answer redirects the user towards another (following) question. When no more information is necessary, the tool shows all possible interventions on the considered component.

These ideas can be illustrated using the example sketched in Figure VII.4 concerning the roof retrofitting. Here, the description of the existing roof is made in seven steps: 1) number of roofs, 2) roof type, 3) roof insulation thickness, 4) attic state, 5) roofing materials, 6) underlayment, and finally 7) Retrofitting interventions proposal. At the end of the "branch" the intervention codified as TOI04 ("Installation of an outside thermal insulation between the rafters, of a new underlayment, re-installation of the existing roofing materials beforehand put down") is proposed. It has to be noticed that when switching into expert mode, all expert data are pre-inputted. The simplified mode data are automatically converted into expert data by REFLEX and the user is not obliged to scroll all decision trees to proceed in his case study. If needed, he/she can just verify pre-inputted data or modify a part of them.

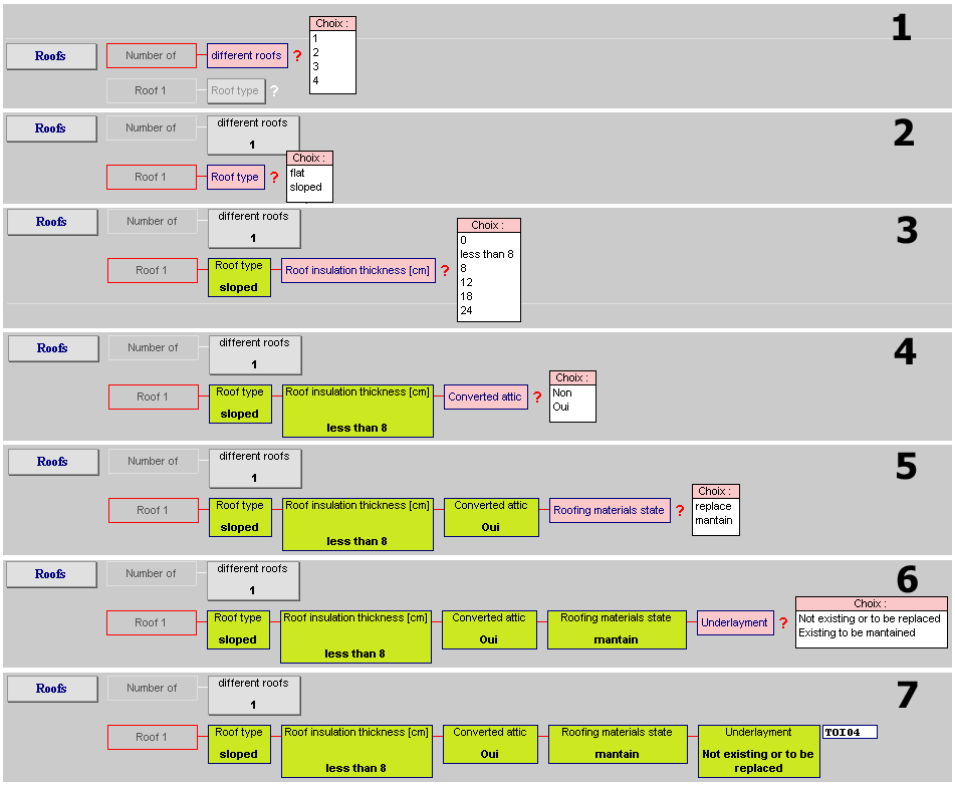


Figure VII.4 – Input data concerning a component to be retrofitted (here, the roof) directly inserted on the decision tree of the expert system. The different steps of the inputting procedure are shown by means of successive screen captures of REFLEX, numbered from 1 to 7.

The next step in the solutions generation is the combination of interventions. The problem is tricky as the naive combination of the variants of the interventions proposed for each component (and between them the alternative "no

intervention" has to be considered) gives generally raise to an infeasible number of possible solutions.

In a similar context, [Kaklaukas *et al.* 2005] proposed a method based on nested multi-criteria analyses for coping with the problem of the great amount of possible scenarios which can be obtained by combining all possible variants of retrofitting actions. A first set of multi-criteria analyses defines the best variants for the retrofitting of each element of the building. The set of criteria used for each analysis is different for each of the components to be examined. Then, retained variants (for example the best four) are combined obtaining thus the retrofitting scenarios to be ranked.

The REFLEX solutions generation algorithm "smartly" explores the space of the possible combinations of variants by using two constraints imposed by the customer: the maximal budget and the required improvement of the energy efficiency. When analysing the combinations of a number of given interventions, the algorithm first focuses on the combination including the cheapest variants. If the obtained solution exceeds the customer's budget, then all possible combinations (including different variants) are a priori excluded, without being explored, as they are all more expensive than the customer's budget and thus unaffordable.

A similar filter is set up by using the constraint relating to the improvement of the energy efficiency: if the combination of the most efficient variants does not allow to obtain the required performances, then all other combinations are excluded. This algorithm proved quite effective: the exploration of a very high number of combinations, which can easily reach several millions, is done in a reasonable computing time. Each generated solution is described by a set of files which are the main inputs of the evaluation module.

3.2 Retrofitting solutions evaluation

The solutions generated at the previous stage of the process are all consistent with respect to the customer's situation and constraints from three points of view: (i) technical (they are applicable to the customer's actual dwelling), (ii) financial (their cost complies with the maximum budget of the customer), (iii) energy efficiency (their potential of energy saving complies with the customer's expectations).

The solutions' generation is essentially based on a technical "expert system" approach, as described above. To go forward in the decision-making process, REFLEX users shall leave the technical space of the retrofitting solutions and move the space of their performances.

As REFLEX considers two players, the customer and the energy supplier, the evaluation module considers two different sets of criteria. For confidentiality reasons, no details about EDF criteria set are given here. The point of view of the customer is described by the following criteria (grouped by types) :

- financial performance: investment cost, cumulated financial gain;
- energy: annual energy savings;
- comfort: thermal comfort indices (winter and summer), Indoor Air Quality (IAQ) index, ease-of-use;
- environment: reduction of CO_2 emissions, environmental-friendliness of materials and systems;
- miscellaneous: renovation work inconveniences.

A total of fourteen criteria are calculated: ten refer to the customer, four refer to the energy supplier. The criteria assessment process includes two main stages:

- assessment of the hourly energy needs and consumptions as well as indoor temperatures of the dwelling over a complete "typical" year (i.e. under reference weather conditions);

- criteria calculation, part of them using the results above.

To do this, the evaluation module includes a dynamic thermal simulation core (the EDF R&D *SimFast* software [Déqué *et al.* 1999]). The benefit of using a specific dynamic tool is to increase the quality of the evaluation of thermal performances: each simulation is tailored to the specific features of the customer and the dwelling: dwelling type and size, thermal properties of the envelope, technical characteristics of the heating system, occupancy patterns, weather conditions, thermal inertia of the building.

As far as financial criteria are concerned, the estimation of investment costs is based on common cost ratios applied to the dwelling features. A wide range of energy tariffs are implemented, allowing a relevant assessment of the bill savings.

Two thermal comfort indices are calculated, one for the heating season and one for the rest of the year. They are based on an analysis of the indoor temperature curve, involving the PMV (Predicted Mean Vote) model [ISO 2006] and the adaptive approach to thermal comfort.

Some criteria, e.g. environmental-friendliness or work inconveniences, are qualitative and not quantitatively related to the energy consumption and indoor temperature of the dwelling. They are calculated against a number of input variables by means of scoring tables.

Each criterion is implemented in a specific module. This brings a high level of flexibility of the software tool. A new criterion can be easily introduced by just adding the corresponding evaluation module. Any existing module can be replaced by a more efficient one, if needed, without affecting the others.

In the next (and last) step of the process, the retrofitting solutions are represented by a vector of performances including 14 values. These ones are all summarized in a performances matrix (cf. page 133) which has as many rows as the number of solutions, and as many columns as the number of criteria.

3.3 Retrofitting solutions' filtering and ranking

Once proposed scenarios have been evaluated, a graphical user interface (GUI) gives a representation of them in the criteria space. Each point in the graph represents a given alternative: blue points if only the building's envelope (walls, roofs, etc.) has been improved, green points if the solution concerns only the thermal systems (heating, cooling, DHW) and red ones if the solution concerns both (envelope and systems). The user can select the criteria that he/she wishes to put on the x and y axes of the graph. In the example of Figure VII.5, the cost of the operation is on the x-axis and the reduction of the energy consumption on the y-axis. By setting the cursors in the left part of the GUI, it is possible to define some constraints on the performances of the solutions and to rule out the ones who do not satisfy these constraints (in the dark grey zone of the GUI).

The retained solutions are ranked using MCDA techniques. The user can choose between four different methods: Compromise Programming (weighted distance from an ideal alternative), ELECTRE III, PROMETHEE I and II. Two different rankings are obtained using the two sets of criteria described above: one from the customer's point of view and another from the energy supplier's point of view.

The results of MCDA calculations are shown in a graph (Figure VII.6) where each solution is represented by a point the coordinates of which are the ranks in the two rankings: the best alternatives for the customer are in the lower part of the graph, the best alternatives for the energy supplier are in the left part of the graph and solutions close to the origin are well ranked in both rankings.

By a "right-click" on a point, the user can obtain a detailed description of the corresponding alternative. It is also possible to show the performances of a given solution and to find out on the graph the solutions including a given technology. For instance, in the graph of Figure VII.6, an arrow highlights solutions including a condensing boiler.

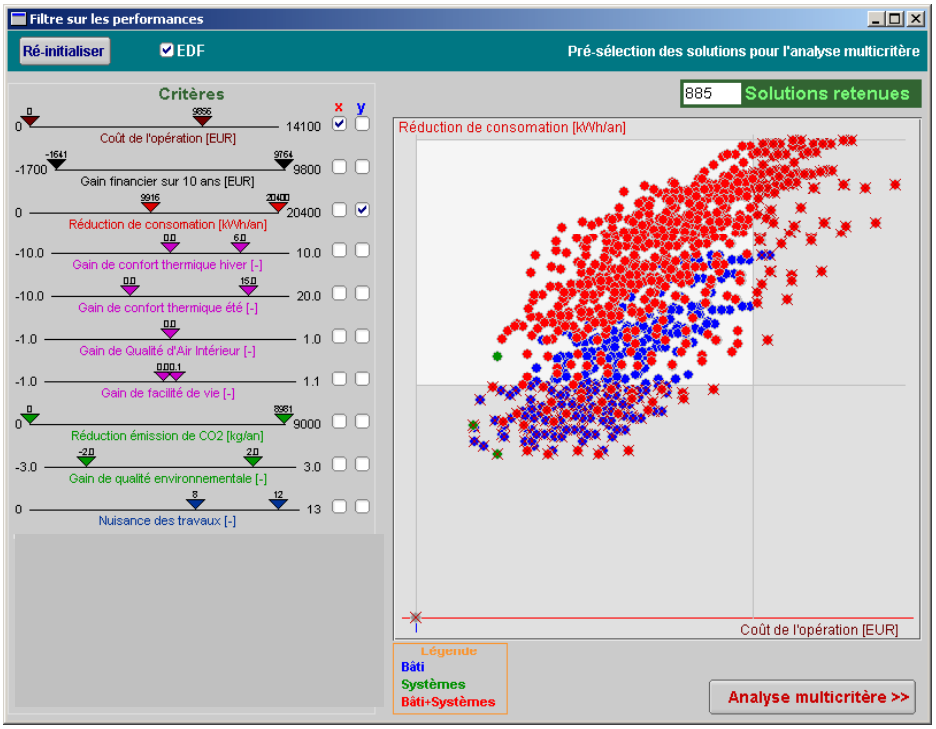


Figure VII.5 – Example of application of REFLEX to the retrofitting of a two-bedrooms house: multi-criteria representation and filtering of the proposed solutions.

The example featured in Figures VII.5 and VII.6 must be regarded as a simple test, the only purpose of which is to show the outputs provided by the REFLEX tool. It concerns the retrofitting of a two-bedrooms mid-terrace house, built in the 50's, located in the Paris Area. The retrofitting solutions have been ranked using ELECTRE III method. The weights and thresholds of criteria used for this example reflect the point of view of a customer mainly concerned with financial and thermal comfort criteria. They are shown in Table VII.1.

In this example, the best solutions are obtained combining the replacement of the existing oil boiler by a new natural gas condensing boiler and the thermal insulation of the roof and/or the front wall.

Criterion	Unit	q	p	v	weight (%)
Investment cost	€	500	1500	5000	19
Cumulated gain	€	10%	50%	100%	15
Annual energy savings	kWh	10%	30%	150%	5
Winter thermal comfort	Index between 0 and 100	10	30	100	15
Summer therm. comfort	Index between 0 and 100	10	30	100	15
Home friendliness	Index between 0 and 3	0	1	3	4
Indoor Air Quality	Index between 0 and 3	0	1	3	4
Annual CO_2 savings	kg of CO_2	10%	30%	200%	9
Environment-friendly construction	Index between 0 and 10	0	5	10	7
Work inconveniences	Index between 0 and 10	0	5	10	7

Tableau VII.1 – Weights and thresholds of the customer's criteria, used in the example shown in Figures VII.5 and VII.6.

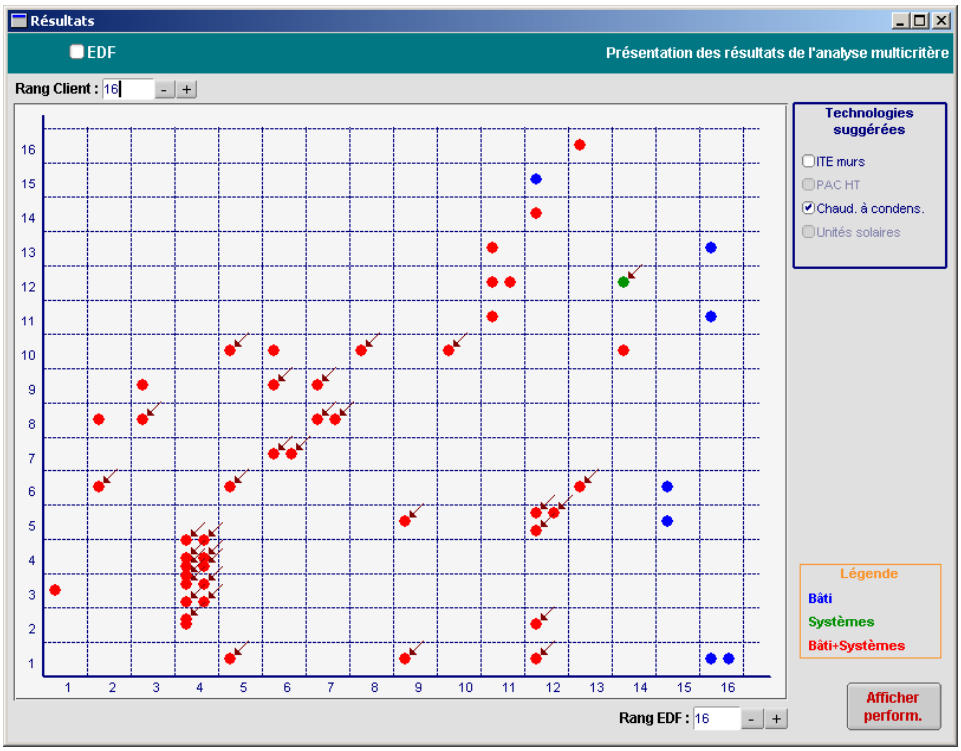


Figure VII.6 – Example of application of REFLEX to the retrofitting of a two-bedrooms house: ranks of retained solutions.

4 Discussion and perspectives

REFLEX is the putting together of an expert system, encoding technical expertise of building retrofitting, of a dynamic energy simulation tool and of MCDA methods. From this point of view it is a comprehensive tool which aims at fully tackling the problem in all its complexity. The most interesting points of the REFLEX approach are:

- the possibility to explore the whole solutions' space;
- the accounting for both the "multiple criteria" and the "multiple decision makers" dimensions of the problem;
- the separation of the technical expertise and the computer code. The catalogue of retrofitting interventions and the decision trees are described in specific data bases (XML format) which are used by a "generic algorithm" generating the retrofitting alternatives. That means, for instance, that a new intervention can be added, without modifying the computer code.

This last point and its modular architecture allows REFLEX to be easily adapted to other particular frameworks (like apartment buildings or office buildings retrofitting), as witnessed in particular by the experience of the REFLEX-Bio [Flory-Celini 2008] tool.

REFLEX has been essentially conceived for internal R&D purposes. In a more long-term view, REFLEX can be viewed as the archetype of a new family of tools to be used for providing more and more tailored energy efficiency advices to EDF customers.

Chapter VIII

Afterword

La strada che non ha strisce
Sarà la rotta sotto questa luna
Coi suoi problemi, coi suoi compromessi
E che ogni volta non ritrovi mai la stessa.
Liftiba, Lacio Drom - Buon Viaggio, 1995

The road that has no stripes
Will be the route under this moon
With its problems and its compromises
So that each time you never find the same.
Litfiba, Lacio Drom - Good trip (personal translation)

1 Concluding remarks

At the end of this manuscript, which is definitely a kind of summary of this first part of my career, it is interesting to take a final quick look at the main teachings of these last fifteen years.

I started my career in a very technical engineering environment. The interest (and actually the passion) for the applied research came a little later. Initially motivated by the quest for solutions to problems demanding more powerful tools than the one I had in my engineer's toolbox, I had the opportunity to discover (again, like when I was a student) the beauty of learning and being challenged with difficult (at last for me) mathematical and intellectual problems, the joy of understanding ... but also the feeling of vertigo, when confronted with the number of methods, tools, problems still to be appraised and learnt. In this sense, research learnt me, on the one hand, that most of the problems which seemingly have no solution can definitely be tackled and, on the other hand, to be humble as I quickly found that nobody ever knows enough about a topic, a methods, a tool.

I also had the great opportunity to work in different countries, in different environments (consulting, production engineering, technical support to field personnel, academic labs, industrial R&D) and in different engineering domains of application (coastal and environmental modelling, aerospace, water distribution, energy efficiency, building simulation, nuclear production).

All of these experiences learnt me more and more about my work of engineer (and about myself). I am aware of the fact that, in spite of my research works, the main added-value of my contributions is that they stand on the frontier between research and engineering, frontier that, depending on the topic (and also on the very personal idea that one has), is sometimes clear but most of time extremely fuzzy.

Figure VIII.1 gives a quantitative sketch of my technical and scientific production, in terms of papers of different nature. As one can see, the most intense period of this first part of my technical and scientific career started in 2008-2009. The opportunity to work in a very scientific and methodological framework and to manage a transverse project, with a great number of different applications as well as the management of PhD and post-doc fellows and research partnerships also gave a significant momentum to my production.

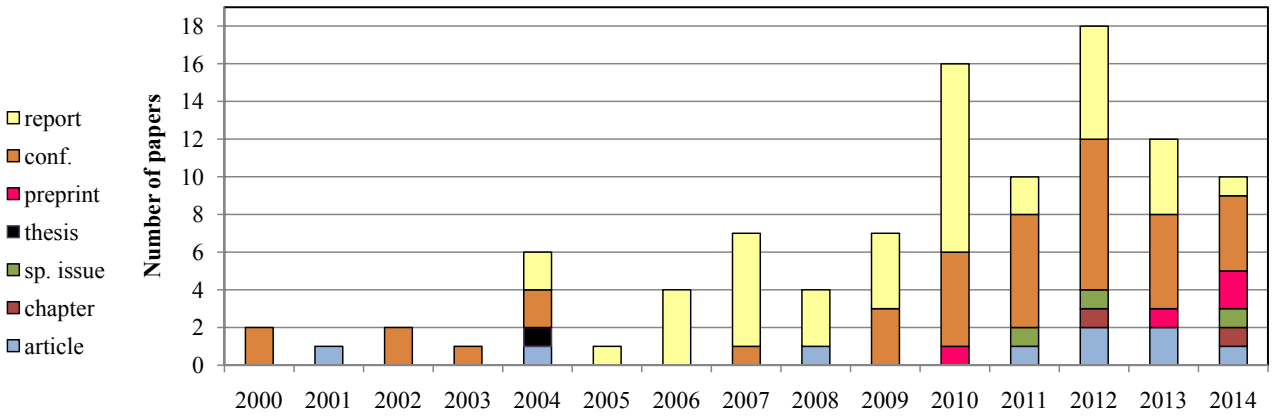


Figure VIII.1 – Summary of the scientific and technical production in terms of "papers" of different nature, written from 2000 to the present day. Most of the "public" production has been accomplished from 2008, while the period 2005-2008 has been most dedicated to internal EDF R&D contributions. **Important remark:** the counting of technical reports here is rather exhaustive while only technical reports most significant with respect to the technical and scientific topic of rest of the manuscript are listed in my personal references list at page 167.

The topics I have been concerned with are (extremely) varied. I made a fun exercise, for highlighting this variety, by building a "word-cloud" (a popular graphical representation of texts, especially used for sketching the content of websites) from the text of this manuscript (cf. Figure VIII.2).

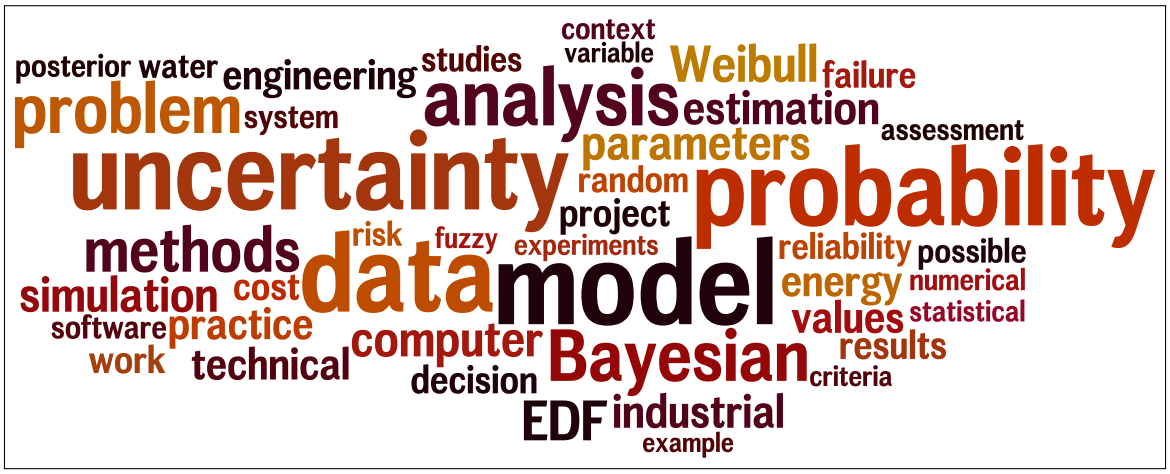


Figure VIII.2 – Word-cloud of the text of the present manuscript (after little adaptations to avoid representing "common words") built thanks to the Wordle™ tool (`www.wordle.net`), showing the most frequent words used.

Although the principle of the analysis, based only on words occurring frequency, is somehow simplistic, a quick look at the word-cloud shows that it actually succeeds in putting into evidence the main topics this document (and my work,

so far) have been concerned with. Most of my activities consisted in making engineering studies by means of modelling and/or simulation, in presence of uncertainties.

Keywords referring to the main methods and tools used are also easily found: in particular the Bayesian analysis is definitely the most powerful tool I used for coping with this different problems.

2 On-going works and perspectives

As a conclusion, some perspectives of the works presented hereinbefore are given. Of course, these perspectives reflect a rather personal viewpoint and, taking into account (in particular) the very large extent of the scientific domain of "computer experiments", they do not pretend to be exhaustive.

As far as discrete lifetime models are concerned, as also reminded at the end of Chapter II, it should be interesting to complete the proposed analysis (restricted to Inverse Pólya and Weibull-1 distributions only) to other probabilistic models. In particular, variants based on more advanced Pólya's urn schemes, in which the number of added balls at the end of a trial depends on the number of trials already made, appears an interesting perspective for giving more credit to a model which, in spite of some handling difficulties, is extremely intuitive and thus potentially useful in engineering.

In addition to the refinements of the proposed algorithm by improving the efficiency of the MCMC procedure and the estimation of the copula within it, the methodology concerning the estimation of Markov models, sketched in Chapter III, can be interestingly extended to more complex cases: one can think, for instance, at time-dependent transition matrices or time-dependent data collection mechanisms. Industrial reliability, but also biomedical survey or ecological simulation are likely to provide interesting use-cases for motivating new methodological works on this rather general class of problems.

In the very wide scientific domain of computer experiments many perspectives and areas of further works can be imagined.

The most important is in my opinion the extension of the common methodological framework sketched in Chapter IV (where uncertainty analysis is essentially seen as the propagation of "parametric" uncertainties through a black-box code) to take systematically into account, in standard practice, the uncertainty tainting the model itself. The current vision, inherited from the industrial practice of "structural reliability", implicitly admits that the engineer who uses a numerical code for risk analysis purposes considers the code as a representation of reality, the approximation of which is sufficient in view of the purposes of the study. In other words, while admitting the importance of the process of verification and validation (V&V) of a code, this step is often considered as a prerequisite, to be achieved (separately) prior to any study of uncertainty or reliability analysis [1].

Nowadays, the technical community is moving towards a more integrated approach in which the steps of verification, validation, uncertainty propagation form a single unified framework, often named VVUQ (Verification, Validation, Uncertainty Quantification) . Several definitions of "*verification*" and "*validation*" exist, but they all refer to the same idea: verification aims at ensuring the proper implementation in mathematical terms of the phenomenon or system under investigation, while validation is focused on the agreement between physical reality and results (predictions of that reality)

[1] Cf. the following sentence excerpted from [Aven 2010]: "*The issue we raise here is how to deal with this 'error' [model inaccuracy], should we quantify it? No is our clear answer. It is not meaningful to quantify the model inaccuracy. The point we make is that if the model is not considered good enough for its purpose, it should be improved. The uncertainty assessments are based on the model used. Of course, when observations of Z are available, we would compare the assessments of Z, which are conditional on the use of the model G, with these observations. The result of such a comparison provides a basis for improving the model and accepting it for use. But at a certain stage we accept the model and apply it for comparing options and making judgments about for example risk acceptance (tolerability). Then it has no meaning in quantifying the model inaccuracy.*
The results are conditional on the model used. Instead of specifying $\mathbb{P}[Z \leq z]$ *directly we compute* $\mathbb{P}[G(X) \leq z|K]$ *and G is a part of the background knowledge K*."
NB. Here, the difference is made between the actual variable of interest Z and its forecast $G(X)$ by means of the computer model G.

provided by the code. If the verification task is primarily the responsibility of the teams that implement and provide the code, validation is to be be seen as a statistical problem: after all, the keystone of validation is the comparison between physical and numerical experiments in order to quantify the error that analysts make when replacing reality with the computer model (model error).

This problem, in practice, shares the same underlying skills, together with a number of mathematical tools (in particular metamodelling and Monte Carlo methods), with the uncertainty propagation. In the framework of an on-going PhD program (in cooperation with AgroParisTech) we investigate the popular framework for model calibration and validation of [Kennedy & O'Hagan 2001]. More precisely, let $\{y_i(x_i)\}$ be a sample of data observed under given experimental conditions x_i (controlled variables) tainted with measurement error ε_i. Let $z(x,u)$ be a computer code used for reproducing the same physical phenomenon; in this writing are distinguished, among the inputs, the controlled variables and other "parameters" of the experience, out of the operator's control (e.g. the value of an intrinsic physical property of the system under investigation). The statistical model is noted:

$$y_i(x_i) = z(x_i, u^\star) + b(x_i) + \varepsilon_i, \tag{VIII.1}$$

in which $u^\star$ is the value of u for which the model fits the data "in the best way possible". In this writing, the term $b(x_i)$ is a bias, modelling the gap between model and reality. The problem here is twofold: the calibration of the numerical code (i.e. estimating $u^\star$) and the estimation of the bias (model error) $b(\cdot)$. Both $z(\cdot,\cdot)$ and $b(\cdot)$ are supposed to be realizations of Gaussian processes.

Inspired from recent researches in this domain [Le Gratiet 2013, Bachoc 2013], the work (currently in progress) aims at analysing the feasibility of this approach in an industrial framework, in which experimental data are generally scarce and costly and the numerical code is possible CPU time consuming [Damblin *et al.* 2013b]. In particular, the focus is put on the problems of non-identifiability of the error carried by the imperfect knowledge of $u^\star$ and the one carried by the bias $b(\cdot)$. Several use-cases has already been imagined for this work and in particular a problem of building thermal simulation [Damblin *et al.* 2014a] with an interesting underlying decision problem, namely the trust an energy provider, like EDF, can put in energy consumption forecasts, provided to its customers.

Another part of this work is concerned with the problem of enhancing the prediction quality of the kriging metamodel which, in practice, replaces the actual code in the algorithmic implementation of the methodology sketched above, in the case of complex (and time consuming) computer model. In this framework, a sequential method for the design of computer experiments for solving the calibration problem, inspired by the popular EGO (Efficient Global Optimization) algorithm of [Jones *et al.* 1998] and the SUR (Stepwise Uncertainty Reduction) criterion of [Bect *et al.* 2012] is currently investigated [Damblin *et al.* 2014b].

More generally, in many other cases concerned with metamodelling, building designs of computer experiments, adapted to the problem to be tackled, for instance intended to well approximating the actual model in specific regions of interest [Picheny *et al.* 2010, Bect *et al.* 2012] (particularly useful when estimating a failure probability or a quantile), or respecting good space-filling properties [Roustant *et al.* 2010, Pronzato & Müller 2012, Damblin *et al.* 2013a] is a very promising field of investigation.

From a more technological viewpoint, the increasing of computing power is an interesting opportunity in particular for metamodelling, as distributed designs of experiments (DOE) seem to be ideal targets and uses cases for HPC (high performance computing). Among the recommendations provided by the on-going EESI2 project (European Exascale Software Initiative, 2012-2015) for the future EU activities in the domain of intensive computer simulation [Erbacci *et al.* 2013], some specific points are dedicated to design of numerical experiments.

In particular, has been highlighted the importance of (i) taking into account DOE methods when developing future middleware[2] (a good flexibility in terms of easily switching from large number of small jobs to small number of large jobs should be targeted in order to make easier the use of generic software tools implementing the DOE), of (ii) managing the problem of the resilience of software tools to calculation failures (rerunning points in DOE that have not completed, better distinguish between cases that failed for numerical reasons or HPC-infrastructure related reasons etc.) and of (iii) managing in the best way possible the multiple levels of parallelism that can be involved in multi-physics codes and combining them with the "external" parallel level imposed by the software tool ruling the DOE.

Other extensions of the usual practice of uncertainty analysis are also to be mentioned as promising areas for further methodological and engineering works in order to more and more effectively cope with (i) the problems of functional inputs [Iooss & Ribatet 2009, Goffart *et al.* 2013, Fruth *et al.* 2014] or outputs [Auder *et al.* 2012] as well as (ii) with computer models returning stochastic outputs [Marrel *et al.* 2012, Douard & Iooss 2013] or with a "chaotic" behaviour (i.e. extremely sensitive to some parameters, which can be, in addition, tainted with significant uncertainty) as it can be the case in computational fluid dynamics [Camy 2013].

To conclude this (non-exhaustive) personal list of perspectives on computer experiments, it is worth mentioning the effort that has to be made on the visualization of "uncertain" results in numerical simulation. Actually, common visualization techniques may fail to represent the uncertainty tainting 2D or 3D (or possibly 4D if one includes time) uncertain outputs of complex computer models because, on the one hand, "*our visual channels can be overwhelmed when increasing the amount and the dimensionality of the data*" and, on the other hand, "*when moving from quantified uncertainty to visualized uncertainty, we often simplify the uncertainty to make it fit into the available visual representation*" [Potter *et al.* 2012]. Cf. also [Popelin & Iooss 2013] for an example of more specific issues concerning uncertainty analysis of thermal-hydraulic transients in nuclear engineering. As a matter of fact, the problem is how to represent at the same time the results and their uncertainty avoiding excessive simplifications but also without saturating the human perception capability. It is a challenging and promising area of research, as witnessed, for instance, by the works (at the frontier between numerical simulation, computer science, human physiology) of [Coninx 2012].

As far as extra-probabilistic frameworks for uncertainty analysis are concerned, an important challenge for further works is enhancing the effectiveness of uncertainties propagation. Actually, a limitation for the practical use of these *alternative* settings consists in the computational cost of the propagation of the α-cuts of input variables (cf. Chapter V, page 99) in case of non-monotonic codes. The use of mathematical methods and tools specifically intended for interval propagation [Jaulin *et al.* 2001] in this hybrid framework could be an interesting perspective. Moreover, following ahead the argument often presented in favour of extra-probabilistic settings introducing them as less-informative and more intuitive than probability, one can also think at interval analysis as another interesting alternative in engineering studies (cf. for instance the recent works of [Merheb 2013] in the field of thermal simulation of building's components).

The decision-making framework for building point estimators of safety criteria could be extended to account, in a more general way than the one presented in Chapter VI (page 120), also for the computer model uncertainty. The case study described in [Damblin *et al.* 2014a] ("guaranteed" forecast of energy consumption of a dwelling) is expected to provide interesting methodological results. Nevertheless, the computational cost of this full-Bayesian methodology is definitely an issue ... that can be tackled by means of metamodelling; but in this case, metamodelling error should also be taken into account (or at least it has to be checked that it does not introduce a significant bias in the estimation).

Another interesting avenue for further investigations, inspired from Chapter VI, concerns the elicitation of cost functions tailored to the case of safety studies. Here, as already sketched in the previous chapters, the "optimal" value can

[2] A "middleware" is a software specifically intended for enabling communication between applications and managing data in a distributed computing environment.

be hardly imagined as the one minimizing a financial cost. Although the simple cost functions evoked in Chapter VI (quadratic and weighted absolute loss) are rooted in common sense and provide useful arguments for justifying the conservatism of results in mathematical terms, further investigations could be carried to propose more adapted cost functions.

Finally, as far as multi-criteria decision-aid methods are concerned, although sensitivity analysis is a traditionally consolidated step in studies involving these families of methods, it should be interesting to explore how the standard practice could take more benefit of the recent dramatic improvements made in the domain of uncertainty and sensitivity analysis of computer experiments; after all, given the performance matrix and the set of parameters, e.g. weights, indifference, preference and veto threshold (inputs), the MCDA algorithm can be seen as a "black-box" procedure providing one or several rankings as outputs. This avenue could be more investigated in the future than it has been done, to the best of my knowledge, so far.

Appendix A: Proofs and additional material

Chapter II

Sampling from the inverse Pólya distribution (page 18). Simulating a $n-$sample from the inverse Pólya distribution can be numerically done using the following algorithm:

- `Set` $n_1 = n$.
- `Step` $k \to k+1$, `for` $k \geq 1$:

  1. `simulate a` n_k-`sample` $x_1^{(k)}, \cdots, x_{n_k}^{(k)}$ `from the Bernoulli distribution with parameter`

  $$\alpha_k \quad = \quad \frac{\alpha + (k-1)\zeta}{1+(k-1)\zeta};$$

  2. `compute` $n_{k+1} = \sum_{i=1}^{n_k} \left(1 - x_i^{(k)}\right)$;
  3. `add to the sample the value` k `replicated` $\sum_{i=1}^{n_k} x_i^{(k)}$ `times;`
  4. `stop if` $n_{k+1} = 0$.

Maximum likelihood estimation of the inverse Pólya distribution (page 18). Assume that among the available data $\mathcal{D}$ there are s survival data and r failure observations. Besides, assume that k_i components have survived until the n_i-th solicitation, for $i = 1, \ldots, s$, and that the r components have broken down after $n_{s+1}, \ldots, n_{s+r}$ solicitations, respectively. Denoting

$$\alpha \quad = \quad \sum_{i=1}^{s} k_i n_i + \sum_{j=s+1}^{s+r} n_j - r,$$

the likelihood of observed data $\mathcal{D}$ is written as

$$\begin{aligned} \mathcal{L}(\mathcal{D}|\alpha,\zeta) \quad &= \quad \frac{(1-\alpha)^{\alpha} \prod_{i=s+1}^{s+r} (\alpha + (n_i - 1)\zeta)}{\left[\prod_{j=1}^{s} \left(\prod_{k=1}^{n_j - 1} (1 + j\zeta)\right)^{k_j}\right] \left[\prod_{l=s+1}^{s+r} \left(\prod_{p=1}^{n_l - 1} (1 + p\zeta)\right)\right]}, \\ &= \quad \frac{(1-\alpha)^{\alpha} \Gamma^{\mu}(1/\zeta) \prod_{i=s+1}^{s+r} (\alpha + (n_i - 1)\zeta)}{\zeta^{\alpha+r} \left[\prod_{j=1}^{s} \Gamma^{k_j}(n_j + 1/\zeta)\right] \left[\prod_{k=s+1}^{s+r} \Gamma(n_k + 1/\zeta)\right]} \quad \text{with } \mu = \sum_{i=1}^{s} k_i + r. \end{aligned} \tag{A.1}$$

Appendix A: Proofs and additional material

The parameters (α,ζ) are assumed to take their values in spaces $\Theta\in[0,1]$ and $\Delta\in[0,\infty[$. A Newton-Raphson descent algorithm for assessing the maximum likelihood estimator (MLE) can be carried out, based on the following rationale. Denoting $\ell(\alpha,\zeta)=\log\mathcal{L}(\mathcal{D}|\alpha,\zeta)$, the equation $\frac{\partial\ell}{\partial\alpha}=0$ implies that

$$\sum_{i=s+1}^{s+r}\frac{1-\alpha}{\alpha+(n_i-1)\zeta} \quad = \quad \alpha, \tag{A.2}$$

and since $\partial^2\ell/\partial\alpha^2<0$, (A.2) has a unique solution maximizing (A.2) given ζ. Furthermore, $\frac{\partial\ell}{\partial\zeta}=0$ implies that

$$\sum_{i=s+1}^{s+r}\frac{\zeta^2}{\alpha+(n_i-1)\zeta}-(\alpha+r)\zeta-\sum_{j=1}^{s+r+1}k'_j\Psi(n'_j+1/\zeta) \quad = \quad 0 \tag{A.3}$$

where $k'_j=k_j$ for $j=1,\ldots,s$, $k'_j=1$ for $j=s+1,\ldots,s+r$ and $k'_{s+r+1}=\mu$. Besides, $n'_j=n_j$ for $j=1,\ldots,s+r$ and $n'_{s+r+1}=0$. Denoting Ψ the digamma function, a descent algorithm can solve (A.2) and (A.3), following the iterative scheme

$$\begin{pmatrix}\alpha_{n+1}\\ \zeta_{n+1}\end{pmatrix} \quad = \quad \begin{pmatrix}\alpha_n\\ \zeta_n\end{pmatrix}-\lambda_n F_n\nabla_n^{-1}$$

where λ_n is an adaptive step, $F_n=F(\alpha_n,\zeta_n)$, $\nabla_n=\nabla(\alpha_n,\zeta_n)$,

$$F(\alpha,\zeta) \quad = \quad \begin{pmatrix}\frac{\alpha}{1-\alpha}-\sum_{i=s+1}^{s+r}\frac{1}{\alpha+(n_i-1)\zeta}\\ \sum_{i=s+1}^{s+r}\frac{\zeta^2}{\alpha+(n_i-1)\zeta}-(\alpha+r)\zeta-\sum_{j=1}^{s+r+1}k'_j\Psi(n'_j+1/\zeta)\end{pmatrix}$$

and

$$\nabla(\alpha,\zeta) \quad = \quad \begin{pmatrix}\frac{\alpha}{(1-\alpha)^2}+\sum_{i=s+1}^{s+r}\frac{1}{(\alpha+(n_i-1)\zeta)^2} & \sum_{i=s+1}^{s+r}\frac{n_i-1}{(\alpha+(n_i-1)\zeta)^2}\\ -\sum_{i=s+1}^{s+r}\frac{\zeta^2}{(\alpha+(n_i-1)\zeta)^2} & \sum_{i=s+1}^{s+r}\frac{2\alpha\zeta}{(\alpha+(n_i-1)\zeta)^3}+(\alpha+r)+\sum_{j=1}^{s+r+1}\frac{k'_j}{\zeta^2}\Psi'(n'_j+1/\zeta)\end{pmatrix}.$$

In practice, the step λ_n can be calibrated in function of the bounds of the parametric space (hence $\alpha_n\in[0,1]\ \forall n$) and the possible lack of inversibility of ∇_n. It is recommended to initialize the method by using a crude likelihood maximization over a grid of $\Theta\times\Delta$.

Concavity of Weibull-1 hazard function when $1<\beta\leq 2$ (page 20). First notice that for $n\geq 3$:

$$\lambda''(n)=\lambda(n)-2\lambda(n-1)+\lambda(n-2)=\exp\left[-\left(\frac{n-1}{\eta}\right)^\beta+\left(\frac{n-2}{\eta}\right)^\beta\right]\cdot A(n)$$

with:

$$A(n)=2-\exp\left[-\left(\frac{n}{\eta}\right)^\beta+2\left(\frac{n-1}{\eta}\right)^\beta-\left(\frac{n-2}{\eta}\right)^\beta\right]-\exp\left[\left(\frac{n-1}{\eta}\right)^\beta-2\left(\frac{n-2}{\eta}\right)^\beta+\left(\frac{n-3}{\eta}\right)^\beta\right].$$

Consider now the function $G(x)=2-\exp(-x)-\exp(x)$. For $x>0$, its first derivative $G'(x)=\exp(-x)-\exp(x)$ is negative; $G(x)$ is strictly decreasing and as $G(0)=0$, it also negative for $x>0$.

For $\beta>1$, $G\left[\left(\frac{n-1}{\eta}\right)^\beta-2\left(\frac{n-2}{\eta}\right)^\beta+\left(\frac{n-3}{\eta}\right)^\beta\right]<0$ as its argument is positive.

Thus:

$$2-\exp\left[\left(-\frac{n-1}{\eta}\right)^\beta+2\left(\frac{n-2}{\eta}\right)^\beta-\left(\frac{n-3}{\eta}\right)^\beta\right]-\exp\left[\left(\frac{n-1}{\eta}\right)^\beta-2\left(\frac{n-2}{\eta}\right)^\beta+\left(\frac{n-3}{\eta}\right)^\beta\right]<0.$$

As for $1<\beta\leq 2$ and $n\geq 3$, $\left(\frac{n}{\eta}\right)^{\beta}-2\left(\frac{n-1}{\eta}\right)^{\beta}+\left(\frac{n-2}{\eta}\right)^{\beta}$ is decreasing, thus:

$$-\exp\left[\left(\frac{n}{\eta}\right)^{\beta}-2\left(\frac{n-1}{\eta}\right)^{\beta}+\left(\frac{n-2}{\eta}\right)^{\beta}\right]\leq-\exp\left[-\left(\frac{n-1}{\eta}\right)^{\beta}+2\left(\frac{n-2}{\eta}\right)^{\beta}-\left(\frac{n-3}{\eta}\right)^{\beta}\right].$$

Consequently, $A(n)<0$ for $1<\beta\leq 2$ and, trivially, $\lambda''(n)$ too.

Proof of Proposition 7.1: bounds for the MTTF of Weibull-1 distribution (page 22). Let us write the expression of the mean of $\mathrm{W}_1(\eta,\beta)$, taking into account Equation II.8, showing the link between the density $f_{\mathrm{W}}(\cdot)$ and the discrete pdf $p_{\mathrm{W}_1}(\cdot)$:

$$\mathbb{E}_{\mathrm{W}_1}(N)=\sum_{i=1}^{\infty}i\,p_{\mathrm{W}_1}(i)=\sum_{i=1}^{\infty}\int_{i-1}^{i}i\,f_{\mathrm{W}}(t)dt=\sum_{i=1}^{\infty}\int_{i-1}^{i}([t]+1)\,f_{\mathrm{W}}(t)dt=\int_{0}^{\infty}([t]+1)\,f_{\mathrm{W}}(t)dt,$$

$[t]$ being the floor function of the random variable $t\sim\mathrm{W}(\eta,\beta)$. As $t-1<[t]\leq t\Rightarrow tf_{\mathrm{W}}(t)<([t]+1)f_{\mathrm{W}}(t)\leq(t+1)f_{\mathrm{W}}(t)$ by integrating this inequality over t one concludes that:

$$\int_{0}^{\infty}t\,f_{\mathrm{W}}(t)dt<\mathbb{E}_{\mathrm{W}_1}(N)\leq\int_{0}^{\infty}t\,f_{\mathrm{W}}(t)dt+\int_{0}^{\infty}1\,f_{\mathrm{W}}(t)dt,$$

i.e. $\mathbb{E}_{\mathrm{W}}(T)<\mathbb{E}_{\mathrm{W}_1}(N)\leq\mathbb{E}_{\mathrm{W}}(T)+1$.

Proof of Proposition 7.2: L^{∞} convergence of Weibull-1 to continuous Weibull (page 23). Since

$$\sup_{t\in\mathbb{R}^+}|p(t|\beta,\eta)-f_{\mathrm{W}}(t|\beta,\eta)|\leq\sup_{t\in\mathbb{R}^+}p(t|\beta,\eta)+\sup_{t\in\mathbb{R}^+}f_{\mathrm{W}}(t|\beta,\eta),$$

we only need to show that:

$$\lim_{\eta\to\infty}\sup_{t\in\mathbb{R}^+}p(t|\beta,\eta)=\lim_{\eta\to\infty}\sup_{t\in\mathbb{R}^+}f_{\mathrm{W}}(t|\beta,\eta)=0 \tag{A.4}$$

The second equality comes from the fact that η is a scale parameter, and that the Weibull density is bounded for $\beta\geq 1$. Hence:

$$\begin{aligned}\sup_{t\in\mathbb{R}^+}f_{\mathrm{W}}(t|\beta,\eta) &= \sup_{t\in\mathbb{R}^+}\frac{1}{\eta}f_{\mathrm{W}}\left(\frac{t}{\eta}|\beta,1\right)\\ &= \frac{1}{\eta}\sup_{t\in\mathbb{R}^+}f_{\mathrm{W}}(t|\beta,1)\overset{\eta\to\infty}{\longrightarrow}0.\end{aligned}$$

The first equality in (A.4) comes from the fact that, for all $t\in[n-1,n]$:

$$\begin{aligned}p(t|\beta,\eta) &= \int_{n-1}^{n}f_{\mathrm{W}}(t|\beta,\eta)dt\\ &\leq \sup_{t\in[n-1,n]}f_{\mathrm{W}}(t|\beta,\eta),\end{aligned}$$

hence

$$\sup_{t\in\mathbb{R}^+}p(t|\beta,\eta)\leq\sup_{t\in\mathbb{R}^+}f_{\mathrm{W}}(t|\beta,\eta).$$

Chapter III

Proof of Equation III.8 (page 35). The probability of a given sequence of $T+1$ states (t from 0 to T) with only one observed state at $t = t_k$ can be written as the sum of T sums fits into each other (with index from 1 to r), one for each unobserved state:

$$\mathbb{P}(\bullet,...,\bullet,s_j,\bullet,...,\bullet) = \sum_{i_0} p_{i_0}(0)\left[\sum_{i_1}\theta_{i_0,i_1}\left[\sum_{i_2}\theta_{i_1,i_2}...\left[\sum_{i_{t_k-1}}\theta_{i_{t_k-2},i_{t_k-1}}\theta_{i_{t_k-1},j}\left[\sum_{i_{t_k}}\theta_{j,i_{t_k}}...\left[\sum_{i_T}\theta_{i_{T-1},i_T}\right]...\right]\right]...\right]\right] \tag{A.5}$$

In the expression above the sum of the first t_k sums (indices from i_0 to i_{t_k-1}) is the unconditional probability $p_j(t_k)$ for the system to be in state j at time t_k. That can be showed by developing the recursive formula (III.1):

$$\begin{aligned} p_j(t_k) &= \sum_{g_1} p_{g_1}(t_k-1)\cdot\theta_{g_1,j} \\ &= \sum_{g_1}\sum_{g_2} p_{g_2}(t_k-2)\cdot\theta_{g_2,g_1}\cdot\theta_{g_1,j} \\ &= \sum_{g_1}\sum_{g_2}\sum_{g_3} p_{g_3}(t_k-3)\cdot\theta_{g_3,g_2}\cdot\theta_{g_2,g_1}\cdot\theta_{g_1 j} \\ &= \dots \end{aligned}$$

and renaming the index $g_1, g_2, g_3, \dots$ as $i_{t_k-1}, i_{t_k-2}, i_{t_k-3}, \dots$ The sum of the remaining sums in the expression of $\mathbb{P}(\bullet,...,\bullet,s_j,\bullet,...,\bullet)$ is one as $\sum_j \theta_{i,j} = 1$. The probability of the sequence is then $p_j(t_k)$. Thus, the likelihood of m incomplete sequences where each individual is observed only once can be written:

$$\prod_{k=1}^{m}\prod_{j=1}^{r}\prod_{t=0}^{T} p_j(t)^{\mathbb{1}_{\{t=t_k, y(k,t)=s_j\}}} = \prod_{j=1}^{r}\prod_{t=0}^{T} p_j(t)^{\sum_k \mathbb{1}_{\{t=t_k, y(k,t)=s_j\}}}$$

which is, under the hypothesis that probabilities $p_j(0)$ are know, the Equation III.8 up to constant of proportionality.

Gibbs sampler for MNAR data when missingness only depends on the actual state (page 36). First initialize the algorithm by arbitrarily completing state sequences. Then at each step $h = 1, 2, \dots$, perform the following two-step procedure:

1. parameters estimation:

$$\boldsymbol{\theta}_i^{[h]} | \boldsymbol{y}^{[h-1]} \sim \text{Dir}\left(\gamma_{i1} + w_{i,1}^{[h-1]}, ..., \gamma_{ir} + w_{1,r}^{[h-1]}\right)$$

and

$$\boldsymbol{\eta}_i^{[h]} | \boldsymbol{y}^{[h-1]} \sim \text{Beta}\left(\alpha_i + a_i^{[h-1]}, \beta_i + b_i^{[h-1]}\right),$$

where $a_i^{[h-1]} = \sum_{t=1}^{T}\sum_{k=1}^{m} \mathbb{1}_{\left\{y_{(k,t)}^{[h-1]} = s_i, x_{(k,t)}=1\right\}}$, $b_i^{[h-1]} = \sum_{t=1}^{T}\sum_{k=1}^{m} \mathbb{1}_{\left\{y_{(k,t)}^{[h-1]} = s_i, x_{(k,t)}=0\right\}}$ and $w_{i,j}^{[h-1]}$ are the same as in Section 2.1.

2. data augmentation: drawing $z_{\texttt{mis}(k,t)}^{[h]}$ conditional on the following probabilities:

$$\mathbb{P}\left(y_{(k,1)}^{[h]} = s_j | y_{(k,2)}^{[h-1]} = s_i, \boldsymbol{\theta}^{[h]}, \boldsymbol{\eta}^{[h]}\right) \propto \eta_j^{[h]} \cdot \theta_{j,i}^{[h]}, \text{ for } t = 1,$$

$$\mathbb{P}\left(y_{(k,T)}^{[h]} = s_j | y_{(k,T-1)}^{[h]} = s_i, \boldsymbol{\theta}^{[h]}, \boldsymbol{\eta}^{[h]}\right) \propto \eta_j^{[h]} \cdot \theta_{i,j}^{[h]}, \text{ for } t = T$$

and

$$\mathbb{P}\left(y_{(k,t)}^{[h]} = s_j | y_{(k,t-1)}^{[h]} = s_{i_1}, y_{(k,t+1)}^{[h-1]} = s_{i_2}, \boldsymbol{\theta}^{[h]}, \boldsymbol{\eta}^{[h]}\right) \propto \eta_j^{[h]} \cdot \theta_{i_1,j}^{[h]} \cdot \theta_{j,i_2}^{[h]}, \text{ otherwise.}$$

Chapter V

Basics of Dempster-Shafer Theory (pages 99 and following.)

Among the many introductions to the the "Dempster-Shafer Theory", also known as "Evidence Theory", we particularly refer to [Helton *et al.* 2006a] and [Limbourg 2007], from which this brief presentation is inspired.

Let us consider an "uncertain" variable (assumed scalar for the sake of simplicity) $X \in \mathbb{R}$. Let us consider the pair $\{\mathcal{A}, v(\cdot)\}$ where $\mathcal{A}$ is a set of subsets of $\mathbb{R}$ and $v(\cdot)$ is a function, such as:

$$\begin{cases} v(A) > 0 \text{ if } A \in \mathcal{A} \\ v(A) = 0 \text{ if } A \notin \mathcal{A} \\ \sum_{A \in \mathcal{A}} v(A) = 1. \end{cases}$$

Under this conditions, the numeric value $v(A)$ is said the "*Basic Probability Assignment*" (BPA), or the "*Basic Belief Assignment*" for the set A. The subsets of $\mathbb{R}$ with non-zero BPA are called "*focal elements*". In practice, in applications of the Dempster-Shafer theory the sets A are intervals and the number of focal sets is finite. Let us note n_f the number of focal sets: $A_1, A_2, ..., A_{n_f}$.

From the function $v(\cdot)$ two different measures of uncertainty can be defined, named "Belief" and "Plausibility":

$$\begin{cases} \text{Bel}(B) = \sum_{i=1}^{n_f} \mathbb{1}_{\{A_i \subset B\}} \cdot v(A_i) \\ \text{Pl}(B) = \sum_{i=1}^{n_f} \mathbb{1}_{\{A_i \cap B \neq \varnothing\}} \cdot v(A_i). \end{cases}$$

An example of calculation of Belief and Plausibility for an interval B, from a given BPA is shown in Figure A.1.

As far as the interpretation of these measures are concerned, quoting the clear explanation of [Helton *et al.* 2006a], "*$Bel(B) = 0$ indicates that none of the available information unambiguously supports B being true (i.e., no focal element [...] is a subset of B, and $Pl(B) = 1$ indicates that none of the available information unambiguously supports B being false (i.e., every focal element [...] intersects B)*".

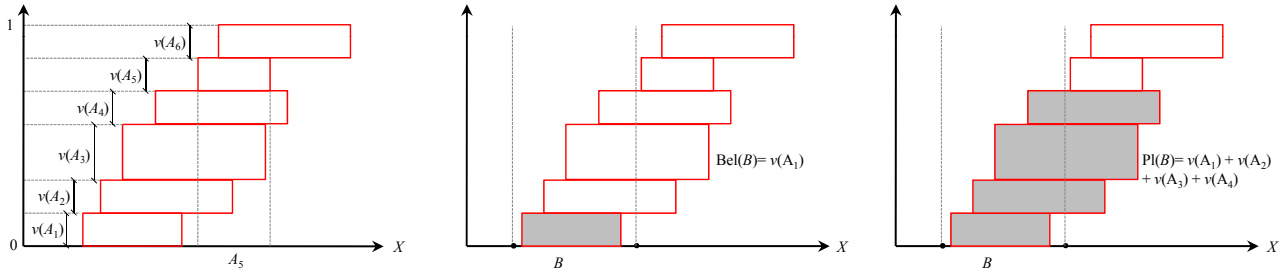


Figure A.1 – Graphical illustration of Belief and Plausibility function in Dempster-Shafer Theory (inspired from [Limbourg 2007]). Left (example of BPA): a finite number of focal sets A_i is defined and a belief mass $v(A_i)$ is given to each of them (here the height of the "rectangles" the projection of which on the x-axis is the set A_i). Center: given a set B, the shaded rectangles are the ones the bases of which are contained in B and thus the ones contributing to Bel(B). Right: Rectangles the bases of which intersects B and thus contributing to Pl(B).

In particular, one is often interested in the belief and the plausibility functions of sets $B =]-\infty, x]$. The functions mapping x to Bel$(]-\infty, x])$ and to Pl$(]-\infty, x])$ are referred as cumulative belief and plausibility functions, noted CBel(x) and CPl(x) respectively.

Intuitively, belief and plausibility can be easily interpreted as lower and upper bounds respectively of the degree of credibility (or degree of belief) an analyst puts in the statement: "$X \in B$". That suggests the common interpretation of Belief and Plausibility as probability bounds for the event under consideration. Mathematical arguments for establishing this link between probability and

Dempster-Shafer theory are provided and discussed by [Halpern & Fagin 1992].[1]

According to this "probabilistic" interpretation, the functions CBel(x) and CPl(x) are bounds for the cumulative distribution function $F(x)$ and define a so-called "probability box" (or pbox), cf. [Ferson & Tucker 2006].

The link between possibility and belief/plausibility functions can be made by considering that the α-cuts of a possibility distribution define nested intervals, which can be imagined as nested focal sets to which a mass $v(A_i) = \alpha_i - \alpha_{i-1}$ is given (cf. Figure A.2).

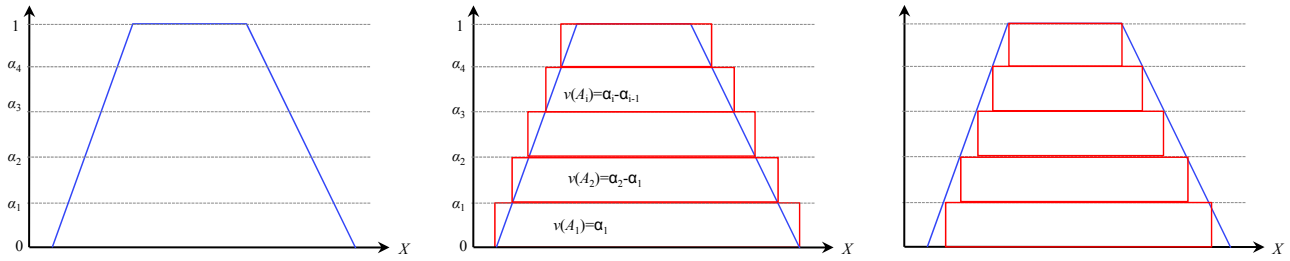


Figure A.2 – Example of construction of a BPA from a possibility distribution (blue curve on the left graph). The α-cuts of levels $\alpha_0 = 0$, α_1, ..., $\alpha_n = 1$ define $n-1$ nested focal sets of masses α_1, $\alpha_2 - \alpha_1$ etc. It is worth noting that (at least) two different ways of building focal sets are available, leading to the so-called "*lower*" and "*upper*" approximation of the possibility function [Baudrit 2005].

[1] However, in spite of the mathematical arguments summarized in [Halpern & Fagin 1992] (essentially based on a discussion about the axioms underlying the Belief functions, which can be seen as an extension of Kolmogorov's ones defining the probability measure), it has to be noticed that the interpretation of belief and plausibility as lower and upper probabilities is controversial from a more "epistemic" viewpoint. On this interpretation, the two co-founders of the theory disagree. If [Dempster 1967] explicitly refers to "Upper and Lower Probabilities", according to [Shafer 1976] probability and belief provide a substantially different information. We will not discuss this point and will follow the Dempster's viewpoint, which is most common within the technical community.

Chapter VI

Proof of theorem 5.1: Predictive estimator as particular Bayes estimator (page 120). According to the definitions of $\hat{\phi}_{\text{pred}}$ and $\hat{H}_{\text{pred}}(\cdot)$ (cf. Equations VI.15 and VI.14 respectively and to the hypothesis 1 of the theorem:

$$\begin{aligned}\hat{\phi}_{\text{pred}} &= \phi\left(\hat{H}_{\text{pred}}(\cdot)\right)\\ &= \underset{d}{\text{Argmin}}\, C\left(d, \hat{H}_{\text{pred}}\right)\\ &= \underset{d}{\text{Argmin}}\, C\left(d, \mathbb{E}_{\Theta}\left[H(\cdot|\Theta)\right]\right).\end{aligned}$$

According to the second hypothesis ("linearity" of the cost):

$$\begin{aligned}\hat{\phi}_{\text{pred}} &= \underset{d}{\text{Argmin}}\, \mathbb{E}_{\Theta}\left[C\left(d, H(\cdot|\Theta)\right)\right]\\ &= \underset{d}{\text{Argmin}} \int C\left(d, H(\cdot|\Theta)\right) p(\Theta)\, d\Theta,\end{aligned}$$

that is the Bayes estimator of ϕ related to the cost function $C\left(d, H(\cdot|\Theta)\right)$.

Gibbs sampler for the estimation of fragility curves (page 125).
First, choose an initial value $\theta^{[0]} = (\mu^{[0]}, \tau^{[0]})$ for the parameters. Then at each step $h = 1, 2, \ldots$, perform the following two-step procedure:

1. data augmentation: generate n values $x_i^{[h]}$ $(i = 1, \ldots, n)$ of the latent log-capabilities from the following truncated normal distributions:

$$\begin{aligned}x_i^{[h]} &\sim \text{Norm}_{]-\infty, x_i]}\left(\mu^{[h-1]}, \tau^{[h-1]}\right) \quad \text{if } y_i = 1 \text{ (damage)}\\ x_i^{[h]} &\sim \text{Norm}_{[x_i, +\infty[}\left(\mu^{[h-1]}, \tau^{[h-1]}\right) \quad \text{if } y_i = 0 \text{ (no damage)}\end{aligned}$$

2. parameters estimation: sample $\theta^{[h]}$ from the following distributions (Gamma-Normal conjugated model)

$$\begin{aligned}\tau^{[h]} &\sim \text{Ga}\left(\frac{2a+n}{2}, \frac{1}{2}\left(S^{[h]} + 2b + \frac{n\lambda\left(\bar{x}^{[h]} - u\right)^2}{n+v}\right)\right)\\ \mu^{[h]}|\tau^{[h]} &\sim \text{Norm}\left(\frac{n\bar{x}^{[h]} + vu}{n+v}, (n+v)\tau^{[h]}\right).\end{aligned}$$

In the expression above $\bar{x}^{[h]} = \sum_i x_i^{[h]}$ and $S = \sum_i (x_i^{[h]} - \bar{x}^{[h]})^2$ are the sufficient statistics of the augmented sample $x_1^{[h]}, x_2^{[h]}, \ldots, x_n^{[h]}$.

Appendix B: Curriculum vitae

Alberto Pasanisi

- Date and place of birth: 27 September 1972, Naples, Italy.
- Nationalities: Italian and French.

Education

- **2004.** PhD in Water Science (specialty: Statistics) from the École Nationale du Génie Rural des Eaux et des Forêts (ENGREF), Paris, France.
  Title of the dissertation: Decision aid in the management of networks of drinkable water meters. Final mark: *Très honorable.*
  PhD Supervisor: Prof. Éric Parent.
  Jury Composition: Gilles Celeux (INRIA Paris Sud, President), Éric Parent (ENGREF, PhD Supervisor), Jacques Bernier (ENGREF, reviewer), Furio Cascetta (University of Naples 2, reviewer), Jean-Pierre Raoult (University of Marne-la-Vallée, reviewer), Pascal Arnac (Veolia Water, Examiner), Dominique Olivier (Veolia Water, Examiner), Bernard Brémond (CEMAGREF Bordeaux, Invited), Lucien Duckstein (ENGREF, Invited)
- **1998.** Engineering habilitation diploma from the University Federico II, Naples, Italy.
- **1997.** Equivalent MSci in Civil Engineering (specialty: Hydraulics) from the University Federico II, Naples, Italy. Final mark: 110/110 *cum laude*.
  Title of the dissertation: *Protecting an eroding shoreline with artificial nourishment interventions*.
- **1990.** High School Diploma (Maturità Classica) from the Instituto Pontano (Naples, Italy). Final mark: 60/60.

Professional activities

- **Since 2008.** Research Engineer and Project Manager (2009-2014) at the *Industrial Risks Management* Department of EDF R&D, Chatou (Paris Area), France.
  Domain of activity: uncertainty in computer simulation, computer experiments, reliability, application of probabilistic and statistical methods to several engineering domain, such as hydraulic and seismic risk, nuclear power plants components, environmental assessment, buildings thermal simulation, eco-toxicology.
- **2004-2008.** Research Engineer at the *Energy in Buildings and Territories* Department of EDF R&D, Les Renardières (Paris Area).
  Domain of activity: buildings thermal simulation, prediction of energy consumption, decision making in thermal retrofitting of existing buildings, technical and economical study of multi-energy HVAC (heating, ventilation, air-conditioning) systems.

- **2000-2004.** Project Engineer and Statistician at the *Networks, Metering, Investments* Departement of the Compagnie Générale des Eaux (Head Office), Paris.
  Domain of activity: optimal renewal of water meters networks, statistical and economical studies, design of experiments.
- **2000.** Project Engineer at the *Production Engineering* Department of Alenia Aerospazio (Head Office), Naples Area, Italy.
  Domain of activity: optimization of sub-contractors production procedures (Boeing 717 and Airbus A340 programs), sub-contractors audits.
- **1999-2000.** Research Engineer at the University of Naples 2, Italy.
  Domain of activity: finite difference and fuzzy modelling of pollutants' transport phenomena in rivers end estuaries.
- **1998-2000.** Consulting Engineer (own-account worker).
  Domain of activity: technical design of several hydraulic works (breakwaters, marinas, sewers, water distribution networks) and study of coastal erosion phenomena, development of a software for coastal erosion forecasting (AquaTecno Company, Italy), consultant of the Environment Agency of Espírito Santo State (Brazil) on coastal erosion.
- **1997-1998.** Military duty in Italian Army (Recruitment Unit, Naples, Italy). Rank: Corporal. Member of the Office of the Troop Commander: various activities and, in particular, reception and redeployment of deserters and jailed soldiers.

Award

- **2008.** EDF Group New Executives Forum, Best twinning project (ex aequo): *Tools for buildings energy efficiency advices for supporting trading development of EDF Belgium*.

Research activities

Topics

- Uncertainty in numerical simulation, computer experiments: contribution of different mathematical settings for coping with uncertainties, code validation, evaluation of failure probabilities by mean of computer experiments.
- Industrial components and structures reliability: Markov deterioration models, discrete and continuous lifetime models, seismic fragility of structures.
- Bayesian statistical analysis: Monte Carlo methods for Bayesian inference, building point estimators for quantities of interest in reliability problems.
- Multi-criteria decision aid: application of methods from the ELECTRE family to engineering problems.

Partnerships/Cooperations (multi-partners)

- **2013-2017**. French National Research Funding Agency (ANR) Project CHORUS (*Common Horizons of Research on Uncertainties in Simulation*).
- **2013-2015.** UK Technology Strategy Board (TSB) Project *Fracture of graphite fuel bricks*.
- **2012-2015.** European Project EESI2 (*Europen Exascale Software Initiative 2*). Task Vice-chair (2012-2013) and Chair (since 2014).
- **2011-2013.** French National Research Funding Agency (ANR) Project FIABILITÉ (*Reliability of building energy performance modeling and simulation*).
- **2009-2012.** French Ministry of Economy Project CSDL (*Complex Systems Design Lab*).

- **2008-2011**. French National Research Funding Agency (ANR) Project OPUS (*Open Source Platform for Uncertainties in Simulation*). Coordinator (2009-2011).
- **1998-2000.** European Project ALICE-QFView (*Quantitative Flow Field Visualisation*).
- **1998-2000.** Italian-Brazilian Cooperation Project NIKE 1 (*Forecasting program of coastal erosion in Espírito Santo*).

Partnerships/Cooperations (bilateral)

- **2012.** EDF R&D - AgroParisTech (France). Topic: *Bayesian estimation of seismic fragility curves*.
- **2011.** EDF R&D - AgroParisTech (France). Topic: *Bayesian optimal design of physical experiments for the evaluation of the toughness of steels*.
- **2010-2012.** EDF R&D - Politecnico di Milano (Italy). Topic: *Advanced computational methods for modelling the mechanisms of degradation in equipments of electricity production plants and uncertainty modelling and propagation*.
- **2008.** EDF R&D - INRIA Paris Sud (France). Topic: *Analysis of discrete lifetime models for industrial components*.
- **2005-2006.** EDF R&D - University of Liège (Belgium). Topic: *Multi-criteria approach in thermal retrofitting of existing buildings*.

Research management

- **Since 2013.** Member of the Council of the French Statistical Society (SFdS).
- **Since 2012.** Vice-chair (2012-2013) and Chair (since 2014) of the transverse task Uncertainty Quantification, Verification & Validation within the European Project EESI2 (*European Exascale Software Initiative 2*).
- **Since 2012.** Corresponding member of the *French Statistical Society* (SFdS) towards ENBIS (*European Network For Industry and Business Statistics*).
- **Since 2012.** Member of the Community Steering Group of MUCM (*Managing Uncertainty in Complex Models*).
- **Since 2012.** Member of the Board of the EDF R&D network *Statistics and Probability*.
- **Since 2009.** President (and founder) of the working group *Reliability and Uncertainties* of the French Statistical Society (SFdS).
- **Since 2009.** Member of the Steering Committee of the OpenTURNS Consortium (*Open source Treatment of Uncertainties, Risk'N Statistics*).
- **Since 2009.** Coordinator of the working group *Uncertainty and Industry* of the *French Institute for Risk Management* (IMdR).
- **2009-2014.** Manager of the EDF R&D Project *Uncertainties*, gathering the the EDF R&D activities concerning uncertainty in numerical simulation.
- **2011-2012.** Founding member of the Special Interest Group *Uncertainty Quantification in Industrial Analysis and Design*, of the *European Research Community On Flow, Turbulence and Combustion* (ERCOFTAC).
- **2009-2011.** Coordinator of National Research Funding Agency (ANR) Project OPUS (*Open Source Platform for Uncertainties in Simulation*).
- **2009.** Member of the Steering Committee of the DICE consortium (*Deep Inside Computer Experiments*).

Scientific editing activities

- **Guest editor** of scientific reviews:
  - 2014. Quality Technology and Quantitative Management: special issue on *Graphical causality models: Trees, Bayesian Networks and Big Data*, with J.M. Poggi and R. Kenett.

  - 2012. Statistics and Computing: special issue on *Modeling of computer experiments for uncertainty propagation and sensitivity analysis*, with A. Antoniadis.
  - 2011. Journal de la Société Française de Statistique: special issue on *Stochastic methods for Sensitivity Analysis*, with C. Prieur and F. Wahl.

- **Since 2013.** Member of the editorial Board of the Journal *Statistique et Société* (SFdS).
- **Peer reviews:**
  - Annals of Operations Research.
  - Communications in Statistics - Simulation and Computation.
  - Computational Statistics and Data Analysis.
  - Environmental Modelling and Software.
  - International Journal of Quality, Statistics, and Reliability.
  - Journal de la Société Française de Statistique.
  - Journal of Statistical Computation and Simulation.
  - Quality Technology and Quantitative Management.
  - Reliability Engineering and System Safety.
  - Statistics and Computing.

Organization of scientific events

- **2014.** ENBIS-SFdS Spring Meeting *Graphical Causality Models: Tree, Bayesian Networks and Big Data*, Institut Henri Poincaré, Paris, France.
- **2013.** SFdS / GdR MASCOT NUM / IMdR workshop *Validation of numerical codes*, Institut Henri Poincaré, Paris, France.
- **2012.** SFdS scientific day *Decision statistics and engineering under uncertainties ... in the steps of Jacques Bernier*, Institut Henri Poincaré, Paris, France.
- **2009-2011.** ANR OPUS Project Workshops:
  - 2011. *Numerical Simulation and Uncertainty Analysis*, OPUS closing workshop, Institut Henri Poincaré, Paris.
  - 2011. *Uncertainty Quantification, High Performance Computing, Calculation Environments and Software*, University Joseph Fourier, Grenoble, France.
  - 2010. *Uncertainty Propagation, Estimation of Rares Quantiles and Low Probabilities of Failure*, Institut Henri Poincaré, Paris.
  - 2009. *Spectral Methods and Polynomial Chaos*, EADS Innovation Works, Suresnes, France.
  - 2009. *Learning and Model Selection*, CEA, Saclay, France.
- **2010.** SFdS / IMdR scientific day *Uncertainties and Inverse Problems: Methodological Challenges and Applications*, French National Metrology Lab, Paris, France.

Organization of special sessions of congresses

- **Since 2010.** Annual Congresses of the French Statistical Society (Journées de Statistique), sessions of the group *Reliability and Uncertainties*:
  - 2013. 45th SFdS Congress, Toulouse, France.
  - 2012. 44th SFdS Congress, Bruxelles, Belgium.
  - 2011. 43th SFdS Congress, Gammarth, Tunisia.
  - 2010. 42 th SFdS Congress, Marseille, France.
- **2013.** European Network for Business and Industrial Statistics Annual Conference ENBIS-13. Invited sessions: *Financial, insurance and natural risks* and *Computer experiments & reliability*, Ankara, Turkey.
- **2012.** European Network for Business and Industrial Statistics Annual Conference ENBIS-12. Invited session: *Reliability and Uncertainty Analysis in Industry*, Ljubljana, Slovenia.
- **2011.** 58th World Statistics Congress, ISI-2011. Special Topics Session: *Uncertainty, Industry and Statistics*, Dublin, Ireland.

Participation in congress program committees

- **2014.** Lambda-Mu 19 Congress, *Deciding in an uncertain world: a major stake in risk management,* Dijon, France.
- **2014**. ENBIS-SFdS Spring Meeting on *Graphical Causality Models: Tree, Bayesian Networks and Big Data*, Paris (co-chair).
- **2013.** 7th International Conference on Sensitivity Analysis of Model Output (SAMO 2013), Nice, France.
- **2013.** 45th SFdS Congress, Toulouse, France.
- **2011.** European Safety and Reliability Conference (ESREL 2011), Troyes, France (responsible for the thematic area: Bayesian methods).

Teaching activities

Academic courses

- **Since 2008.** Université de Technologie de Troyes: *Industrial feedback and safety* (course on Bayesian reliability).
- **2009-2010.** École Centrale Paris: *Bayesian Uncertainty Analysis* (introductory course).
- **2003-2004.** École Nationale des Ponts et Chaussées: *Statistics and data analysis*.
- **2001-2003.** École Nationale du Génie Rural, des Eaux et des Forêts: *Risk assessment and management*.

Professional training courses

- **Since 2008.** EDF R&D: *Modelling the reliability of components: probabilistic and statistical methods, uncertainty analysis*, Chatou, France.
- **Since 2008.** EDF R&D: *Uncertainty management in numerical simulation* (manager of the course between 2008 and 2012), Chatou, France.
- **Since 2011.** The Materials Ageing Institute: *Management of Uncertainties in Computational Mechanics*, Les Renardières, France.
- **2011-2012.** ERCOFTAC Society: *Awareness course on Uncertainty Management and Quantification in Industrial Analysis and Design*, Münich, Germany (2011), Hampton, VA USA (2011), Chatou, France (2012).

- **2009-2012.** French National Metrology Lab (LNE): *Managing uncertainty in engineering studies*, Paris (training course in three modules, manager of two of them).
- **2012.** Spring School SIMUREX on the *Reliability of thermal simulation based forecasts of energy consumption*, Cargèse, France.
- **2010.** French Statistical Society. Tutorial on *Methods and tools for uncertainty management in industry*, Marseille, France.

Management and scientific follow-up of students

PhD students

- **Since 2012.** Guillaume Damblin (AgroParisTech). Topic: *Uncertainty and computer experiments: a Bayesian decision approach from validation to forecasting* (joint work with with É. Parent, P. Barbillon and M. Keller).
- **2010-2011.** Khouloud Ghorbel (AgroParisTech). Topic: *Bayesian uncertainty quantification in computer experiments within an industrial framework* (joint work with with É. Parent and M. Keller).
- **2005-2008.** Caroline Flory-Celini (University of Lyon 1). Topic: *Modelling and ranking bio-climatic solutions for existing residential buildings* (joint work with D. Covalet and J. Virgone, follow-up on multi-criteria methods).

Post doctoral fellows

- **2010.** Merlin Keller. Topic: *Bayesian setting and uncertainty quantification* (joint work with É. Parent and N. Bousquet).

Graduate students

- **2013.** Pratnya Paramitha Oktaviana (University Pierre et Marie Curie Paris 6). Topic: *Bayesian assessment of statistical distributions of soil-to-plant transfer factors for metals* (joint work with with M. Keller, T. Tanaka and M.P. Étienne).
- **2013.** Raphaël Vinet (University Pierre et Marie Curie Paris 6). Topic: *Uncertainty analysis in numerical simulation by means of Uranie and OpenTURNS software* (joint work with M. Baudin).
- **2011.** Elisa Ferrario (Politecnico di Milano). Topic: *Uncertainty analysis in risk assessment for environmental applications* (reviewer with E. Zio of the final dissertation "tesi di laurea").
- **2010.** Marjorie Jala (ISUP, University Pierre et Marie Curie Paris 6). Topic: *Critical analysis of different frameworks for uncertainty assessment* (joint work with B. Iooss).
- **2010.** Nicolas Chrysanthos and Thomas Juan (École Supérieure d'Électricité, SUPÉLEC). Topic: *Using metamodels for estimating failure probabilities* (with A. Arnaud, E. Vazquez, J. Bect).
- **2009.** Amine Bennabi (Conservatoire National des Arts et Métiers). Topic: *Bayesian approach and uncertainty assessment in industrial practice* (joint work with É. Parent).
- **2009.** Shuai Fu (University Pierre et Marie Curie Paris 6). Topic: *Estimation of discrete Markov models for industrial reliability* (joint work with N. Bousquet).
- **2007.** Audrey Laude (École Spéciale des Travaux Publics). Topic: *Multi-criteria decision aid in energy retrofitting of buildings*.
- **2006.** Nicolas Bûcher (École Nationale Supérieure de l'Aéronautique et de l'Espace). Topic: *Multi-criteria decision aid in energy retrofitting of buildings*.
- **2005.** François Fernandez (University of La Rochelle). Topic: *Evaluation of multi-energy systems for buildings heating and air-conditioning*.

Participation in PhD juries

- **2013.** Rania Merheb (University of Bordeaux 1). Topic: *Reliability of tools for forecasting the behaviour of complex thermal systems* (examiner).
- **2011.** Nicolas Durrande (École des Mines de St. Étienne). Topic: *Study of classes of kernels adapted to the simplification and the interpretation of approximation models. A functional and probabilistic approach* (examiner).
- **2011.** Samir Touzani (University Joseph Fourier Grenoble 1). Topic: Surface response methods based on decomposition of the functional variance and application to sensitivity analysis (examiner).
- **2008.** Caroline Flory-Celini (University Claude Bernard Lyon 1). Topic: *Modelling and ranking bio-climatic solutions for existing residential buildings* (invited member).

Publications and communications

Articles in peer-reviewed journals

- [Damblin *et al.* 2014c] G. Damblin, M. Keller, A. Pasanisi, P. Barbillon and É. Parent. *Approche décisionnelle bayésienne pour estimer une courbe de fragilité*. Journal de la Société Française de Statistique (accepted), 2014.
- [Pedroni *et al.* 2013a] N. Pedroni, E. Zio, E. Ferrario E., Pasanisi A. and M. Couplet. *Hierarchical propagation of probabilistic and non-probabilistic uncertainty in the parameters of a risk model*. Computers and Structures, vol. 126, pages 199–213, 2013.
- [Ciffroy *et al.* 2013] Ph. Ciffroy, M. Keller and A. Pasanisi. *Estimating Hazardous Concentrations by an informative Bayesian approach*. Environmental Toxicology and Chemistry, vol. 32, no. 3, pages 602–611, 2013.
- [Pasanisi *et al.* 2012c] A. Pasanisi, M. Keller and É. Parent. *Estimation of a quantity of interest in uncertainty analysis: some help from Bayesian decision Theory*. Reliability Engineering and System Safety, vol. 100, pages 93–101, 2012.
- [Pasanisi *et al.* 2012a] A. Pasanisi, S. Fu and N. Bousquet. *Estimating Discrete Markov Models From Various Incomplete Data Schemes*. Computational Statistics and Data Analysis, vol. 56, no. 9, pages 2609–2625, 2012.
- [Keller *et al.* 2011c] M. Keller, A. Pasanisi and É. Parent. *Réflexions sur l'analyse d'incertitudes dans un contexte industriel : Information disponible et enjeux décisionnels*. Journal de la Société Française de Statistique, vol. 152, no. 4, pages 60–77, 2011.
- [Pasanisi & Ojalvo 2008] A. Pasanisi and J. Ojalvo. *A multi-criteria decision tool to improve the energy efficiency of residential buildings*. Foundations of Computing and Decision Sciences, vol. 33, no. 1, pages 71–82, 2008.
- [Pasanisi & Parent 2004] A. Pasanisi and É. Parent. *Modélisation bayésienne du vieillissement des compteurs d'eau par mélange de classes d'appareils de différents états de dégradation*. Revue de Statistique Appliquée, vol. 52, no. 1, pages 39–65, 2004.
- [Di Natale *et al.* 2001] M. Di Natale, L. Duckstein and A. Pasanisi. *Forecasting pollutants transport in a river by a fuzzy rule-based model*. Belgian Journal of Operations Research, Statistics and Computer Science, vol. 41, no. 3-4, pages 129–138, 2001.

Book chapters

- [Parent *et al.* 2014] É. Parent, A. Pasanisi, N. Bousquet, M. Keller and J. Bernier. *Considérations décisionnelles pour la construction d'un ouvrage de protection contre les crues*. In J.J. Droesbeke, Saporta G. and C. Thomas-Agnan (Eds.), Approches statistiques du risque. Technip, 2014.
- [Pasanisi & Dutfoy 2012] A. Pasanisi and A. Dutfoy. *An Industrial Viewpoint on Uncertainty Quantification in Simulation: Stakes, Methods, Tools, Examples*. In A.M. Dienstfrey and R.F. Boisvert (Eds.), Uncertainty Quantification in Scientific Computing, pages 27–45. Springer, 2012.

Submitted articles and preprints

- [Keller *et al.* 2014b] M. Keller, A. Pasanisi, M. Marcilhac, T. Yalamas, R. Secanell and G. Senfaute. *A Bayesian methodology applied to the estimation of earthquake recurrence parameters for seismic hazard assessment*. Submitted, 2014.
- [Pedroni *et al.* 2014] N. Pedroni, E. Zio, A. Pasanisi and M. Couplet. *A critical discussion and practical recommendations on some issues relevant to the non-probabilistic treatment of uncertainty in engineering risk assessment*. Submitted, 2014.
- [Pasanisi *et al.* 2013c] A. Pasanisi, C. Roero, N. Bousquet and E. Remy. *On the practical interest of discrete lifetime models in industrial reliability studies*. Submitted, 2013.
- [Keller *et al.* 2010b] M. Keller, É. Parent and A. Pasanisi. *On the Role of Decision Theory in Uncertainty Analysis*. ArXiv e-prints, 1009.4342, 2010.

Special issues of journals

- [Poggi *et al.* 2014] J.M. Poggi, R. Kenett and A. Pasanisi Eds. *Special issue: Graphical Causality Models: Trees, Bayesian Networks and Big Data*. Quality Technology and Quantitative Management, vol. 11, no. 1, pages 1–147, 2014.
- [Antoniadis & Pasanisi Eds. 2012] A. Antoniadis and A. Pasanisi Eds. *Special Issue: Modeling of computer experiments for uncertainty propagation and sensitivity analysis*. Statistics and Computing, vol. 22, no. 3, pages 677–847, 2012.
- [Prieur *et al.* 2011] C. Prieur, A. Pasanisi and F. Wahl Eds. *Special issue: Stochastic methods for sensitivity analysis*. Journal de la Société Française de Statistique, vol. 152, no. 1, pages 1–130, 2011.

PhD dissertation

- [Pasanisi 2004a] A. Pasanisi. *Aide à la décision dans la gestion des parcs de compteurs d'eau potable*. PhD thesis, École Nationale du Génie Rural, des Eaux et des Forêts, 2004.

Invited talks

- [Pasanisi 2012a] A. Pasanisi. *Analyse d'incertitudes en simulation numérique : méthodes, outils, défis scientifiques*. In 30èmes rencontres AUGC et IBPSA-France, Chambéry (France), June 2012.
- [Pasanisi 2012b] A. Pasanisi. *Approche décisionnelle pour l'estimation de critères fiabilistes à l'aide de la simulation numérique : point de vue industriel et exemples d'application*. In Journées JSTAR 2012, Rennes (France), October 2012.
- [Gaudier *et al.* 2011] F. Gaudier, A. Pasanisi and A. Dutfoy. *Analyse d'incertitudes et simulation dans le domaine de l'énergie*. In 28ème Forum ORAP, Paris, October 2011.
- [Pasanisi 2011] A. Pasanisi. *An Industrial Viewpoint on Uncertainty Quantification in Simulation: Stakes, Methods, Tools, Examples*. In IFIP Working Conference on Uncertainty Quantification in Scientific Computing, Boulder (CO, USA), August 2011.

Communications in international conferences

- [Damblin *et al.* 2014a] G. Damblin, P. Barbillon, M. Keller, A. Pasanisi and E. Parent. *Bayesian validation of a computer model for the energy consumption of a building, with application to optimal electric bill pricing*. In 2014 International Society for Bayesian Analysis World Meeting, Cancun (Mexico), July 2014.
- [Keller *et al.* 2014a] M. Keller, A. Pasanisi, G. Damblin, M. Marcilhac, T. Yalamas, R. Secanell and G. Senfaute. *A Bayesian methodology for the estimation of the failure probability of a structure submitted to seismic hazard.* In 2014 International Society for Bayesian Analysis World Meeting, Cancun (Mexico), July 2014.

- [Pedroni *et al.* 2013b] N. Pedroni, E. Zio, E. Ferrario, A. Pasanisi and M. Couplet. *Bayesian update of the parameters of probability distributions for risk assessment in a two-level hybrid probabilistic-possibilistic uncertainty framework*. In ESREL 2013 Conference, Amsterdam (Netherlands), September 2013.
- [Pasanisi *et al.* 2013b] A. Pasanisi, M. Keller, M. Marcilhac, T. Yalamas, R. Secanell, C. Martin and G. Senfaute. *Bayesian assessment of seismic hazard curves*. In ENBIS13 Conference, Ankara (Turkey), September 2013.
- [Damblin *et al.* 2013b] G. Damblin, M. Keller, A. Pasanisi, É. Parent, P. Barbillon and J. Bernier. *A Bayes decision approach to code validation in an industrial context*. In 7th International Conference on Sensitivity Analysis of Model Output, Nice (France), July 2013.
- [Damblin *et al.* 2012b] G. Damblin, M. Keller, A. Pasanisi, I. Zentner and É. Parent. *How to Choose a Fragility Curve? Bayesian Decision Theory Applied to Uncertainty Analysis in an Industrial Context*. In ENBIS12 Conference, Ljubljana (Slovenia), September 2012.
- [Pedroni *et al.* 2012] N. Pedroni, E. Zio, E. Ferrario and A. Pasanisi. *Propagation of aleatory and epistemic uncertainties in the model for the design of a flood protection dike*. In ESREL 2012 / PSAM11 Conference, Helsinki (Finland), June 2012.
- [Baraldi *et al.* 2011] P. Baraldi, N. Pedroni, E. Zio, E. Ferrario, A. Pasanisi and M. Couplet. *Monte Carlo and fuzzy interval propagation of hybrid uncertainties on a risk model for the design of a flood protection dike*. In ESREL 2011 Conference, pages 2167–2175, Troyes (France), September 2011.
- [Keller *et al.* 2011a] M. Keller, É. Parent and A. Pasanisi. *Should industrial uncertainty analysis be Bayesian?* In 58th World Statistics Congress of the International Statistical Institute, Dublin (Ireland), August 2011.
- [Keller *et al.* 2010a] M. Keller, É. Parent, N. Bousquet and A. Pasanisi. *Bayesian and frequentist parametric prediction of a tail probability in an industrial reliability context*. In 9th Valencia International Meeting on Bayesian Statistics, Benidorm (Spain), June 2010.
- [Pasanisi *et al.* 2009a] A. Pasanisi, E. de Rocquigny, N. Bousquet and É. Parent. *Some useful features of the Bayesian setting while dealing with uncertaintiesin industrial practice*. In ESREL 2009 Conference, Prague (Cech Rep.), September 2009.
- [Pasanisi & Ojalvo 2007] A. Pasanisi and J. Ojalvo. *A multicriteria decision tool to improve the energy efficiency of residential houses*. In 65th Meeting of the European Working Group on Multiple Criteria Decision Aiding, Poznan (Poland), April 2007.
- [Pasanisi *et al.* 2002] A. Pasanisi, É. Parent, P. Arnac and F. Paquet. *Describing the Ageing of Water Meters in Vivendi Water Distribution Networks by a Dynamic State Model*. In 7th Valencia International Meeting on Bayesian Statistics, Tenerife (Spain), June 2002.
- [Vucinic *et al.* 2000] D. Vucinic, J. Favaro, B. Sünder, I. Jenkinson, G. Tanzini, B.K. Hazarika, M.R. d'Alcalà, D. Vicinanza, R. Greco and A. Pasanisi. *Fast and convenient access to fluid dynamics data via the World Wide Web*. In ECCOMAS 2000 Conference, Barcelona (Spain), September 2000.
- [Di Natale *et al.* 2000] M. Di Natale, L. Duckstein and A. Pasanisi. *Forecasting pollutants transport in river by a fuzzy rule-based model*. In Belgium Fuzzy II Conference, Mons (Belgium), March 2000.

Communications in French national conferences

- [Damblin *et al.* 2014b] G. Damblin, P. Barbillon, M. Keller, A. Pasanisi and É. Parent. *Plans d'expérience séquentiels pour la calibration de modèles numériques coûteux*. In 46èmes Journées de Statistique, Rennes (France), June 2014.
- [Pasanisi 2014] A. Pasanisi. *Fragilité et aléa sismique : apports de l'approche bayésienne*. In 7èmes Rencontres Stat. Au Sommet de Rochebrune, Megève, France, April 2014.
- [Caruso & Pasanisi 2013] A. Caruso and Pasanisi. *Incertitudes en simulation numérique à EDF : enjeux et activités*. In WorkStat 2013, Clamart (France), November 2013.

Appendix B: Curriculum vitae

- [Rapenne *et al.* 2013] S. Rapenne, M. Baudin, E. Demay, K. Shakourzadeh, A. Pasanisi and O. Alos-Ramos. *Une boîte à outils Scilab pour prédire l'entartrage dans les circuits de refroidissement des centrales nucléaires équipées de tours aéroréfrigérantes et son couplage avec OpenTURNS*. In ScilabTEC, Palaiseau (France), June 2013.
- [Ardillon *et al.* 2012] E. Ardillon, A. Arnaud, N. Bousquet, M. Couplet, A. Dutfoy, Iooss B., M. Keller, A. Pasanisi, E. Remy and V. Verrier. *Identification des problématiques de recherche pour la durée de vie des composants et les incertitudes*. In 18ème Congrès Lambda-Mu, Tours (France), October 2012.
- [Ciffroy *et al.* 2012] Ph. Ciffroy, M. Keller, A. Pasanisi and B. Richard. *Constructing species sensitivity distributions (SSD) using informative Bayesian approach*. In Colloque 2012 de la Société Française d'Ecotoxicologie Fondamentale et Appliquée, Lyon (France), July 2012.
- [Damblin *et al.* 2012a] G. Damblin, M. Keller, A. Pasanisi, I. Zentner and É. Parent. *Comment décider d'une courbe de fragilité ? Approche décisionnelle bayésienne pour traiter les incertitudes dans un contexte industriel*. In 44èmes Journées de Statistique, Bruxelles (Belgium), May 2012.
- [Pasanisi *et al.* 2012b] A. Pasanisi, M. Keller and É. Parent. *Estimation de quantités d'intérêt en fiabilité industrielle... avec un petit coup de main de Thomas Bayes*. In 6èmes Rencontres Stat. Au Sommet de Rochebrune, Megève (France), April 2012.
- [Keller *et al.* 2011d] M. Keller, A. Pasanisi and É. Parent. *Sur l'intérêt du bayésien pour l'analyse d'incertitudesdans un contexte industriel : Information disponible et enjeux décisionnels.* In Rencontres AppliBUGS, Paris (France), December 2011.
- [Keller *et al.* 2011b] M. Keller, A. Pasanisi, K. Ghorbel and É. Parent. *Réflexions sur l'analyse d'incertitude dans un contexte industriel: information disponible et enjeux décisionnels*. In 43èmes Journées de Statistique, Gammarth (Tunisia), May 2011.
- [Pasanisi *et al.* 2010] A. Pasanisi, S. Fu and N. Bousquet. *Estimation de modèles markoviens discrets dans un cadre industriel fiabiliste à données manquantes*. In 42èmes Journées de Statistique, Marseille (France), May 2010.
- [Arnaud *et al.* 2010] A. Arnaud, J. Bect, M. Couplet, A. Pasanisi and E. Vazquez. *Évaluation d'un risque d'inondation fluviale par planification séquentielle d'expériences*. In 42èmes Journées de Statistique, Marseille (France), May 2010.
- [Couplet *et al.* 2010] M. Couplet, L. Lebrusquet and A. Pasanisi. *Caractérisation des coefficients de Strickler d'un fleuve par inversion probabiliste*. In 42èmes Journées de Statistique, Marseille (France), May 2010.
- [Pasanisi 2010] A. Pasanisi. *Bayesian uncertainty assessment in industrial practice: EDF-R&D framework and on-going works*. In 5èmes Rencontres Stat. au sommet de Rochebrune, Megève (France), March 2010.
- [Pasanisi *et al.* 2009b] A. Pasanisi, C. Roero, G. Celeux and E. Remy. *Quelques considérations sur l'utilisation pratique des modèles discrets de survie en fiabilité industrielle*. In 41èmes Journées de Statistique, Bordeaux (France), May 2009.
- [Dutfoy *et al.* 2009] A. Dutfoy, I. Dutka-Malen, A. Pasanisi, R. Lebrun, F. Mangeant, J. Sen Gupta, M. Pendola and T. Yalamas. *OpenTURNS, an Open Source initiative to Treat Uncertainties, Risks'N Statistics in a structured industrial approach*. In 41èmes Journées de Statistique, Bordeaux (France), May 2009.
- [Pasanisi 2004c] A. Pasanisi. *La subjectivité dans l'approche scientifique : Aristote*. In 2èmes Rencontres stat. au sommet de Rochebrune, Megève (France), March 2004.
- [Pasanisi 2004d] A. Pasanisi. *Transitions latentes, données inadaptées et compteurs ... histoire d'une thèse*. In 2èmes Rencontres stat. au sommet de Rochebrune, Megève (France), March 2004.
- [Pasanisi 2003] A. Pasanisi. *Modélisation bayésienne de la dégradation métrologiquedes compteurs d'eau avec la prise en compte de sources multiples d'information*. In 35èmes Journées de Statistiques, Lyon (France), June 2003.
- [Pasanisi 2002] A. Pasanisi. *Modéliser le vieillissement des compteurs d'eau sur la base d'informations incomplètes*. In 1ères Rencontres stat. au sommet de Rochebrune, Megève (France), March 2002.

Technical reports (selection)

- [Baudin *et al.* 2013] M. Baudin, A. Pasanisi and F. Gaudier. *Comparaison entre Uranie et OpenTURNS*. Tech. Report H-T57-2013-03841-FR, EDF R&D, 2013.
- [Erbacci *et al.* 2013] G. Erbacci, F. Bodin, V. Bergeaud, A. Pasanisi, S. McIntosh-Smith, T. Ludwig, F. Cappello, C. Cavazzoni, M.C. Sawley *WP5 First Intermediate Report: Cross Cutting Issues Working Groups*. Tech. Report D5.1, European Exascale Software Initiative - EESI2 Project, 2013.
- [Pasanisi *et al.* 2013a] A. Pasanisi, M. Couplet and A. Dutfoy. *Guide Méthodologique pour le Traitement des Incertitudes - Mise à jour 2013*. Tech. Report H-T57-2013-02207-FR, EDF R&D, 2013.
- [Baudin & Pasanisi 2012] M. Baudin and A. Pasanisi. *Traitement des incertitudes dans la simulation numérique descircuits de refroidissement des CNPE avecaéroréfrigérant: couplage entre les logiciels CooliSS ST et Open TURNS*. Tech. Report H-T57-2012-02834-FR, EDF R&D, 2012.
- [OPUS 2011] OPUS. *Final Report of the ANR OPUS project*. Tech. report D-WP0/11/03/A, OPUS Project, 2011. www.opus-project.fr.
- [Iooss *et al.* 2011] B. Iooss, E. Ardillon, A. Arnaud, G. Arnaud, J. Baccou, M. Berveiller, G. Blatman, M. Couplet, A. Dutfoy, F. Gaudier, N. Gilardi, J.M. Martinez, A. Pasanisi, A.L. Popelin, Y. Richet and I. Zentner. *Benchmark entre logiciels de traitement des incertitudes*. Tech. Report F2.1.1.2, Complex System Design Lab Project, 2011.
- [Iooss *et al.* 2010] B. Iooss, A. Pasanisi and M. Jala. *Étude bibliographique et méthodologique du traitement des incertitudes en phase de conception*. Tech. Report H-T56-2010-02503-FR, EDF R&D, 2010.
- [Roero *et al.* 2008] C. Roero, A. Pasanisi and G. Celeux. *Éclairages sur les modèles discrets de survie en fiabilité industrielle*. Tech. Report H-T56-2008-04499-FR, EDF R&D, 2008.
- [Pasanisi 2008a] A. Pasanisi. *Cahier des charges d'un module générique d'aide multicritère à la décision*. Tech. Report H-E14-2008-00073-FR, EDF R&D, 2008.
- [Pasanisi 2008b] A. Pasanisi. *Réflexions préalables au développement d'un logiciel d'évaluation probabiliste du débit de fuite d'un générateu rde vapeur en conditions d'EHP*. Tech. Report H-T56-2008-02960-FR, EDF R&D, 2008.
- [Le Mouel *et al.* 2007] A. Le Mouel, A. Pasanisi and P. Girault. *SimUK (UK Domestic Energy Model) - Design Reference Guide*. Tech. Report H-E14-2007-03370-EN, EDF R&D, 2007.
- [Pasanisi & Évin 2007] A. Pasanisi and F. Évin. *Compétitivité technico-économique des solutions multi-énergies "PAC + Bois"*. Tech. Report H-E14-2007-02874-FR, EDF R&D, 2007.
- [Pasanisi & Heijmans 2007] A. Pasanisi and C. Heijmans. *Les outils de conseil en efficacité énergétique des bâtiments au service du développement commercial d'EDF-Belgium*. Tech. Report H-E14-2007-02023-FR, EDF R&D, 2007.
- [Pasanisi & Bieret 2007] A. Pasanisi and J. Bieret. *Réflexions autour d'un outil de définition et d'évaluation de solutions intégrées visant différents niveaux de performance énergétique pour les maisons individuelles neuves ou à rénover. Comparaison des outils existants*. Tech. Report H-E14-2007-01765-FR, EDF R&D, 2007.
- [Pasanisi 2007] A. Pasanisi. *Aide multicritère à la décision en rénovation énergétique des logements existants : l'outil REFLEX*. Tech. Report H-E14-2007-00425-FR, EDF R&D, 2007.
- [Pasanisi & Coince 2006] A. Pasanisi and A.S. Coince. *Évaluation technico-économique de solutions multi-énergies pour le marché résidentiel*. Tech. Report H-E14-2006-04298-FR, EDF R&D, 2006.
- [Pasanisi 2006] A. Pasanisi. *Outil d'aide à la décision "Multi-énergies" version 1.0. Principes de modélisation et fonctionnement*. Tech. Report H-E14-2006-00159-FR, EDF R&D, 2006.
- [Pasanisi 2005] A. Pasanisi. *Éléments d'aide multicritère à la décision*. Tech. Report H-E14-2005-00770-FR, EDF R&D, 2005.

- [Pasanisi 2004b] A. Pasanisi. *Analyse technico-économique de systèmes mono-énergie et multi-énergie en résidentiel*. Tech. Report H-E14-2004-02934-FR, EDF R&D, 2004.
- [Pasanisi & Sivadier 2004] A. Pasanisi and C. Sivadier. *Poids de la climatisation dans la facture énergétique en résidentiel. Analyse des facteurs d'influence et mesures de réduction des consommations*. Tech. Report H-E14-2004-01571-FR, EDF R&D, 2004.

References

[Abbott & Price 2005] M.B. Abbott and W.A. Price. Coastal, Estuarial and Harbour Engineer's Reference Book. Taylor & Francis, 2005. 8

[Abraham & Cadre. 2004] C. Abraham and B. Cadre. *Asymptotic global robustness in Bayesian decision theory*. Annals of Statistics, vol. 32, pages 1341–1366, 2004. 117

[Accoto 2009] N. Accoto. *The hidden mixture Markov model with application to the analysis of electricity supply: a Bayesian perspective*. In Complex Data Modeling and Computationally Intensive Statistical Methods for Estimation and Prediction (SCo 2009), Milano (Italy), September 2009. 38

[Aggarwal 1993] K.K. Aggarwal. Reliability Engineering. Kluwer Academic Publishers, 1993. 15

[Ahmed & Wu 2013] W. Ahmed and Y.W. Wu. *A survey on reliability in distributed systems*. Journal of Computer and System Sciences, vol. 79, no. 8, pages 1243–1255, 2013. 14

[Alajaji & Fuja 1993] F. Alajaji and T.E. Fuja. *A Communication Channel Modeled on Contagion*. Tech. Report 93.78, ISR. University of Maryland, 1993. 16

[Allman *et al.* 2009] E.S. Allman, C. Matias and J.R. Rhodes. *Identifiability of parameters in latent structure models with many observed variables*. The Annals of Statistics, vol. 37, pages 3099–3132, 2009. 32

[Almalki & Nadarajah 2014] S.J. Almalki and S. Nadarajah. *Modifications of the Weibull distribution: A review*. Reliability Engineering and System Safety, vol. 124, pages 32–55, 2014. 16, 20

[Alzaatreh *et al.* 2012] A. Alzaatreh, C. Lee and F. Famoye. *On the discrete analogues of continuous distributions*. Statistical Methodology, vol. 9, no. 6, pages 589 – 603, 2012. 19

[Ancelet *et al.* 2012a] S. Ancelet, N. Bousquet and É. Parent. *Optimisation bayésienne de plans d'expériences*. In 6èmes Rencontres stat. au sommet de Rochebrune, Megève (France), April 2012. 60

[Ancelet *et al.* 2012b] S. Ancelet, N. Bousquet and É. Parent. *Projet Incertitudes 2015 - Calibration bayésienne de plans d'expérience pour la quantification de sources d'incertitudes - Application à l'estimation de la distribution statistique de la ténacité de l'acier*. Tech. Report H-T56-2012-01024-FR, EDF R&D, 2012. 60

[Andrieu & Atchadé 2007] C. Andrieu and Y. Atchadé. *On the efficiency of adaptive MCMC algorithms*. Electronic Communications in Probability, vol. 12, pages 336–349, 2007. 39

[Andrieu & Moulines 2006] C. Andrieu and E. Moulines. *On the ergodicity properties of some adaptive MCMC algorithms*. Annals of Applied Probability, vol. 16, pages 1462–1505., 2006. 39, 45

[Andrieu & Thoms 2008] C. Andrieu and J. Thoms. *A tutorial on adaptive MCMC*. Statistics and Computing, vol. 18, pages 343–373, 2008. 37, 38, 39, 41

References

[Andrieux 2011] S. Andrieux. *L'Open Source à EDF R&D. Politique générale et panorama*. In R&D industrielle d'expertise et open source : un avenir commun?, Clamart (France), July 2011. 64

[Antoniadis & Pasanisi Eds. 2012] A. Antoniadis and A. Pasanisi Eds. *Special Issue: Modeling of computer experiments for uncertainty propagation and sensitivity analysis*. Statistics and Computing, vol. 22, no. 3, pages 677–847, 2012. 60, 168

[Ardillon *et al.* 2012] E. Ardillon, A. Arnaud, N. Bousquet, M. Couplet, A. Dutfoy, Iooss B., M. Keller, A. Pasanisi, E. Remy and V. Verrier. *Identification des problématiques de recherche pour la durée de vie des composants et les incertitudes*. In 18ème Congrès Lambda-Mu, Tours (France), October 2012. 60, 170

[Arefi *et al.* 2010] M Arefi, R. Viertl and S.M. Taheri. *A Possibilistic Analogue to Bayes Estimation with Fuzzy Data*. Tech. Report SM-2010-3, Institut für Statistik und Wahrscheinlichkeitstheorie. Technische Universität Wien, 2010. 103

[Arnaud *et al.* 2010] A. Arnaud, J. Bect, M. Couplet, A. Pasanisi and E. Vazquez. *Évaluation d'un risque d'inondation fluviale par planification séquentielle d'expériences*. In 42èmes Journées de Statistique, Marseille (France), May 2010. 60, 85, 170

[Atchadé *et al.* 2011] Y.F. Atchadé, G. Fort, É. Moulines and P. Priouret. *Adaptive Markov Chain Monte Carlo: Theory and Methods*. In D. Barber, A. Taylan Cemgil and S. Chiappa (Eds.), Bayesian Time Series Models, pages 32–47. Cambridge University Press, 2011. 41

[Auder *et al.* 2012] B. Auder, A. de Crecy, B. Iooss and M. Marquès. *Screening and metamodeling of computer experiments with functional outputs. Application to thermal-hydraulic computations*. Reliability Engineering & System Safety, vol. 107, pages 122–131, 2012. 151

[Aven & Kvaloy 2002] T. Aven and J.T. Kvaloy. *Implementing the Bayesian paradigm in risk analysis*. Reliability Engineering and System Safety, vol. 78, pages 195–201, 2002. 108, 118

[Aven & Steen 2010] T. Aven and R. Steen. *The concept of ignorance in a risk assessment and risk management context*. Reliability Engineering & System Safety, vol. 95, pages 1117–1122, 2010. 97

[Aven & Zio 2011] T. Aven and E. Zio. *Some considerations on the treatment of uncertainties in risk assessment for practical decision making*. Reliability Engineering and System Safety, vol. 96, pages 64–74, 2011. 97, 107

[Aven 2010] T. Aven. *Some reflections on uncertainty analysis and management*. Reliability Engineering and System Safety, vol. 95, pages 195–201, 2010. 107, 120, 149

[Aven 2011] T. Aven. *Interpretations of alternative uncertainty representation in a reliability and risk analysis context*. Reliability Engineering and System Safety, vol. 96, pages 353–360, 2011. 107

[Ayyub *et al.* 2009] B.M. Ayyub, J. Foster and W.L. McGill. *Risk Analysis of a Protected Hurricane-Prone Region. I: Model Development*. Natural Hazards Review, vol. 10, no. 2, pages 38–53, 2009. 122

[Azaïs *et al.* 2010] R. Azaïs, A. Gégout-Petit, M. Touzet and C. Élegbede. *Estimation, simulation et prévision d'un modèle de propagation de fissures par des processus markoviens déterministes par morceaux*. In 17ème Congrès Lambda-Mu, La Rochelle (France), October 2010. 47

[Azar & Hauglustaine 2001] S. Azar and J.M. Hauglustaine. *Multicriteria and multiple actors tool aiding to optimize building envelope at the architectural sketch design*. Informatica, vol. 12, no. 1, pages 3–24, 2001. 139

[Bachoc 2013] F. Bachoc. *Parametric estimation of covariance function in Gaussian-process based Kriging models. Application to uncertainty quantification for computer experiments*. PhD thesis, Université Paris 7, 2013. 150

[Balcomb & Curtner 2000] J.D. Balcomb and A. Curtner. *Multi-criteria decision-making process for buildings*. In 2000 Energy Conversion Engineering Conference and Exhibit, pages 528–535, Las Vegas (NE, USA), 2000. 140

[Baraldi *et al.* 2011] P. Baraldi, N. Pedroni, E. Zio, E. Ferrario, A. Pasanisi and M. Couplet. *Monte Carlo and fuzzy interval propagation of hybrid uncertainties on a risk model for the design of a flood protection dike*. In ESREL 2011 Conference, pages 2167–2175, Troyes (France), September 2011. 90, 99, 101, 169

[Barate 2013] R. Barate. *Calcul Haute Performance avec OpenTURNS*. In Journée du GDR MASCOT NUM Quantification d'incertitude et calcul intensif, Grenoble (France), March 2013. 79

[Bárdossy & Duckstein 1995] A. Bárdossy and L. Duckstein. Fuzzy rule-based modeling with applications to geophysical, biological and engineering systems. CRC Press, 1995. 90, 93

[Bartholomew 1973] D.J. Bartholomew. Stochastic Models for Social Processes, second edition. Wiley, 1973. 36

[Basile *et al.* 2007] O. Basile, P. Dehombreux and F. Riane. *Evaluation of the uncertainty affecting reliability models*. Journal of Quality in Maintenance Engineering, vol. 13, pages 137–151, 2007. 42

[Baudin & Martinez 2010] M. Baudin and J.M. Martinez. *Polynômes de chaos sous Scilab via la librairie NISP*. In 42èmes Journées de Statistique, Marseille (France), May 2010. 86

[Baudin & Pasanisi 2012] M. Baudin and A. Pasanisi. *Traitement des incertitudes dans la simulation numérique descircuits de refroidissement des CNPE avecaéroréfrigérant: couplage entre les logiciels CooliSS ST et Open TURNS*. Tech. Report H-T57-2012-02834-FR, EDF R&D, 2012. 60, 171

[Baudin *et al.* 2013] M. Baudin, A. Pasanisi and F. Gaudier. *Comparaison entre Uranie et OpenTURNS*. Tech. Report H-T57-2013-03841-FR, EDF R&D, 2013. 60, 171

[Baudrit *et al.* 2006] C. Baudrit, D. Dubois and D. Guyonnet. *Joint Propagation and Exploitation of Probabilistic and Possibilistic Information in Risk Assessment*. IEEE Transactions on Fuzzy Systems, vol. 14, no. 5, pages 593–608, 2006. 98, 99

[Baudrit *et al.* 2008] C. Baudrit, D. Dubois and N. Perrot. *Representing parametric probabilistic models tainted with imprecision*. Fuzzy Sets and System, vol. 159, no. 15, pages 1913–1928, 2008. 97, 99, 100

[Baudrit 2005] C. Baudrit. *Représentation et propagation de connaissances imprécises et incertaines : application à l'évaluation des risque liés aux sites et aux sols pollués*. PhD thesis, Université Toulouse 3 Paul Sabatier, 2005. 98, 99, 100, 101, 102, 158

[Bayle 1946] G. Bayle. Cours de statique graphique. Eyrolles, 1946. 4

[Baysse *et al.* 2012] C. Baysse, D. Bihannic, A. Gégout-Petit, M. Prenat and J. Saracco. *Detection of a degraded operating mode of optronic equipment using Hidden Markov Model*. In ESREL 2012 / PSAM 11 Conference, Helsinki (Finland), June 2012. 31

[Bebbington *et al.* 2012] M. Bebbington, C.D. Lai, M. Wellington and R. Zitikis. *The discrete additive Weibull distribution: A bathtub-shaped hazard for discontinuous failure data*. Reliability Engineering and System Safety, vol. 106, pages 37 – 44, 2012. 19

[Bect *et al.* 2012] J. Bect, D. Ginsbourger, L. Li, V. Picheny and E. Vazquez. *Sequential design of computer experiments for the estimation of a probability of failure*. Statistics and Computing, vol. 22, no. 3, pages 773–793, 2012. 78, 84, 150

[Benassai 1973] E. Benassai. *Costruzioni marittime*. In M. Lenti (Ed.), Enciclopedia dell'Ingenieria. ISEDI, 1973. 4

[Benton & Krishnamoorthy 2003] D. Benton and K. Krishnamoorthy. *Computing discrete mixtures of continuous distributions: noncentral chisquare, noncentral t and the distribution of the square of the sample multiple correlation coefficient*. Computational Statistics and Data Analysis, vol. 43, pages 249–267, 2003. 79

[Berge-Thierry *et al.* 2003] C. Berge-Thierry, F. Cotton, O. Scotti, D.A. Griot-Pommera and Y. Fukushima. *New empirical response spectral attenuation laws for moderate European earthquakes*. Journal of Earthquake Engineering, vol. 7, no. 2, pages 193–222, 2003. 129

[Berger 1985] J. O. Berger. Statistical Decision Theory and Bayesian Analysis. Springer, 1985. 107, 108, 109, 111

[Bernardo & Smith 1994] J. M. Bernardo and A. F. M. Smith. Bayesian Theory. Wiley, 1994. 32, 96, 107, 108

[Bernardo 1996] J.M. Bernardo. *The concept of exchangeability and its applications*. Far East Journal of Mathematical Sciences, vol. 4, pages 111–121, 1996. 96

[Bernardo 2011] J.M. Bernardo. *Modern Bayesian Inference: Foundations and Objective Methods*. In P.S. Bandyopadhyay and M.R. Forster (Eds.), Handbook of the Philosophy of Science. Volume 7: Philosophy of Statistics, pages 263–306. Elsevier, 2011. 108

[Bernier & Sabaton 1972] J. Bernier and C. Sabaton. *Un modèle stochastique au court terme pour la DBO et l'oxygène dissous en rivière et ses applications*. Tech. Report HC 040/1, EDF R&D, 1972. 71

References

[Bernier 1975] J. Bernier. *La prise en compte des phénomènes aléatoires dans les modèles mathématiques hydrauliques et hydrométéorologiques de l'environnement*. Tech. Report C45/75.15, EDF R&D, 1975. 72

[Bernier 1980] J. Bernier. *Utilisation statistique des codes de calcul en mécanique des fluides*. Tech. Report E40-80-08, EDF R&D, 1980. 73, 120

[Bernier 2003] J. Bernier. *Décisions et comportement des décideurs face au risque hydrologique*. Hydrological Sciences Journal, vol. 48, no. 3, pages 301–316, 2003. 121

[Blatman & Sudret 2010] G. Blatman and B. Sudret. *An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis*. Probabilistic Engineering Mechanics, vol. 25, no. 2, pages 183–197, 2010. 77, 79

[Blatman & Sudret 2011] G. Blatman and B. Sudret. *Principal component analysis and Least Angle Regression in spectral stochastic finite element analysis*. In ICASP11 Conference, Zurich (Switzerland), August 2011. 60

[Blatman 2009] G. Blatman. *Chaos polynomial creux et adaptatif pour la propagation d'incertitudes et l'analyse de sensibilité*. PhD thesis, Université Blaise Pascal Clermont-Ferrand 2, 2009. 60, 79

[Blondeau *et al.* 2002] P. Blondeau, M. Sperandio and F. Allard. *Multicriteria analysis of ventilation in summer period*. Building and Environment, vol. 37, pages 165–176, 2002. 139

[Bois 2001] F.Y. Bois. *Applications of population approaches in toxicology*. Toxicology Letters, vol. 120, pages 385–394, 2001. 64

[Boreux *et al.* 2010] J.J. Boreux, E. Parent and J. Bernier. Pratique du calcul bayésien. Springer, 2010. 118

[Bousquet 2005] N. Bousquet. *Eliciting prior distributions for Weibull inference in an industrial context*. Communications in Dependability and Quality Management, vol. 8, pages 12–19, 2005. 20

[Bousquet 2008] N. Bousquet. *Diagnostics of prior-data agreement in applied Bayesian analysis*. Journal of Applied Statistics, vol. 35, pages 1011–1029, 2008. 20

[Bousquet 2010] N Bousquet. *Eliciting vague but proper maximal entropy priors in Bayesian experiments*. Statistical Papers, vol. 51, no. 3, pages 613–628, 2010. 20

[Bousquet 2012] N. Bousquet. *Accelerated Monte Carlo estimation of exceedance probabilities under monotonicity constraints*. Annales de la Faculté des sciences de Toulouse, Sér. 6, vol. 21, no. 3, pages 557–591, 2012. 60

[Box & Draper 1987] G.E. Box and N.R. Draper. Empirical model building and response surfaces. Wiley, 1987. 76

[Bracquemond & Gaudoin 2003] C. Bracquemond and O. Gaudoin. *A survey on discrete lifetime distributions*. International Journal on Reliability, Quality, and Safety Engineering, vol. 10, no. 1, pages 69–98, 2003. 15, 16, 17, 19, 21

[Bracquemond 2001] C. Bracquemond. *Modélisation stochastique du vieillissement en temps discret*. PhD thesis, Institut National Polytechnique de Grenoble, 2001. 16

[Brans *et al.* 1984] J.P. Brans, B. Mareschal and P. Vincke. *PROMETHEE: a new family of outranking methods in multicriteria analysis*. In Brans J.P. (Ed.), Operations Research '84, pages 408–421. North Holland Publishing Co, 1984. 137

[Brochot *et al.* 2007] C. Brochot, T.J. Smith and F.Y. Bois. *Development of a physiologically based toxicokinetic model for butadiene and four major metabolites in humans: Global sensitivity analysis for experimental design issues*. Chemico-Biological Interactions, vol. 167, no. 3, pages 168–183, 2007. 64

[Brooks & Gelman 1998] S. Brooks and A. Gelman. *General Methods for Monitoring Convergence of Iterative Simulations*. Journal of Computational and Graphical Statistics, vol. 47, pages 434–455, 1998. 38

[Bucher 1988] C.G. Bucher. *Adaptive sampling - an iterative fast Monte Carlo procedure*. Structural Safety, vol. 5, no. 2, pages 119–126, 1988. 70

[Buffa *et al.* 2012] A. Buffa, Y. Maday, T. Patera, C. Prud'Homme and G. Turinici. *A priori convergence of the Greedy algorithm for the parametrized reduced basis method*. ESAIM: Mathematical Modelling and Numerical Analysis, vol. 46, no. 3, pages 595–603, 2012. 85

[Bukowski & Goble 1995] J.V. Bukowski and W.M. Goble. *Using Markov models for safety analysis of programmable electronic systems*. ISA Transactions, vol. 34, pages 193–198, 1995. 31

[Butucea *et al.* 2013] C. Butucea, J.F. Delmas, A. Dutfoy and R. Fischer. Maximum entropy copula with given diagonal section. ArXiv e-prints, 1312.5219, 2013. 60

[Caccavelli & Gügerli 2002] D. Caccavelli and H. Gügerli. *TOBUS - an European diagnosis and decision-making tool for office building upgrading*. Energy and Buildings, vol. 34, pages 113–119, 2002. 140

[Cacuci 1981] D.G. Cacuci. *Sensitivity theory for nonlinear systems I. Nonlinear functional analysis approach*. Journal of Mathematical Physics, vol. 22, pages 2794–2802, 1981. 70

[Camy 2013] R. Camy. *An approach to evaluate uncertainties in complex CFD simulations*. In Workshop SFdS-IMdR-GdR MASCOT NUM Validation de codes numériques, Paris (France), November 2013. 151

[Canavese & Ortega 2013] D. Canavese and N.R.S. Ortega. *A proposal of a fuzzy rule-based system for the analysis of health and health environments in Brazil*. Ecological Indicators, vol. 34, pages 7–14, 2013. 90

[Carter & Kohn 1994] C. K. Carter and P. Kohn. *On Gibbs sampling for state space models*. Biometrika, vol. 81, pages 541–553, 1994. 38

[Caruso & Pasanisi 2013] A. Caruso and Pasanisi. *Incertitudes en simulation numérique à EDF : enjeux et activités*. In WorkStat 2013, Clamart (France), November 2013. 60, 169

[CASL 2011] CASL. *Consortium for Advanced Simulation of LWRs. A Project Summary*. Tech. report, US Department of Energy, 2011. 63

[Castino *et al.* 1998] F. Castino, R. Festa and C.F Ratto. *Stochastic modelling of wind velocities time series*. Journal of Wind Engineering and Industrial Aerodynamics, vol. 74-76, pages 141 – 151, 1998. 20

[Chang *et al.* 2012] N.B. Chang, C. Qi and Y. Jeffrey Yang. *Optimal expansion of a drinking water infrastructure system with respect to carbon footprint, cost-effectiveness and water demand*. Journal of Environmental Management, vol. 110, pages 194–206, 2012. 134

[Chatterjee *et al.* 2008] D. Chatterjee, G. Sutradhar and B. Oraon. *Fuzzy rule-based prediction of hardness for sintered HSS components*. Journal of Materials Processing Technology, vol. 200, no. 1-3, pages 212–220, 2008. 90

[Chauliac *et al.* 2011] C. Chauliac, J.M. Aragonés, D. Bestion, D.G. Cacuci, N. Crouzet, F.P. Weiss and M.A. Zimmermann. *NURESIM. A European simulation platform for nuclear reactor safety: Multi-scale and multi-physics calculations, sensitivity and uncertainty analysis*. Nuclear Engineering and Design, vol. 241, no. 9, pages 3416–3426, 2011. 65

[Chojnacki & Ounsy 1996] E. Chojnacki and A. Ounsy. *Description of the IPSN method for the uncertainty and sensitivity analysis and the associated software: SUNSET*. In 4th International Conference on Nuclear Engineering, New Orleans (LA, USA), March 1996. 78

[Christensen & Huffman 1985] R. Christensen and M.D. Huffman. *Bayesian Point Estimation Using the Predictive Distribution*. The American Statistician, vol. 39, no. 4, pages 319–321, 1985. 111

[Chutia *et al.* 2013] R. Chutia, S. Mahanta and D. Datta. *Non-probabilistic sensitivity and uncertainty analysis of atmospheric dispersion*. Annals of Fuzzy Mathematics and Informatics, vol. 5, no. 1, pages 213–228, 2013. 104

[Ciffroy *et al.* 2012] Ph. Ciffroy, M. Keller, A. Pasanisi and B. Richard. *Constructing species sensitivity distributions (SSD) using informative Bayesian approach*. In Colloque 2012 de la Société Française d'Ecotoxicologie Fondamentale et Appliquée, Lyon (France), July 2012. 106, 170

[Ciffroy *et al.* 2013] Ph. Ciffroy, M. Keller and A. Pasanisi. *Estimating Hazardous Concentrations by an informative Bayesian approach*. Environmental Toxicology and Chemistry, vol. 32, no. 3, pages 602–611, 2013. 106, 167

[Clarotti *et al.* 1997] A. Clarotti, A. Lannoy and H. Procaccia. *Probability risk analysis of ageing components which fail on demand : a Bayesian model.* In R.K. Penny (Ed.), Ageing of Materials and Methods for the Assessment of Lifetimes of Engineering Plant, pages 85–93. Balkema, Rotterdam, 1997. 15, 17, 26

[Cochran *et al.* 2001] J.K. Cochran, A. Murugan and V. Krishnamurthy. *Generic Markov models for availability estimation and failure characterization in petroleum refineries*. Computers & Operations Research, vol. 28, pages 1–12, 2001. 31

References

[Cole *et al.* 2005] B.F. Cole, M. Bonetti, A.M. Zaslavsky and R.D. Gelber. *A multistate Markov chain model for longitudinal categorical quality-of-life data subject to non-ignorable missingness*. Statistics in Medicine, vol. 24, pages 2317–2334, 2005. 36, 46

[Coles 2001] S. Coles. An Introduction to Statistical Modelling of Extreme Values. Springer-Verlag, 2001. 113

[Coninx 2012] A. Coninx. *Visualisation interactive de grands volumes de données incertaines: pour une approche perceptive*. PhD thesis, Université Joseph Fourier Grenoble 1, 2012. 151

[Cornell 1968] C.A. Cornell. *Engineering seismic risk analysis*. Bulletin of the Seismological Society of America, vol. 58, no. 1, pages 1583–1606, 1968. 128

[Couplet *et al.* 2010] M. Couplet, L. Lebrusquet and A. Pasanisi. *Caractérisation des coefficients de Strickler d'un fleuve par inversion probabiliste*. In 42èmes Journées de Statistique, Marseille (France), May 2010. 60, 87, 170

[Cox *et al.* 2012] M. Cox, A. Forbes, P. Harris and C. Matthews. *Numerical Aspects in the Evaluation of Measurement Uncertainty*. In A.M. Dienstfrey and R.F. Boisvert (Eds.), Uncertainty Quantification in Scientific Computing, pages 180–194. Springer, 2012. 66

[Craiu 2011] R. Craiu. *Tuning of Markov Chain Monte Carlo Algorithms using Copulas*. UPB Scientific Bulletin Series A, vol. 73, pages 5–12, 2011. 41

[Crestaux 2011] T. Crestaux. *Méthode adaptative d'intégration multi-dimensionnelle et sélection d'une base de polynômes de chaos*. PhD thesis, Université Paris 13, 2011. 77

[Cronvall & Männistö 2009] O. Cronvall and I. Männistö. *Combining discrete-time Markov processes and probabilistic fracture mechanics in RI-ISI risk estimates*. International Journal of Pressure Vessels and Piping, vol. 86, pages 732–737, 2009. 31

[Cunningham 2012] M.A. Cunningham. *Considerations of Uncertainty in Regulatory Decision Making*. In A.M. Dienstfrey and R.F. Boisvert (Eds.), Uncertainty Quantification in Scientific Computing, pages 15–26. Springer, 2012. 66

[Damblin *et al.* 2012a] G. Damblin, M. Keller, A. Pasanisi, I. Zentner and É. Parent. *Comment décider d'une courbe de fragilité ? Approche décisionnelle bayésienne pour traiter les incertitudes dans un contexte industriel*. In 44èmes Journées de Statistique, Bruxelles (Belgium), May 2012. 106, 170

[Damblin *et al.* 2012b] G. Damblin, M. Keller, A. Pasanisi, I. Zentner and É. Parent. *How to Choose a Fragility Curve? Bayesian Decision Theory Applied to Uncertainty Analysis in an Industrial Context*. In ENBIS12 Conference, Ljubljana (Slovenia), September 2012. 106, 169

[Damblin *et al.* 2013a] G. Damblin, M. Couplet and B. Iooss. *Numerical studies of space filling designs: optimization of Latin hypercube samples and subprojection properties*. Journal of Simulation, vol. 7, pages 276–289, 2013. 150

[Damblin *et al.* 2013b] G. Damblin, M. Keller, A. Pasanisi, É. Parent, P. Barbillon and J. Bernier. *A Bayes decision approach to code validation in an industrial context*. In 7th International Conference on Sensitivity Analysis of Model Output, Nice (France), July 2013. 60, 150, 169

[Damblin *et al.* 2014a] G. Damblin, P. Barbillon, M. Keller, A. Pasanisi and E. Parent. *Bayesian validation of a computer model for the energy consumption of a building, with application to optimal electric bill pricing*. In 2014 International Society for Bayesian Analysis World Meeting, Cancun (Mexico), July 2014. 150, 151, 168

[Damblin *et al.* 2014b] G. Damblin, P. Barbillon, M. Keller, A. Pasanisi and É. Parent. *Plans d'expérience séquentiels pour la calibration de modèles numériques coûteux*. In 46èmes Journées de Statistique, Rennes (France), June 2014. 150, 169

[Damblin *et al.* 2014c] G. Damblin, M. Keller, A. Pasanisi, P. Barbillon and É. Parent. *Approche décisionnelle bayésienne pour estimer une courbe de fragilité*. Journal de la Société Française de Statistique (accepted), 2014. 106, 167

[Datta 2011] D. Datta. *Non-Probabilistic Uncertainty Analysis of Analytical and Numerical Solution of Heat Conduction*. International Journal of Energy, Information and Communications, vol. 2, no. 4, pages 143–156, 2011. 104

[De Bonis *et al.* 2002] P. De Bonis, G. Fattoruso, A. Pagano and F. Pasanisi. *Transport of pollutants in coastal areas: a comparison among different mathematical models*. In 2002 Annual Conference of the International Solid Waste Association, Istanbul (Turkey), July 2002. 8

[de Finetti 1930a] B. de Finetti. *Funzione caratteristica di un fenomeno aleatorio*. Memorie della R. Accademia Nazionale dei Lincei, vol. 4, no. 5, pages 86–133, 1930. 96

[de Finetti 1930b] B. de Finetti. *La prévision : ses lois logiques, ses sources subjectives*. Annales de l'Instiut Henri Poincaré, vol. 7, pages 1–68, 1930. 97

[de Rocquigny *et al.* 2008] É. de Rocquigny, N. Devictor and S. Tarantola (Eds.). Uncertainty in Industrial Practice. Wiley, 2008. 67, 81

[de Rocquigny 2006] É. de Rocquigny. *La maîtrise des incertitude dans un contexte industriel. Seconde partie : revue des méthodes de modélisation statistique, physique et numérique*. Journal de la Société Française de Statistique, vol. 3, pages 73–106, 2006. 119

[Decker 1991] K. Decker. *The Monte Carlo method in science and engineering: Theory and application*. Computer Methods in Applied Mechanics and Engineering, vol. 89, pages 463–483, 1991. 73

[Delmas *et al.* 2011] J. Delmas, A. Assire and I. Fournier. *Salome-Meca : une plate-forme au service de la simulation mécanique*. In CSMA 2011. 10ème Colloque National en Calcul des Structures, Hyères (France), May 2011. 125

[Deltout *et al.* 1999] I. Deltout, S. Richardson and J.Y. Le Hersan. *Stochastic algorithms for Markov models estimation with intermittent missing data*. Biometrics, vol. 55, pages 565–573, 1999. 34

[Dempster 1967] A.P. Dempster. *Upper and Lower Probabilities Induced by a Multivalued Mapping*. Annals of Statistics, vol. 28, pages 325–339, 1967. 158

[Déqué *et al.* 1999] F. Déqué, L. Catusse and S. Castanet. *Residentiel Sim-Fast: Software for rapid technical appraisals*. In 6th Internation IBPSA Conference, Kyoto (Japan), August 1999. 144

[Der Kiureghian 2008] A. Der Kiureghian. *Analysis of structural reliability under parameter uncertainties*. Probabilistic Engineering Mechanics, vol. 23, pages 351–358, 2008. 119

[Désenfant *et al.* 2007] M. Désenfant, N. Fischer, B. Blanquart and N. Bédiat. *Évaluation de l'incertitude en utilisant les simulations de Monte Carlo*. In 13ème Congrès International de Métrologie, Lille (France), June 2007. 66

[Di Natale *et al.* 2000] M. Di Natale, L. Duckstein and A. Pasanisi. *Forecasting pollutants transport in river by a fuzzy rule-based model*. In Belgium Fuzzy II Conference, Mons (Belgium), March 2000. 90, 169

[Di Natale *et al.* 2001] M. Di Natale, L. Duckstein and A. Pasanisi. *Forecasting pollutants transport in a river by a fuzzy rule-based model*. Belgian Journal of Operations Research, Statistics and Computer Science, vol. 41, no. 3-4, pages 129–138, 2001. 90, 167

[Diebold *et al.* 1994] F.X. Diebold, J.H. Lee and G. Weinbach. *Regime Switching with Time-Varying Transition Probabilities*. In C. Hargreaves (Ed.), Nonstationary Time Series Analysis and Cointegration, pages 283–302. Oxford University Press, 1994. 31

[Ditlevsen & Madsen 1996] O. Ditlevsen and H.O. Madsen. Structural reliability methods. Wiley, 1996. 69

[Ditlevsen *et al.* 1988] O. Ditlevsen, P. Bjerager, R. Olesen and A.M. Hasofer. *Directional simulation in Gaussian processes*. Probabilistic Engineering Mechanics, vol. 3, no. 4, pages 207–217, 1988. 70

[DOD 1981] DOD. *Definitions of terms for reliability and maintainability. Military Standard MIL-STD-721C*. Department of Defense, Washington, DC, 1981. 14

[Dodgson *et al.* 2009] J.S. Dodgson, M. Spackman, A. Pearman and L.D. Phillips. *Multi-criteria analysis: a manual*. Tech. report, Department for Communities and Local Government: London, 2009. 133

[Dolezal 2008] W.R. Dolezal. *Success Factors for Digital Mock-ups (DMU) in complex Aerospace Product Development*. PhD thesis, Technische Universität München, 2008. 63

[Dolinski 1983] K. Dolinski. *First-order second-moment approximation in reliability of structural systems: critical review and alternative approach*. Structural Safety, vol. 1, no. 3, pages 211–231, 1983. 70, 75

[Doornik 2005] J.A. Doornik. *An Improved Ziggurat Method to Generate Normal Random Samples*. Tech. report, Nuffield College, Oxford University, 2005. 79

[Douard & Iooss 2013] F. Douard and B. Iooss. *Dealing with uncertainty in technical and economic studies of investment strategy optimization*. In 7th International Conference on Sensitivity Analysis of Model Output, Nice (France), July 2013. 151

References

[Downing *et al.* 1985] D.J. Downing, R.H. Gardner and F.O. Hoffman. *An Examination of Response-Surface Methodologies for Uncertainty Analysis in Assessment Models*. Technometrics, vol. 27, no. 2, pages 151–163, 1985. 73, 76

[Doyeux *et al.* 2012] V. Doyeux, V. Chabannes, C. Prud'Homme and I. Mourad. *Simulation of vesicle using level set method solved by high order finite element*. ESAIM: Proceedings, vol. 38, pages 335–347, 2012. 63

[Dubois & Prade 2011] D. Dubois and H. Prade. *Possibility Theory and its Applications: Where Do we Stand?* Mathware and Soft Computing, vol. 18, no. 1, pages 18–31, 2011. 98

[Dubois *et al.* 2004] D. Dubois, L. Foulloy, G. Mauris and H. Prade. *Probability-Possibility Transformations, Triangular Fuzzy Sets and Probabilistic Inequalities*. Reliable Computing, vol. 10, pages 273–297, 2004. 100

[Dubois 2006] D. Dubois. *Possibility theory and statistical reasoning*. Computational Statistics and Data Analysis, vol. 51, pages 47–69, 2006. 97, 98

[Duckstein & Bogardi 1979] L. Duckstein and I. Bogardi. *Uncertainties in lake management.* In Hipel K.W. Unn T.E. McBean E.A. (Ed.), Reliability in Water Resources Management, pages 253–279. Water Resources Publications, 1979. 31

[Duffau 1903] Duffau. Guide du constructeur ou analyse des prix raisonnée des travaux de bâtiments et ouvrages d'art. Gounouilhou Éd., 1903. 4

[Dupuis & Schwarz 2007] J. Dupuis and C.J. Schwarz. *A Bayesian approach to the Multistate Jolly-Seber capture-recapture model*. Biometrics, vol. 63, pages 1015–1022, 2007. 31, 35

[Dupuis 1995] J. Dupuis. *Bayesian Estimation of Movement and Survival Probabilities from Capture-Recapture Data*. Biometrika, vol. 82, pages 761–772, 1995. 31, 32, 36

[Durrande *et al.* 2012] N. Durrande, D. Ginsbourger and O. Roustant. *Additive covariance kernels for high-dimensional Gaussian process modeling*. Annales de la Faculté des sciences de Toulouse, Sér. 6, vol. 21, no. 3, pages 481–499, 2012. 78

[Dutfoy & Lebrun 2009] A. Dutfoy and R. Lebrun. *Practical approach to dependence modelling using copulas*. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, vol. 223, no. 4, pages 347–361, 2009. 75

[Dutfoy *et al.* 2009] A. Dutfoy, I. Dutka-Malen, A. Pasanisi, R. Lebrun, F. Mangeant, J. Sen Gupta, M. Pendola and T. Yalamas. *OpenTURNS, an Open Source initiative to Treat Uncertainties, Risks'N Statistics in a structured industrial approach*. In 41èmes Journées de Statistique, Bordeaux (France), May 2009. 60, 79, 170

[Dutfoy *et al.* 2012] A. Dutfoy, B. Iooss and R. Lebrun. *Probabilistic modelling of variables with inequality constraints*. In ESREL 2012 / PSAM 11 Conference, Helsinki (Finland), June 2012. 60

[Eckert *et al.* 2009] N. Eckert, É. Parent, T. Faug and M. Naaim. *Bayesian optimal design of an avalanche dam using a multivariate numerical avalanche model*. Stochastic Environmental Research and Risk Assessment, vol. 23, no. 8, pages 1123–1141, 2009. 117

[Eggenberger & Pólya 1923] F. Eggenberger and G. Pólya. *Über die Statistik verketter Vorgänge*. Zeitschrift für Angewandte Mathematik und Mechanik, no. 1, pages 279–289, 1923. 15, 16

[El Ghaoui 2002] L. El Ghaoui. *Inversion error, condition number, and approximate inverses of uncertain matrices*. Linear algebra and its applications, vol. 343, pages 171–193, 2002. 41

[El-Nashar 2008] A.M. El-Nashar. *Optimal design of a cogeneration plant for power and desalination taking equipment reliability into consideration*. Desalination, vol. 229, pages 21–32, 2008. 31

[Eldred *et al.* 1996] M.S. Eldred, D.E. Outka, W.J. Bohnhoff, W.R. Witkowski, V.J. Romero, E.R. Ponslet and K.S. Chen. *Optimization of Complex Mechanics Simulations with Object-Oriented Software Design*. Computer Modeling and Simulation in Engineering, vol. 1, no. 3, pages 323–352, 1996. 78

[EPRI 1994] EPRI. *Methodology for Developing Seismic fragilities*. Tech. Report TR-103959, Electric Power Research Institute, 1994. 123

[Erbacci *et al.* 2013] G. Erbacci, F. Bodin, V. Bergeaud, A. Pasanisi, S. McIntosh-Smith, T. Ludwig, F. Cappello, C. Cavazzoni and M.C. Sawley. *WP5 First Intermediate Report: Cross Cutting Issues Working Groups*. Tech. Report D 5.1, European Exascale Software Initiative - EESI2 Project, 2013. 150, 171

[Faissol *et al.* 2009] D.M. Faissol, P.M. Griffin and J.L. Swann. *Bias in Markov models of disease*. Mathematical Biosciences, vol. 220, no. 2, pages 143–156, 2009. 31

[Faivre *et al.* 2013] R. Faivre, B. Iooss, S. Mahévas, D. Makowski and H. Monod. Analyse de sensibilité et exploration de modèles: Application aux sciences de la nature et de l'environnement. Quae Éditions, 2013. 60, 67, 75

[Fernholm 2013] A. Fernholm. *The Nobel Prize in Chemistry 2013 - Popular Information*. Tech. report, The Royal Swedish Academy of Sciences, 2013. 62

[Ferson & Tucker 2006] S. Ferson and W.T. Tucker. *Sensitivity in Risk Analyses with Uncertain Numbers*. Tech. Report SAND2006-2801, SANDIA National Laboratories, 2006. 104, 158

[Fischer *et al.* 1979] H.B. Fischer, J.E. List, R.C. Koh, J. Imberger and N.H. Brooks. Mixing in Inland and Coastal Waters. Academic Press, 1979. 91, 92

[Fischer 1968] H.B. Fischer. *Methods for Predicting Dispersion Coefficients in Natural Streams, with Applications to Lower Reaches of the Green and Duwamish Rivers, Washington*. Tech. Report Geological Survey Professional Paper 582-A, United States Department of the Interior, 1968. 91

[Fletcher 1990] C.A.J. Fletcher. Computational Fluid Dynamics. Spinger, 1990. 92

[Flory-Celini 2008] C. Flory-Celini. *Modélisation et positionnement de solutions bioclimatiques dans le bâtiment résidentiel existant*. PhD thesis, Université Claude Bernard Lyon 1, 2008. 140, 146

[Flourentzou & Roulet 2002] F. Flourentzou and C.A. Roulet. *Elaboration of retrofit scenarios*. Energy and Buildings, vol. 34, pages 185–192, 2002. 140

[Fortin *et al.* 2009] M. Fortin, S. Bédard, J. DeBlois and S. Meunier. *Assessing and testing prediction uncertainty for single tree-based models: A case study applied to northern hardwood stands in southern Québec, Canada*. Ecological Modelling, vol. 220, no. 20, pages 2770 – 2781, 2009. 20

[Frisch 1968] U. Frisch. *Wave propagation in random media*. In A.T. Bharucha-Reid (Ed.), Probabilistic methods in applied mathematics, volume 1, pages 75–198. Academic Press, 1968. 71

[Fruth *et al.* 2014] J. Fruth, O. Roustant and S. Kuhnt. Sequential designs for sensitivity analysis of functional inputs in computer experiments. HAL preprints 00943509, 2014. 151

[Fu 2012] S. Fu. *Inversion probabiliste bayésienne en analyse d'incertitude*. PhD thesis, Université Paris Sud, 2012. 60, 100

[Fuertes & Kalotychou 2007] A.M. Fuertes and E. Kalotychou. *On sovereign credit migration: A study of alternative estimators and rating dynamics*. Computational Statistics and Data Analysis, vol. 51, pages 3448–3469, 2007. 31

[Fuh 1993] C.D. Fuh. *Statistical inquiry for Markov chains by bootstrap method*. Statistica Sinica, vol. 3, pages 53–66, 1993. 32

[Galambosi *et al.* 1999] A. Galambosi, L. Duckstein, E. Özelkan and I. Bogardi. *Fuzzified effect of ENSO and macrocirculation patterns on precipitation: an Arizona case study*. International Journal of Climatology, vol. 19, pages 1411–1426, 1999. 90

[Gamboa 2013] F. Gamboa. *Sobol pick freeze (and FAST) methods in the Costa BRAVA sauce*. In 7th International Conference on Sensitivity Analysis of Model Output, Nice (France), July 2013. 87

[Garbade & Dolezal 2007] R. Garbade and W.R. Dolezal. *DMU@Airbus. Evolution of the Digital Mock-up (DMU) at Airbus to the Centre of Aircraft Development*. In F.L. Krause (Ed.), The Future of Product Development, pages 3–12. Springer, 2007. 63

[Garcia 2013] E. Garcia. *AIRBUS Virtual Hybrid Testing Framework: focus on V&V concerns*. In Workshop SFdS-IMdR-GdR MASCOT NUM Validation de codes numériques, Paris (France), November 2013. 63

[Garnero & Montgomery 2006] M.A. Garnero and N. Montgomery. *Pronostic de la profondeur de fissuration d'un rotor de turbine*. In 15ème Congrès Lambda-Mu, Lille (France), October 2006. 47

[Gaudier *et al.* 2011] F. Gaudier, A. Pasanisi and A. Dutfoy. *Analyse d'incertitudes et simulation dans le domaine de l'énergie*. In 28ème Forum ORAP, Paris (France), October 2011. 60, 168

References

[Gaudier 2010] F. Gaudier. *URANIE : The CEA/DEN Uncertainty and Sensitivity platform*. In Sixth International Conference on Sensitivity Analysis of Model Output, Milano (Italy), July 2010. 78

[Gebrezgabher *et al.* 2014] S.A. Gebrezgabher, M.P.M. Meuwissen and A. Oude Lansink. *A multiple criteria decision making approach to manure management systems in the Netherlands*. European Journal of Operational Research, vol. 232, no. 3, pages 643–653, 2014. 134

[Gégout-Petit 2012] A. Gégout-Petit. *Contribution à la statistique des processus : modélisation et applications*. HDR Qualification Thesis, Université Bordeaux 1, 2012. 47

[Geisser 1971] S. Geisser. *The Inferential Use of Predictive Distributions*. In V.P. Godambe and D.A. Sprott (Eds.), Foundations of Statistical Inference, pages 456–469. Holt, Rinehart, and Winston, 1971. 111

[Gelman *et al.* 2004] A. Gelman, J.B. Carlin, H.S. Stern and D.B. Rubin. Bayesian data analysis, second edition. Chapman & Hall, 2004. 33

[Genest & Favre 2007] C. Genest and A.C. Favre. *Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask*. Journal of Hydrologic Engineering, vol. 12, pages 347–368, 2007. 41, 74

[Genest *et al.* 2006] C. Genest, J.F. Quessy and B. Rémillard. *Goodness-of-fit procedures for copula models based on the probability integral transformation*. Scandinavian Journal of Statistics, vol. 33, pages 337–366, 2006. 41, 46

[Gentle *et al.* 2004] J.E. Gentle, W. Härdle and Y. Mori (Eds.). Handbook of computational statistics. Springer-Verlag, 2004. 38

[Gentleman *et al.* 1994] R.C. Gentleman, J.F. Lawless, Lindsey J.C. and Yan P. *Multi-state Markov models for analyzing incomplete disease history data with illustration for HIV disease*. Statistics in Medicine, vol. 13, pages 805–821, 1994. 31

[Ghanem & Spanos 1990] R. Ghanem and P.D. Spanos. *Polynomial chaos in stochastic finite elements*. Journal of Appled Mechanics, vol. 57, pages 197–202, 1990. 70, 76

[Gilks & Roberts 1996] W.R. Gilks and G.O. Roberts. *Strategies for improving MCMC*. In W.R. Gilks, S. Richardson and D.J. Spiegelhalter (Eds.), Markov Chain Monte Carlo in Practice, pages 89–114. Chapman & Hall, 1996. 38

[Girard & Parent 2004] P. Girard and É. Parent. *The deductive phase of statistical analysis via predictive simulations: test, validation and control of a linear model with autocorrelated errors representing a food process*. Journal of Statistical Planning and Inference, vol. 124, pages 99–120, 2004. 32

[Goffart *et al.* 2013] J. Goffart, T. Mara and É Wurtz. *Impact of weather data on the performance evaluation of a passive house: an application of sensitivity analysis*. In 7th International Conference on Sensitivity Analysis of Model Output, Nice (France), July 2013. 151

[Gouno *et al.* 2011] E. Gouno, L. Courtrai and M. Fredette. *Estimation from aggregate data*. Computational Statistics and Data Analysis, vol. 55, pages 615–626, 2011. 32, 37

[Grassi *et al.* 2000] M. Grassi, I. Colombo and R. Lapasin. *Drug release from an ensemble of swellable crosslinked polymer particles*. Journal of Controlled Release, vol. 68, no. 1, pages 97 – 113, 2000. 20

[Grimshaw & Alexander 2011] S.D. Grimshaw and W.P. Alexander. *Markov chain models for delinquency: Transition matrix estimation and forecasting*. Applied Stochastic Models in Business and Industry, vol. 27, no. 3, pages 267–279, 2011. 31, 46

[Grinstead & Laurie Snell 1997] C.M. Grinstead and J. Laurie Snell. Introduction to Probability, 2nd revised edition. American Mathematical Society Press, 1997. 48

[Gutenberg & Richter 1944] B. Gutenberg and C.F. Richter. *Frequency of earthquakes in California*. Bulletin of the Seismological Society of America, vol. 34, no. 4, pages 185–188, 1944. 128

[Haario *et al.* 2001] H. Haario, E. Saksman and J. Tamminen. *An adaptive Metropolis algorithm*. Bernoulli, vol. 7, pages 223–242, 2001. 41

[Haberlandt *et al.* 2002] U. Haberlandt, V. Krysanova and A. Bárdossy. *Assessment of nitrogen leaching from arable land in large river basins. Part II: regionalisation using fuzzy rule based modelling*. Ecological Modelling, vol. 150, pages 277–294, 2002. 91

[Halpern & Fagin 1992] J.Y. Halpern and R. Fagin. *Two views of belief: belief as generalized probability and belief as evidence*. Artificial Intelligence, vol. 54, pages 275–317, 1992. 158

[Hamada *et al.* 2008] M.S. Hamada, A.G. Wilson and C.S. Reese. Bayesian Reliability. Springer, 2008. 15

[Helton & Johnson 1989] J.C. Helton and J.D. Johnson. *An uncertainty/sensitivity study for the station blackout sequence at a Mark I boiling water reactor*. Reliability Engineering & System Safety, vol. 26, no. 4, pages 293 – 328, 1989. 70

[Helton & Sallaberry 2012] J.C. Helton and C.J. Sallaberry. *Uncertainty and Sensitivity Analysis: From Regulatory Requirements to Conceptual Structure and Computational Implementation*. In A.M. Dienstfrey and R.F. Boisvert (Eds.), Uncertainty Quantification in Scientific Computing, pages 60–77. Springer, 2012. 66, 97, 107, 122

[Helton *et al.* 1992] J.C. Helton, J.A. Rollstin, J.L. Sprung and J.D. Johnson. *An exploratory sensitivity study with the MACCS reactor accident consequence model*. Reliability Engineering & System Safety, vol. 36, pages 137–164, 1992. 70

[Helton *et al.* 2006a] J.C. Helton, J.D. Johnson, W.L. Oberkampf and C.B. Storlie. *A Sampling Based Computational Strategy for the Representation of Epistemic Uncertainty in Model Prediction with Evidence Theory*. Tech. Report SAND2006-5557, Sandia National Laboratories, 2006. 102, 103, 157

[Helton *et al.* 2006b] J.C. Helton, J.D. Johnson, C.J. Sallaberry and C.B. Storlie. *Survey of sampling-based methods for uncertainty and sensitivity analysis*. Reliability Engineering & System Safety, vol. 91, no. 10-11, pages 1175–1209, 2006. 75

[Helton *et al.* 2011] J Helton, J.D. Johnson and C.J. Sallaberry. *Quantification of margins and uncertainties: example analyses from reactor safety and radioactive waste disposal involving the separation of aleatory and epistemic uncertainty*. Reliability Engineering and System Safety, vol. 96, pages 1014–1033, 2011. 97, 107

[Helton 1996] J. C. Helton. *Probability, conditional probability and complementary cumulative distribution functions in performance assessment for radioactive waste disposal*. Reliability Engineering and System Safety, vol. 54, no. 3, pages 145–163, 1996. 70, 108, 122

[Hewitt & Savage 1955] E. Hewitt and L.J. Savage. *Symmetric Measures on Cartesian Products*. Transactions of the American Mathematical Society, vol. 80, pages 470–501, 1955. 96

[Heymann 2010] M. Heymann. *Understanding and misunderstanding computer simulation: The case of atmospheric and climate science - An introduction*. Studies in History and Philosophy of Modern Physics, vol. 41, pages 193–200, 2010. 61

[Hickey *et al.* 2009] G. L. Hickey, P. S. Craig and A. Hart. *On The Application Of Loss Functions In Determining Assessment Factors For Ecological Risk*. Ecotoxicology and Environmental Safety, vol. 72, no. 2, pages 293–300, 2009. 118

[Higgins *et al.* 2014] A. Higgins, C. McNamara and G. Foliente. *Modelling future uptake of solar photo-voltaics and water heaters under different government incentives* . Technological Forecasting and Social Change, vol. 83, pages 142–155, 2014. 20

[Hirsch 1988] C. Hirsch. Numerical computation of internal and external flows. Wiley, 1988. 92

[Hoeffding 1948] W. Hoeffding. *A class of statistics with asymptotically normal distributions*. Annals of Mathematical Statistics, vol. 19, pages 293–325, 1948. 76

[Hornberger & Spear 1981] G.M. Hornberger and R.C. Spear. *An approach to the preliminary analysis of environmental systems*. Journal of environmental management, vol. 12, pages 7–18, 1981. 61

[Hoshi *et al.* 2011] S. Hoshi, Y. Maruyama and F. Yamazaki. *Reevaluation Method of Fragility Curves of Wooden House Based on Collected Damage Information*. In Proceedings of the Twelfth East Asia-Pacific Conference on Structural Engineering and Construction, pages 227–232, Hong Kong (China), January 2011. 123

[Huang *et al.* 2011] Y.N. Huang, A.S. Whittaker and N. Luco. *A probabilistic seismic risk assessment procedure for nuclear power plants: (I) Methodology*. Nuclear Engineering and Design, vol. 241, pages 3996–4003, 2011. 122

[Huard *et al.* 2006] D. Huard, G. Évin and A.C. Favre. *Bayesian copula selection*. Computational Statistics and Data Analysis, vol. 51, pages 809–822, 2006. 46

[IEC 1990] IEC. *International electrotechnical vocabulary, chapter 191: dependability and quality of service. Standard IEC 60050:191*. International Electrotechnical Commission, Geneva, 1990. 14

References

[Iooss & Ribatet 2009] B. Iooss and M. Ribatet. *Global sensitivity analysis of computer models with functional inputs*. Reliability Engineering and System Safety, vol. 94, pages 1194–1204, 2009. 151

[Iooss *et al.* 2010] B. Iooss, A. Pasanisi and M. Jala. *Étude bibliographique et méthodologique du traitement des incertitudes en phase de conception*. Tech. Report H-T56-2010-02503-FR, EDF R&D, 2010. 60, 171

[Iooss *et al.* 2011] B. Iooss, E. Ardillon, A. Arnaud, G. Arnaud, J. Baccou, M. Berveiller, G. Blatman, M. Couplet, A. Dutfoy, F. Gaudier, N. Gilardi, J.M. Martinez, A. Pasanisi, A.L. Popelin, Y. Richet and I. Zentner. *Benchmark entre logiciels de traitement des incertitudes*. Tech. Report F2.1.1.2, Complex System Design Lab Project, 2011. 60, 78, 171

[Iooss 2009] B. Iooss. *Contribution au traitement des incertitudes en simulation numérique : propagation d'ondes en milieu aléatoire et analyse d'expériences simulées*. HDR Qualification Thesis, Université Paul Sabatier Toulouse 3, 2009. 69

[Iooss 2011] B. Iooss. *Revue sur l'analyse de sensibilité globale de modèles numériques*. Journal de la Société Française de Statistique, vol. 152, no. 1, pages 3–25, 2011. 67, 75

[Ischwaran & James 2001] J. Ischwaran and L.F. James. *Gibbs sampling methods for stick-breaking priors*. Journal of the American Statistical Association, vol. 96, pages 161–174, 2001. 38

[ISO 1993] ISO. *Measurement of water flow in closed conduits. Meters for cold potable water. Part 1: Specifications; Norm 4064-1*. International Organization for Standardization, Geneva, 1993. 49, 50

[ISO 2006] ISO. *Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria; Norm 7730*. International Organization for Standardization, Geneva, 2006. 144

[ISO 2008] ISO. *Buildings and constructed assets, Service-life planning. Part 5: Life-cycle costing; Norm 15686-5*. International Organization for Standardization, 2008. 134

[ISO 2009] ISO. *Bases for design of structures - General principles on risk assessment of systems involving structures, Norm 13824*. International Organization for Standardization, Geneva, 2009. 122

[Ivanova *et al.* 2000] A. Ivanova, W.F. Rosenberger, S.D. Durham and N. Flournoy. *A birth and death urn for randomized clinical trials: asymptotic methods*. Sankhya: The Indian Journal of Statistics, Series B, vol. 62, pages 104–118, 2000. 16

[Jacquet *et al.* 1998] J. Jacquet, B. Bobée, J. Miquel and É. Parent. *Hommage à Jacques Bernier*. In É. Parent, P. Hubert, B. Bobée and J. Miquel (Eds.), Méthodes statistiques et approches bayésiennes en hydrologie, volume 20 of *Documents techniques en hydrologie*. UNESCO, 1998. 70

[Jaggs & Palmer 2000] M. Jaggs and J. Palmer. *Energy performance indoor environment quality retrofit - a European diagnosis and decision method for building refurbishment*. Energy and Buildings, vol. 31, pages 97–101, 2000. 140

[Jala *et al.* 2013] M. Jala, E. Conil, N. Varsier, J. Wiart, A. Hadjem, É. Moulines and C. Lévy-Leduc. *Simplified pregnant woman models for the fetus exposure assessment*. Comptes Rendus Physique, vol. 14, no. 5, pages 412–417, 2013. 63

[Jala 2013] M. Jala. *Plans d'expériences adaptatifs pour le calcul de quantiles et application à la dosimétrie numérique*. PhD thesis, Télécom ParisTech, 2013. 64

[James 1993] A. James. An introduction to water quality modeling. Wiley, 1993. 91

[Jaulin *et al.* 2001] L. Jaulin, M. Kieffer, O. Didrit and É. Walter. Applied Interval Analysis: With Examples in Parameter and State Estimation, Robust Control and Robotics. Springer, 2001. 151

[Jazi *et al.* 2010] M.A. Jazi, C.D. Lai and M.H. Alamatsaz. *A discrete inverse Weibull distribution and estimation of its parameters*. Statistical Methodology, vol. 7, no. 2, pages 121 – 132, 2010. 19

[JCGM 2008a] JCGM. *Evaluation of measurement data - Guide to the expression of uncertainty in measurement*. Tech. Report 100:2008, Joint Committee for Guides in Metrology, 2008. 66, 69, 74

[JCGM 2008b] JCGM. *Evaluation of measurement data. Supplement 1 to the "Guide to the expression of uncertainty in measurement" - Propagation of distributions using a Monte Carlo method*. Tech. Report 101:2008, Joint Committee for Guides in Metrology, 2008. 66

[Jedrzejuk & Marks 2002] H. Jedrzejuk and W. Marks. *Optimization of shape and functional structure of buildings as well as heat source utilization example*. Building and Environment, vol. 37, pages 1249–1253, 2002. 139

[Jirka *et al.* 2004] G.H. Jirka, T.B. Bleninger, R. Burrows and T. Larsen. *Management of point source discharges into rivers: where do environmental quality standards in the new EC-water framework directive apply?* International Journal of River Basin Management, vol. 2, no. 3, pages 225–233, 2004. 92

[Johnson *et al.* 2005] N.L. Johnson, S. Kotz and A.W. Kemp. Univariate discrete distributions, 3rd edition. Wiley, 2005. 16

[Jones *et al.* 1998] D.R. Jones, M. Schonlau and W.J. Welch. *Efficient global optimization of expensive black-box functions*. Journal of Global Optimization, vol. 13, no. 13, pages 455–492, 1998. 150

[Jones 2005] M.T. Jones. *Estimating Markov Transition Matrices using Proportions Data: An Application to Credit Risk*. Tech. Report Working Paper 2005/2219, International Monetary Fund, 2005. 31, 36

[Kadane *et al.* 1998] J.B. Kadane, L.J. Wolson, A. O'Hagan and K. Craig. *Papers on elicitation with discussions*. The Statistician, pages 3–53, 1998. 117

[Kadane 2011] J.B. Kadane. Principles of uncertainty. Chapman & Hall, 2011. 32

[Kaklaukas *et al.* 2005] A. Kaklaukas, E.K. Zavadskas and S. Raslanas. *Multivariant design and multiple criteria analysis of building refurbishments*. Energy and Buildings, vol. 37, pages 361–372, 2005. 143

[Kalbfleisch & Lawless 1984] J.D. Kalbfleisch and J.F. Lawless. *Least squares estimation of transition probabilities from aggregate data*. Canadian Journal of Statistics, vol. 12, pages 169–182, 1984. 37

[Kamali *et al.* 2005] M. Kamali, K. Ponnambalam and E.D. Soulis. *Hydrologic model calibration using fuzzy TSK surrogate model*. In NAFIPS 2005. Annual Meeting of the North American Fuzzy Information Processing Society, pages 799–803, Ann Arbor (MI, USA), June 2005. 90

[Kashefipour & Falconer 2002] S.M. Kashefipour and R.A. Falconer. *Longitudinal dispersion coefficients in natural channels*. Water Research, vol. 36, pages 1596–1608, 2002. 92

[Keeney & Raiffa 1993] R.L. Keeney and H. Raiffa. Decisions with Multiple Objectives: Preferences and Value Trade-Offs. Cambridge University Press, 1993. 133

[Keller *et al.* 2010a] M. Keller, É. Parent, N. Bousquet and A. Pasanisi. *Bayesian and frequentist parametric prediction of a tail probability in an industrial reliability context*. In 9th Valencia International Meeting on Bayesian Statistics, Benidorm (Spain), June 2010. 105, 169

[Keller *et al.* 2010b] M. Keller, É. Parent and A. Pasanisi. On the Role of Decision Theory in Uncertainty Analysis. ArXiv e-prints, 1009.4342, 2010. 106, 168

[Keller *et al.* 2011a] M. Keller, É. Parent and A. Pasanisi. *Should industrial uncertainty analysis be Bayesian?* In 58th World Statistics Congress of the International Statistical Institute, Dublin (Ireland), August 2011. 106, 169

[Keller *et al.* 2011b] M. Keller, A. Pasanisi, K. Ghorbel and É. Parent. *Réflexions sur l'analyse d'incertitude dans un contexte industriel: information disponible et enjeux décisionnels*. In 43èmes Journées de Statistique, Gammarth (Tunisia), May 2011. 106, 170

[Keller *et al.* 2011c] M. Keller, A. Pasanisi and É. Parent. *Réflexions sur l'analyse d'incertitudes dans un contexte industriel : Information disponible et enjeux décisionnels*. Journal de la Société Française de Statistique, vol. 152, no. 4, pages 60–77, 2011. 60, 105, 106, 120, 167

[Keller *et al.* 2011d] M. Keller, A. Pasanisi and É. Parent. *Sur l'intérêt du bayésien pour l'analyse d'incertitudesdans un contexte industriel : Information disponible et enjeux décisionnels.* In Rencontres AppliBUGS, Paris (France), December 2011. 106, 170

[Keller *et al.* 2014a] M. Keller, A. Pasanisi, G. Damblin, M. Marcilhac, T. Yalamas, R. Secanell and G. Senfaute. *A Bayesian methodology for the estimation of the failure probability of a structure submitted to seismic hazard.* In 2014 International Society for Bayesian Analysis World Meeting, Cancun (Mexico), July 2014. 106, 168

[Keller *et al.* 2014b] M. Keller, A. Pasanisi, M. Marcilhac, T. Yalamas, R. Secanell and G. Senfaute. *A Bayesian methodology applied to the estimation of earthquake recurrence parameters for seismic hazard assessment*. Submitted, 2014. 106, 168

References

[Kennedy & O'Hagan 2001] M. Kennedy and A. O'Hagan. *Bayesian calibration of computer models*. Journal of the Royal Statistical Society, Series B, vol. 63, pages 425–464, 2001. 70, 78, 150

[Kim *et al.* 2007] G. Kim, M.J. Silvapulle and P. Silvapulle. *Comparison of semiparametric and parametric methods for estimating copulas*. Computational Statistics and Data Analysis, vol. 51, pages 2836–2850, 2007. 41

[Ko & Garnier 2013] J. Ko and J. Garnier. *Multi-element stochastic spectral projection for high quantile estimation*. Journal of Computational Physics, vol. 243, pages 87–108, 2013. 85, 100

[Koutitas 1988] C.G. Koutitas. Mathematical models in coastal engineering. Pentech Press, 1988. 8

[Krzysztofowicz 1999] R. Krzysztofowicz. *Bayesian Forecasting via Deterministic Model*. Risk Analysis, vol. 19, no. 4, pages 739–749, 1999. 108, 118

[Kyburg & Smokler 1980] H.E. Kyburg and H.E. Smokler (Eds.). Studies in Subjective Probability. Krieger Publishing Co., 1980. 97

[Lai 2013] C.D. Lai. *Issues Concerning Constructions of Discrete Lifetime Models*. Quality Technology and Quantitative Management, vol. 10, no. 2, pages 251–262, 2013. 15, 16, 19

[Lamboni *et al.* 2013] M. Lamboni, B. Iooss, A.L. Popelin and F. Gamboa. *Derivative-based global sensitivity measures: general links with Sobol' indices and numerical tests*. Mathematics and Computers in Simulation, vol. 87, pages 45–54, 2013. 60

[Lapointe & Bobée 2000] S. Lapointe and B. Bobée. *Revision of possibility distributions: a Bayesian inference pattern*. Fuzzy Sets and System, vol. 116, pages 119–140, 2000. 103

[Lassen 1991] T. Lassen. *Markov modelling of the fatigue damage in welded structures under in-service inspection*. International Journal of Fatigue, vol. 13, pages 417–422, 1991. 31

[Lawless & McLeish 1984] J.F. Lawless and D.L. McLeish. *The information in aggregate data from Markov chains*. Biometrika, vol. 71, pages 419–430, 1984. 37

[Lawless 2003] J.F. Lawless. Statistical models and methods for lifetime data. 2nd ed. Wiley, 2003. 16

[Le Duy *et al.* 2013] T.D. Le Duy, L. Dieulle, D. Vasseur, C. Bérenguer and M. Couplet. *An alternative comprehensive framework using belief functions for parameter and model uncertainty analysis in nuclear probabilistic risk assessment applications*. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, vol. 227, no. 5, pages 471–490, 2013. 104

[Le Gratiet *et al.* 2013] L. Le Gratiet, C. Cannamela and B. Iooss. A Bayesian approach for global sensitivity analysis of (multifidelity) computer codes. ArXiv e-prints, 1307.2223, 2013. 78

[Le Gratiet *et al.* 2014] L. Le Gratiet, M. Couplet, B Iooss and L. Pronzato. *Planification d'expériences séquentielle pour l'analyse de sensibilité*. In 46èmes Journées de Statistique, Rennes (France), Juin 2014. 60

[Le Gratiet 2013] L. Le Gratiet. *Multi-fidelity Gaussian process regression for computer experiments*. PhD thesis, Université Paris 7, 2013. 150

[Le Mouel *et al.* 2007] A. Le Mouel, A. Pasanisi and P. Girault. *SimUK (UK Domestic Energy Model) - Design Reference Guide*. Tech. Report H-E14-2007-03370-EN, EDF R&D, 2007. 132, 171

[Lebrun & Dutfoy 2009a] R. Lebrun and A. Dutfoy. *A generalization of the Nataf transformation to distributions with elliptical copula*. Probabilistic Engineering Mechanics, vol. 24, no. 2, pages 172–178, 2009. 60, 75

[Lebrun & Dutfoy 2009b] R. Lebrun and A. Dutfoy. *An innovating analysis of the Nataf transformation from the viewpoint of copula*. Probabilistic Engineering Mechanics, vol. 24, no. 3, pages 312–320, 2009. 75

[Lebrun & Dutfoy 2009c] R. Lebrun and A. Dutfoy. *Do Rosenblatt and Nataf isoprobabilistic transformations really differ?* Probabilistic Engineering Mechanics, vol. 24, no. 4, pages 577–584, 2009. 75

[Lee *et al.* 1968] T.C. Lee, G.G. Judge and A. Zellner. *Maximum Likelihood and Bayesian Estimation of Transition Probabilities*. Journal of the American Statistical Association, vol. 63, pages 1162–1179, 1968. 32, 36, 37, 38, 39, 42

[Lemaire *et al.* 2010] M. Lemaire, A. Chateauneuf and J.C. Mitteau. Structural reliability. Wiley, 2010. 69

[Lencioni *et al.* 1979] P Lencioni, J.P. Benqué and J. Bernier. *Rejets de centrales en bord de mer. Essai de détermination de la tache thermique lorsque les courants marins sont aléatoires*. Tech. Report HE 041/79.12, EDF R&D, 1979. 72

[Liao & Sheu 2011] G.L. Liao and S.H. Sheu. *Economic production quantity model for randomly failing production process with minimal repair and imperfect maintenance*. International Journal of Production Economics, vol. 130, no. 1, pages 118–124, 2011. 20

[Liao *et al.* 2009] G.L Liao, Y.H. Chen and S.H. Sheu. *Optimal economic production quantity policy for imperfect process with imperfect repair and maintenance*. European Journal of Operational Research, vol. 195, no. 2, pages 348 – 357, 2009. 20

[Limbourg & de Rocquigny 2010] Ph. Limbourg and É. de Rocquigny. *Uncertainty analysis using evidence theory - confronting level-1 and level-2 approaches with data availability and computational constraints*. Reliability Engineering and System Safety, vol. 95, no. 5, pages 550 – 564, 2010. 60, 100, 102, 107

[Limbourg *et al.* 2010] Ph. Limbourg, É. de Rocquigny and G. Andrianov. *Accelerated uncertainty propagation in two-level probabilistic studies under monotony*. Reliability Engineering and System Safety, vol. 95, no. 9, pages 998 – 1010, 2010. 60

[Limbourg 2007] Ph. Limbourg. *Imprecise Probabilities for Predicting Dependability of Mechatronic Systems in Early Design Stages*. PhD thesis, Universität Duisburg-Essen, 2007. 102, 157

[Linsmeier & Pearson 2000] T.J. Linsmeier and N.D. Pearson. *Value at Risk*. Financial Analysts Journal, vol. 56, no. 2, pages 47–67, 2000. 67

[Little & Rubin 1987] R.J.A. Little and D.B. Rubin. Statistical analysis with missing data. Wiley, 1987. 35

[Little 1993] R.J.A. Little. *Pattern-Mixture models for multivariate incomplete data*. Journal of American Statistical Association, vol. 88, pages 125–134, 1993. 36

[Loftness & Hartkopf 2002] V. Loftness and V. Hartkopf. *Building Investment Decision Support (BIDS): Cost-Benefit Tool to Promote High Performance Components, Flexible Infrastructures and Systems Integration for Sustainable Commercial Buildings and Productive Organizations*. In The Austin Papers, pages 46–52. BuildingGreen Inc, 2002. 140

[Loizou *et al.* 2008] G. Loizou, M. Spendiff, H.A. Barton, J. Bessems, F.Y. Bois, M. Bouvier d'Yvoire, H. Buist, H.J. Clewell III, B. Meek, U. Gundert-Remy, G. Goerlitz and W. Schmitt. *Development of good modelling practice for physiologically based pharmacokinetic models for use in risk assessment: The first steps*. Regulatory Toxicology and Pharmacology, vol. 50, no. 3, pages 400–411, 2008. 62, 64

[Lopez 1996] J.A. Lopez. *Regulatory Evaluation of Value-at-Risk Models*. Tech. Report 96-51, The Wharton School, University of Pensylvania, 1996. 67

[MacRae 1977] E.C. MacRae. *Estimation of Time-Varying Markov Processes with Aggregate Data*. Econometrica, vol. 45, pages 183–198, 1977. 37

[Madsen *et al.* 1986] H.O. Madsen, S. Krenk and N. C. Lind. Methods of structural safety. Prentice Hall, 1986. 70, 75

[Mahmoud 2008] H. Mahmoud. Polya Urn Models. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, 2008. 16

[Maranzano & Krzysztofowicz 2008] C.J. Maranzano and R. Krzysztofowicz. *Bayesian reanalysis of the Challenger O-ring data*. Risk Analysis, vol. 28, no. 4, pages 1053–1067, 2008. 118, 122

[Marin & Robert 2007] J.M. Marin and C. Robert. Bayesian Core: A Practical Approach to Computational Bayesian Statistics. Springer, 2007. 34, 36, 118, 125

[Marks 1997] W. Marks. *Multicriteria optimisation of shape of energy-saving buildings*. Building and Environment, vol. 32, no. 4, pages 331–339, 1997. 139

[Marrel *et al.* 2009] A. Marrel, B. Iooss, B. Laurent and O. Roustant. *Calculations of Sobol indices for the Gaussian process metamodel*. Reliability Engineering & System Safety, vol. 94, pages 742–751, 2009. 78

[Marrel *et al.* 2012] A. Marrel, B. Iooss, S. da Veiga and M. Ribatet. *Global sensitivity analysis of stochastic computer models with joint metamodels*. Statistics and Computing, vol. 22, pages 833–847, 2012. 60, 151

References

[Marsaglia & Tsang 2000] G. Marsaglia and W.W. Tsang. *The Ziggurat Method for Generating Random Variables*. Journal of Statistical Software, vol. 5, no. 8, pages 1–7, 2000. 79

[Marshall & Olkin 1988] A.W. Marshall and I. Olkin. *Families of multivariate distributions*. Journal of the American Statistical Association, vol. 83, pages 834–841, 1988. 41

[Marshall & Olkin 1993] A.W. Marshall and I. Olkin. *Bivariate Life Distributions from Pólya's Urn Model for Contagion*. Journal of Applied Probability, vol. 30, no. 3, pages 497–508, 1993. 16

[Marszal *et al.* 2012] A.J. Marszal, P. Heiselberg, R.L. Jensen and J. Nørgaard. *On-site or off-site renewable energy supply options? Life cycle cost analysis of a Net Zero Energy Building in Denmark*. Renewable Energy, vol. 44, pages 154–165, 2012. 135

[Matheron 1963] G. Matheron. *Principles of geostatistics*. Economic Geology, vol. 58, pages 1246–1266, 1963. 77

[Maystre *et al.* 1994] L.Y. Maystre, J. Pictet, J. Simos and B. Roy. Méthodes multicritères ELECTRE: description, conseils pratiques et cas d'application à la gestion environnementale. Presses polytechniques et universitaires romandes, 1994. 135, 137

[McGuire 1976] R.K. McGuire. *Fortran computer program for seismic risk analysis*. Tech. Report Open-File 76-67, United States Department of the Interior, 1976. 128

[McKay *et al.* 1979] M.D. McKay, R.J. Beckman and W.J. Conover. *A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code*. Technometrics, vol. 21, no. 2, pages 239–245, 1979. 70, 73

[Melchers 1990] R.E. Melchers. *Radial Importance Sampling for Strucutral Reliability*. Journal of Engineering Mechanics, vol. 116, no. 1, pages 189–203, 1990. 70

[Merheb 2013] R. Merheb. *Fiabilité des outils de prévision du comportement des systèmes thermiques complexes*. PhD thesis, Université de Bordeaux 1, 2013. 151

[Micallef *et al.* 2005] S. Micallef, C. Brochot and F.Y. Bois. *L'analyse statistique bayésienne de données toxicocinétiques*. Environnement, Risques & Santé, vol. 4, no. 1, pages 21–34, 2005. 64

[Michal 2011] R. Michal. *Doug Kothe: CASL and the virtual reactor*. Nuclear News, vol. 54, no. 3, pages 88–90, 2011. 63

[Minka 2003] T. Minka. *Bayesian inference, entropy, and the multinomial distribution*. Online tutorial, Microsoft Research, 2003. http://research.microsoft.com/en-us/um/people/minka/papers/multinomial.html. 33

[Miquel 1984] J. Miquel. Guide pratique d'estimation des probabilités de crues . Eyrolles, 1984. 112

[Mira & Sargent 2003] A. Mira and D.J. Sargent. *A new strategy for speeding Markov chain Monte Carlo algorithms*. Statistical Methods and Applications, vol. 12, pages 49–60, 2003. 38

[Montginoul 2002] M. Montginoul. *La consommation d'eau des ménages en France : état des lieux*. Tech. report, UMR Gestion des Services Publics Cemagref & École Nationale du Génie de l'Eau et de l'Environnement de Strasbourg, 2002. 56

[Morio 2013] J. Morio. *Contributions à la simulation des événements rares dans les systèmes complexes*. HDR Qualification Thesis, Université Rennes 1, 2013. 69, 75

[Morris 1991] M.D. Morris. *Factorial Sampling Plans for Preliminary Computational Experiments*. Technometrics, vol. 33, pages 161–174, 1991. 70, 75

[Muehlenstaedt *et al.* 2012] T. Muehlenstaedt, O. Roustant, L. Carraro and S. Kuhnt. *Data-driven Kriging models based on FANOVA-decomposition*. Statistics and Computing, vol. 22, no. 3, pages 723 – 738, 2012. 78

[Munoz-Zuniga *et al.* 2012] M. Munoz-Zuniga, J. Garnier, E. Remy and É. de Rocquigny. *Analysis of adaptive directional stratification for the controlled estimation of rare event probabilities*. Statistics and Computing, vol. 22, no. 3, pages 809–821, 2012. 79, 100

[Munoz-Zuniga 2011] M. Munoz-Zuniga. *Méthodes stochastiques pour l'estimation contrôlée de faibles probabilités sur des modèles physiques complexes : application au domaine nucléaire*. PhD thesis, Université Paris 7, 2011. 79

[Nakagawa & Osaki 1975] T. Nakagawa and S. Osaki. *The discrete Weibull distribution*. IEEE Transactions on Reliability, vol. 5, no. 24, pages 300–301, 1975. 15, 19

[Nelsen 2006] R.B. Nelsen. An Introduction to Copulas, 2nd edition. Springer, 2006. 41

[Ng Cheng Hin & Zmeureanu 2014] J. Ng Cheng Hin and R. Zmeureanu. *Optimization of a residential solar combisystem for minimum life cycle cost, energy use and exergy destroyed*. Solar Energy, vol. 100, pages 102–113, 2014. 135

[Nikoloulopoulos & Karlis 2008] A.K. Nikoloulopoulos and D. Karlis. *Copula model evaluation based on parametric bootstrap*. Computational Statistics and Data Analysis, vol. 52, pages 3342–3353, 2008. 46

[Nott & Kohn 2005] D.J. Nott and R. Kohn. *Adaptive sampling for Bayesian variable selection*. Biometrika, vol. 92, pages 747–763, 2005. 38, 41, 46

[Oakley & O'Hagan 2004] J.E. Oakley and A. O'Hagan. *Probabilistic sensitivity analysis of complex models: a Bayesian approach*. Journal of the Royal Statitistical Society, Series B, vol. 66, pages 751–769, 2004. 78

[O'Hagan *et al.* 2006] A. O'Hagan, C.E. Buck, A. Daneshkhah, J. R. Eiser, P.H. Garthwaite, D.J. Jenkinson, J.E. Oakley and T. Rakow. Uncertain judgements: eliciting expert probabilities. Wiley, 2006. 74, 117

[OPUS 2011] OPUS. *Final Report of the ANR OPUS project*. Tech. Report D-WP0/11/03/A, OPUS Project, 2011. 61, 81, 171

[Ostergaard *et al.* 2011] M.G. Ostergaard, A.R. Ibbotson, O. Le Roux and A.M. Prior. *Virtual testing of aircraft structures*. CEAS Aeronautical Journal, vol. 1, pages 83–103, 2011. 62

[Özelkan & Duckstein 2001] E.C. Özelkan and L. Duckstein. *Fuzzy conceptual rainfall-runoff models*. Journal of Hydrology, vol. 253, no. 1-4, pages 41–68, 2001. 90

[Padgett & Spurrier 1985] W.J. Padgett and J.D. Spurrier. *Discrete failure models*. IEEE Transactions on Reliability, vol. 34, no. 3, pages 253–256, 1985. 19

[Parent & Bernier 2007] É Parent and J. Bernier. Le raisonnement bayésien: Modélisation et inférence. Springer, 2007. 32, 109, 111, 112, 121

[Parent *et al.* 1991] É. Parent, F. Lebdi and P. Hurand. *Stochastic modeling of a water resource system: analytical techniques versus synthetic approaches*. In J. Ganoulis (Ed.), Water Resources engineering Risk Assessment., pages 415–434. Springer-Verlag, 1991. 31

[Parent *et al.* 2014] É. Parent, A. Pasanisi, N. Bousquet, M. Keller and J. Bernier. *Considérations décisionnelles pour la construction d'un ouvrage de protection contre les crues*. In J.J. Droesbeke, Saporta G. and C. Thomas-Agnan (Eds.), Approches statistiques du risque. Technip, 2014. 105, 132, 167

[Parmigiani & Inoue 2009] G. Parmigiani and L. Inoue. Decision Theory: Principles and Approaches. Wiley, 2009. 107

[Pasanisi & Bieret 2007] A. Pasanisi and J. Bieret. *Réflexions autour d'un outil de définition et d'évaluation de solutions intégrées visant différents niveaux de performance énergétique pour les maisons individuelles neuves ou à rénover. Comparaison des outils existants*. Tech. Report H-E14-2007-01765-FR, EDF R&D, 2007. 132, 171

[Pasanisi & Coince 2006] A. Pasanisi and A.S. Coince. *Évaluation technico-économique de solutions multi-énergies pour le marché résidentiel*. Tech. Report H-E14-2006-04298-FR, EDF R&D, 2006. 132, 171

[Pasanisi & Dutfoy 2012] A. Pasanisi and A. Dutfoy. *An Industrial Viewpoint on Uncertainty Quantification in Simulation: Stakes, Methods, Tools, Examples*. In A.M. Dienstfrey and R.F. Boisvert (Eds.), Uncertainty Quantification in Scientific Computing, pages 27–45. Springer, 2012. 60, 61, 167

[Pasanisi & Évin 2007] A. Pasanisi and F. Évin. *Compétitivité technico-économique des solutions multi-énergies "PAC + Bois"*. Tech. Report H-E14-2007-02874-FR, EDF R&D, 2007. 132, 171

[Pasanisi & Heijmans 2007] A. Pasanisi and C. Heijmans. *Les outils de conseil en efficacité énergétique des bâtiments au service du développement commercial d'EDF-Belgium*. Tech. Report H-E14-2007-02023-FR, EDF R&D, 2007. 132, 171

[Pasanisi & Ojalvo 2007] A. Pasanisi and J. Ojalvo. *A multicriteria decision tool to improve the energy efficiency of residential houses*. In 65th Meeting of the European Working Group on Multiple Criteria Decision Aiding, Poznan (Poland), April 2007. 132, 169

References

[Pasanisi & Ojalvo 2008] A. Pasanisi and J. Ojalvo. *A multi-criteria decision tool to improve the energy efficiency of residential buildings*. Foundations of Computing and Decision Sciences, vol. 33, no. 1, pages 71–82, 2008. 132, 167

[Pasanisi & Parent 2004] A. Pasanisi and É. Parent. *Modélisation bayésienne du vieillissement des compteurs d'eau par mélange de classes d'appareils de différents états de dégradation*. Revue de Statistique Appliquée, vol. 52, no. 1, pages 39–65, 2004. 30, 52, 167

[Pasanisi & Sivadier 2004] A. Pasanisi and C. Sivadier. *Poids de la climatisation dans la facture énergétique en résidentiel. Analyse des facteurs d'influence et mesures de réduction des consommations*. Tech. Report H-E14-2004-01571-FR, EDF R&D, 2004. 132, 172

[Pasanisi *et al.* 2002] A. Pasanisi, É. Parent, P. Arnac and F. Paquet. *Describing the Ageing of Water Meters in Vivendi Water Distribution Networks by a Dynamic State Model*. In 7th Valencia International Meeting on Bayesian Statistics, Tenerife (Spain), June 2002. 30, 52, 169

[Pasanisi *et al.* 2009a] A. Pasanisi, É. de Rocquigny, N. Bousquet and É. Parent. *Some useful features of the Bayesian setting while dealing with uncertaintiesin industrial practice*. In ESREL 2009 Conference, Prague (Cech Rep.), September 2009. 102, 169

[Pasanisi *et al.* 2009b] A. Pasanisi, C. Roero, G. Celeux and E. Remy. *Quelques considérations sur l'utilisation pratique des modèles discrets de survie en fiabilité industrielle*. In 41èmes Journées de Statistique, Bordeaux (France), May 2009. 14, 170

[Pasanisi *et al.* 2010] A. Pasanisi, S. Fu and N. Bousquet. *Estimation de modèles markoviens discrets dans un cadre industriel fiabiliste à données manquantes*. In 42èmes Journées de Statistique, Marseille (France), May 2010. 30, 170

[Pasanisi *et al.* 2012a] A. Pasanisi, S. Fu and N. Bousquet. *Estimating Discrete Markov Models From Various Incomplete Data Schemes*. Computational Statistics and Data Analysis, vol. 56, no. 9, pages 2609–2625, 2012. 6, 30, 167

[Pasanisi *et al.* 2012b] A. Pasanisi, M. Keller and É. Parent. *Estimation de quantités d'intérêt en fiabilité industrielle... avec un petit coup de main de Thomas Bayes*. In 6èmes Rencontres stat. au sommet de Rochebrune, Megève (France), April 2012. 106, 170

[Pasanisi *et al.* 2012c] A. Pasanisi, M. Keller and É. Parent. *Estimation of a quantity of interest in uncertainty analysis: some help from Bayesian decision Theory*. Reliability Engineering and System Safety, vol. 100, pages 93–101, 2012. 60, 105, 106, 167

[Pasanisi *et al.* 2013a] A. Pasanisi, M. Couplet and A. Dutfoy. *Guide Méthodologique pour le Traitement des Incertitudes - Mise à jour 2013*. Tech. Report H-T57-2013-02207-FR, EDF R&D, 2013. 60, 171

[Pasanisi *et al.* 2013b] A. Pasanisi, M. Keller, M. Marcilhac, T. Yalamas, R. Secanell, C. Martin and G. Senfaute. *Bayesian assessment of seismic hazard curves*. In ENBIS13 Conference, Ankara (Turkey), September 2013. 106, 130, 169

[Pasanisi *et al.* 2013c] A. Pasanisi, C. Roero, N. Bousquet and E. Remy. *On the practical interest of discrete lifetime models in industrial reliability studies*. Submitted, 2013. 14, 168

[Pasanisi 2002] A. Pasanisi. *Modéliser le vieillissement des compteurs d'eau sur la base d'informations incomplètes*. In 1ères Rencontres stat. au sommet de Rochebrune, Megève (France), March 2002. 30, 170

[Pasanisi 2003] A. Pasanisi. *Modélisation bayésienne de la dégradation métrologique des compteurs d'eau avec la prise en compte de sources multiples d'information*. In 35èmes Journées de Statistiques, Lyon (France), June 2003. 30, 52, 170

[Pasanisi 2004a] A. Pasanisi. *Aide à la décision dans la gestion des parcs de compteurs d'eau potable*. PhD thesis, École Nationale du Génie Rural, des Eaux et des Forêts, 2004. 6, 30, 52, 168

[Pasanisi 2004b] A. Pasanisi. *Analyse technico-économique de systèmes mono-énergie et multi-énergie en résidentiel*. Tech. Report H-E14-2004-02934-FR, EDF R&D, 2004. 132, 172

[Pasanisi 2004c] A. Pasanisi. *La subjectivité dans l'approche scientifique : Aristote*. In 2èmes Rencontres stat. au sommet de Rochebrune, Megève (France), March 2004. 7, 170

[Pasanisi 2004d] A. Pasanisi. *Transitions latentes, données inadaptées et compteurs ... histoire d'une thèse*. In 2èmes Rencontres stat. au sommet de Rochebrune, Megève (France), March 2004. 30, 170

[Pasanisi 2005] A. Pasanisi. *Éléments d'aide multicritère à la décision*. Tech. Report H-E14-2005-00770-FR, EDF R&D, 2005. 132, 171

[Pasanisi 2006] A. Pasanisi. *Outil d'aide à la décision "Multi-énergies" version 1.0. Principes de modélisation et fonctionnement*. Tech. Report H-E14-2006-00159-FR, EDF R&D, 2006. 132, 171

[Pasanisi 2007] A. Pasanisi. *Aide multicritère à la décision en rénovation énergétique des logements existants : l'outil REFLEX*. Tech. Report H-E14-2007-00425-FR, EDF R&D, 2007. 132, 171

[Pasanisi 2008a] A. Pasanisi. *Cahier des charges d'un module générique d'aide multicritère à la décision*. Tech. Report H-E14-2008-00073-FR, EDF R&D, 2008. 132, 171

[Pasanisi 2008b] A. Pasanisi. *Réflexions préalables au développement d'un logiciel d'évaluation probabiliste du débit de fuite d'un générateur de vapeur en conditions d'EHP*. Tech. Report H-T56-2008-02960-FR, EDF R&D, 2008. 60, 171

[Pasanisi 2010] A. Pasanisi. *Bayesian uncertainty assessment in industrial practice: EDF-R&D framework and on-going works*. In 5èmes Rencontres stat. au sommet de Rochebrune, Megève (France), March 2010. 60, 170

[Pasanisi 2011] A. Pasanisi. *An Industrial Viewpoint on Uncertainty Quantification in Simulation: Stakes, Methods, Tools, Examples*. In IFIP Working Conference on Uncertainty Quantification in Scientific Computing, Boulder (CO, USA), August 2011. 60, 168

[Pasanisi 2012a] A. Pasanisi. *Analyse d'incertitudes en simulation numérique : méthodes, outils, défis scientifiques*. In 30èmes rencontres AUGC et IBPSA-France, Chambéry (France), June 2012. 60, 168

[Pasanisi 2012b] A. Pasanisi. *Approche décisionnelle pour l'estimation de critères fiabilistes à l'aide de la simulation numérique : point de vue industriel et exemples d'application*. In Journées JSTAR 2012, Rennes (France), October 2012. 106, 168

[Pasanisi 2014] A. Pasanisi. *Fragilité et aléa sismique : apports de l'approche bayésienne*. In 7èmes Rencontres stat. au sommet de Rochebrune, Megève, France, April 2014 2014. 106, 169

[Pedroni *et al.* 2012] N. Pedroni, E. Zio, E. Ferrario and A. Pasanisi. *Propagation of aleatory and epistemic uncertainties in the model for the design of a flood protection dike*. In ESREL 2012 / PSAM11 Conference, Helsinki (Finland), June 2012. 90, 101, 102, 169

[Pedroni *et al.* 2013a] N. Pedroni, E. Zio, E. Ferrario, Pasanisi A. and M. Couplet. *Hierarchical propagation of probabilistic and non-probabilistic uncertainty in the parameters of a risk model*. Computers and Structures, vol. 126, pages 199–213, 2013. 90, 101, 102, 167

[Pedroni *et al.* 2013b] N. Pedroni, E. Zio, E. Ferrario, A. Pasanisi and M. Couplet. *Bayesian update of the parameters of probability distributions for risk assessment in a two-level hybrid probabilistic-possibilistic uncertainty framework*. In ESREL 2013 Conference, Amsterdam (Netherlands), September 2013. 90, 103, 169

[Pedroni *et al.* 2014] N. Pedroni, E. Zio, A. Pasanisi and M. Couplet. *A critical discussion and practical recommendations on some issues relevant to the non-probabilistic treatment of uncertainty in engineering risk assessment*. Sumitted, 2014. 90, 103, 168

[Pesti *et al.* 1996] G. Pesti, B. Shrestha, L. Duckstein and I. Bogardi. *A fuzzy rule-based approach to drought assessment*. Water Resources Research, vol. 32, pages 1741–1747, 1996. 90

[Picheny *et al.* 2010] V. Picheny, D. Ginsbourger, O. Roustant, R.T. Haftka and N.H. Kim. *Adaptive designs of experiments for accurate approximation of a target region*. Journal of Mechanical Design, vol. 132, no. 7, page 071008, 2010. 78, 150

[Pilkney & Pilkney-Jarvis 2007] O.H. Pilkney and L. Pilkney-Jarvis. Useless Arithmetic: Why Environmental Scientists Can't Predict the Future. Columbia University Press, 2007. 61

[Poggi *et al.* 2014] J.M. Poggi, R. Kenett and A. Pasanisi Eds. *Special issue: Graphical Causality Models: Trees, Bayesian Networks and Big Data*. Quality Technology and Quantitative Management, vol. 11, no. 1, pages 1–147, 2014. 168

[Pollard 1973] J.H. Pollard. Mathematical Models for the Growth of Human Populations. Cambridge University Press, 1973. 36

[Pólya 1930] G. Pólya. *Sur quelques points de la théorie des probabilités*. Annales de l'institut Henri Poincaré, vol. 2, no. 1, pages 117–167, 1930. 16

[Pomerol & Barba-Romero 2000] J.C. Pomerol and S. Barba-Romero. Multicriterion decision in management: Principles and practice. Kluwer Academic Publishers, 2000. 133

References

[Popelin & Iooss 2013] A.L. Popelin and B. Iooss. *Visualization tools for uncertainty and sensitivity analyses on thermal-hydraulic transients*. In Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2013 (SNA + MC 2013), Paris (France), October 2013. 151

[Posthuma *et al.* 2001] L. Posthuma, G.W. Suter and T.P. Traas. Species Sensitivity Distributions in Ecotoxicology. Taylor & Francis, 2001. 106

[Potter *et al.* 2012] K. Potter, P. Rosen and C.R. Johnson. *From Quantification to Visualization: A Taxonomy of Uncertainty Visualization Approaches*. In A.M. Dienstfrey and R.F. Boisvert (Eds.), Uncertainty Quantification in Scientific Computing, pages 226–249. Springer, 2012. 151

[Prieur *et al.* 2011] C. Prieur, A. Pasanisi and F. Wahl Eds. *Special issue: Stochastic methods for sensitivity analysis*. Journal de la Société Française de Statistique, vol. 152, no. 1, pages 1–130, 2011. 60, 168

[Pronzato & Müller 2012] L. Pronzato and W. Müller. *Design of computer experiments: space filling and beyond*. Statistics and Computing, vol. 22, pages 681–701, 2012. 78, 150

[Prud'Homme *et al.* 2011] C. Prud'Homme, G. Pena and V. Chabannes. *High order fluid structure interaction in 2D and 3D: Application to blood flow in arteries*. In ACOMEN 11 - 5th International Conference on Advanced COmputational Methods in ENgineering, pages 1–10, Liège (Belgium), October 2011. 63

[Prud'Homme *et al.* 2012] C. Prud'Homme, V. Chabannes, V. Doyeux, M. Ismail, A. Samake and G. Pena. *Feel++: A Computational Framework for Galerkin Methods and Advanced Numerical Methods*. ESAIM: Proceedings, vol. 38, pages 429 – 455, December 2012. 85

[Puolamäki & Kaski 2009] K. Puolamäki and S. Kaski. *Bayesian Solutions to the Label Switching Problem*. In N. Adams, C. Robardet, A. Siebes and J.F. Boulicaut (Eds.), Proceedings of the 8th International Symposium on Intelligent Data Analysis, pages 381–392, Lyon (France), August 2009. 32

[Rapenne *et al.* 2013] S. Rapenne, M. Baudin, E. Demay, K. Shakourzadeh, A. Pasanisi and O. Alos-Ramos. *Une boîte à outils Scilab pour prédire l'entartrage dans les circuits de refroidissement des centrales nucléaires équipées de tours aéroréfrigérantes et son couplage avec OpenTURNS*. In ScilabTEC, Palaiseau (France), June 2013. 60, 170

[Regan *et al.* 2004] H.M. Regan, S. Ferson and D. Berleant. *Equivalence of methods for uncertainty propagation of real-valued random variables*. International Journal of Approximate Reasoning, vol. 36, no. 1, pages 1–30, 2004. 97

[Ren *et al.* 2004] C. Ren, D. Sun and K. D. Dipak. *Comparison of Bayesian and Frequentist Estimation and Prediction for a Normal Population*. The Indian Journal of Statistics, vol. 66, no. 4, pages 678–706, 2004. 110

[Ren *et al.* 2010] H. Ren, W. Zhou, K. Nakagami, W. Gao and Q. Wu. *Multi-objective optimization for the operation of distributed energy systems considering economic and environmental aspects*. Applied Energy, vol. 87, no. 12, pages 3642–3651, 2010. 134

[Rey 2004] E. Rey. *Office building retrofitting strategies: multicriteria approach of an architectural and technical issue*. Energy and Buildings, vol. 36, pages 367–372, 2004. 140

[Ribes & Caremoli 2007] A. Ribes and C. Caremoli. *SALOME platform component model for numerical simulation*. In 31st Annual International Computer Software and Applications Conference, pages 553–564, Beijing (China), July 2007. 65

[Richet & Munoz-Zuniga 2013] Y. Richet and M. Munoz-Zuniga. *Practical computer experiments ... in R*. In The R User Conference 2013, Albacete (Spain), July 2013. 78

[Rinne 2008] H. Rinne. The Weibull Distribution: A Handbook. Taylor & Francis, 2008. 16, 20

[Robert & Casella 2010] C.P. Robert and G. Casella. Monte Carlo Statistical Methods. Springer, 2010. 34, 37

[Robert 2001] C.P. Robert. The Bayesian Choice: from Decision-Theoretic Motivations to Computational Implementation. Springer, 2001. 32, 110

[Roberts & Rosenthal 2007] G.O. Roberts and J.S. Rosenthal. *Coupling and Ergodicity of Adaptive MCMC*. Journal of Applied Probabilities, vol. 44, pages 458–475, 2007. 39, 41, 45

[Roberts & Rosenthal 2009] G.O. Roberts and J.S. Rosenthal. *Examples of Adaptive MCMC*. Journal of Computational and Graphical Statistics, vol. 18, pages 349–367, 2009. 39, 41

[Roberts & Sahu 1997] G.O. Roberts and S.K. Sahu. *Updating schemes, correlation structure, blocking, and parametrization for the Gibbs sampler*. Journal of the Royal Statistical Society B, vol. 59, pages 291–317, 1997. 38

[Rockafeller & Uryasev 2002] R.T. Rockafeller and S. Uryasev. *Conditional value-at-risk for general loss distributions*. Journal of Banking & Finance, vol. 26, pages 1443–1471, 2002. 67

[Roero *et al.* 2008] C. Roero, A. Pasanisi and G. Celeux. *Éclairages sur les modèles discrets de survie en fiabilité industrielle*. Tech. Report H-T56-2008-04499-FR, EDF R&D, 2008. 14, 171

[Roh & Xi 2000] Y.S. Roh and Y. Xi. *A general formulation for transition probabilities of Markov model and the application to fracture of composite materials*. Probabilistic Engineering Mechanics, vol. 15, pages 241–250, 2000. 47

[Rosenthal 2007] J.S. Rosenthal. *AMCMC: An R interface for adaptive MCMC*. Computational Statistics and Data Analysis, vol. 51, pages 5467–5470, 2007. 39

[Roulet *et al.* 2002] C.A. Roulet, F. Flourentzou, H.H. Labben, M. Santamouris, I. Koronaki, E. Dascalaki and V. Richalet. *ORME: A multicriteria rating methodology for buildings*. Building and Environment, vol. 37, pages 579–586, 2002. 140

[Roustant *et al.* 2010] O. Roustant, J. Franco, L. Carraro and A. Jourdan. *A Radial Scanning Statistic for Selecting Space-filling Designs in Computer Experiments*. In A. Giovagnoli, A.C. Atkinson, B. Torsney and C. May (Eds.), mODa 9 - Advances in Model-Oriented Design and Analysis, pages 189–196. Physica-Verlag HD, 2010. 78, 150

[Roustant *et al.* 2012] O. Roustant, D. Ginsbourger and Y. Deville. *DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization*. Journal of Statistical Software, vol. 51, no. 1, pages 1–55, 2012. 81

[Roustant 2011] O. Roustant. *Statistical models and methods for computer experiments*. HDR Qualification Thesis, Université Jean Monnet Saint Étienne, 2011. 69, 78

[Roy & Bouyssou 1993] B. Roy and D. Bouyssou. Aide multicritère à la décision : méthodes et cas. Économica, 1993. 133, 137

[Roy & Gupta 1992] D. Roy and R.P. Gupta. *Classifications of discrete lives*. Microelectronics and Reliability, vol. 10, no. 32, pages 1450–1459, 1992. 16

[Roy & Oberkampf 2011] C.J. Roy and W.L. Oberkampf. *A comprehensive framework for verification, validation and uncertainty quantification in scientific computing*. Computer Methods in Applied Mechanics and Engineering, vol. 200, pages 2131–2144, 2011. 97

[Roy 1968] B. Roy. *Classement et choix en présence de points de vue multiples (la méthode ELECTRE)*. La Revue d'Informatique et de Recherche Opérationelle, vol. 8, pages 57–75, 1968. 135

[Roy 1978] B. Roy. *ELECTRE III : un algorithme de classements fondé sur une représentation floue des préférences en présence de critères multiples*. Cahiers du Centre d'Études de Recherche Opérationnelle, vol. 20, no. 1, pages 13–24, 1978. 136

[Roy 1996] B. Roy. Multicriteria Methodology for Decision Aiding. Kluwer Academic Publishers, 1996. 132, 133

[Rubin 1976] D.B. Rubin. *Inference and missing data*. Biometrika, vol. 63, pages 581–592, 1976. 36

[Rutherford 1994] J.C. Rutherford. River Mixing. Wiley, 1994. 91, 92

[Sabbagh 1996] K. Sabbagh. 21st Century Jet: The Making and Marketing of the Boeing 777. Scribner, 1996. 63

[Sacks *et al.* 1989] J. Sacks, W.J. Welch, T.J. Mitchell and T.J. Wynn. *Design and analysis of computer experiments*. Statistical Science, vol. 4, pages 409–435, 1989. 70, 77

[Saito & Matsumoto 2006] M. Saito and M. Matsumoto. *SIMD-oriented Fast Mersenne Twister: a 128-bit Pseudorandom Number Generator*. In A. Keller, S. Heinrich and H. Niederreiter (Eds.), Monte Carlo and Quasi-Monte Carlo Methods, pages 607–622. Springer, 2006. 79

[Saltelli *et al.* 2004] A. Saltelli, S. Tarantola, F. Campolongo and M. Ratto. Sensitivity analysis in practice: A guide to assessing scientific models. Wiley, 2004. 67, 75

[Saltelli 2002] A. Saltelli. *Sensitivity analysis for importance assessment*. Risk Analysis, vol. 22, no. 3, pages 579–590, 2002. 61

References

[Salvia & Bollinger 1982] A.A. Salvia and R.C. Bollinger. *On discrete hazard functions*. IEEE Transactions on Reliability, vol. 31, no. 5, pages 458–459, 1982. 16

[Sargent *et al.* 2000] D.J. Sargent, J.S. Hodges and B.P. Carlin. *Structured Markov Chain Monte Carlo*. Journal of Computational and Graphical Statistics, vol. 9, pages 217–234, 2000. 38

[Schuëller *et al.* 1989] G.I. Schuëller, C.G. Bucher, U. Bourgund and W. Ouypornprasert. *On efficient computational schemes to calculate structural failure probabilities*. Probabilistic Engineering Mechanics, vol. 4, no. 1, pages 10–18, 1989. 73

[Schultz *et al.* 2010] T. Schultz, P. Gouldby, D. Simm and L. Wibowo. *Beyond the Factor of Safety: Developing Fragility Curves to Characterize System Reliability*. Tech. Report ERDC SR-10-1, US Army Corps of Engineers, 2010. 122

[Schweber & Wachter 2000] S. Schweber and M. Wachter. *Complex Systems, Modelling and Simulation*. Studies in History and Philosophy of Modern Physics, vol. 31, pages 583–609, 2000. 61

[Sendi *et al.* 1999] P.P. Sendi, H.C. Bucher, B.A. Craig, D. Pfluger and M. Battegay. *Estimating AIDS-free survival in a severely immunosuppressed asymptomatic HIV-infected population in the era of antiretroviral triple combination therapy. Swiss HIV Cohort Study*. Journal of acquired immune deficiency syndromes and human retrovirology, vol. 20, pages 376–381, 1999. 46

[Shafer 1976] G. Shafer. A Mathematical Theory of Evidence. Princeton University Press, 1976. 158

[Shaked *et al.* 1995] M. Shaked, J.G. Shanthikumar and J.B. Valdez-Torres. *Discrete hazard rate functions*. Computers & Operations Research, vol. 22, no. 4, pages 391 – 402, 1995. 16

[Shannon 1948] C Shannon. *A mathematical theory of communication*. The Bell System Technical Journal, vol. 27, pages 379–423 and 623–656, 1948. 74

[Sheu 1998] S.H. Sheu. *A generalized age and block replacement of a system subject to shocks*. European Journal of Operational Research, vol. 108, no. 2, pages 345 – 362, 1998. 20

[Shinozuka *et al.* 2000] M. Shinozuka, M. Feng, J. Lee and T. Naganuma. *Statistical Analysis of Fragility Curves*. Journal of Engineering Mechanics, vol. 126, no. 12, pages 1224–1231, 2000. 124

[Sobol 1993] I.M. Sobol. *Sensitivity estimates for non linear mathematical models*. Mathematical Modelling and Computational Experiments, vol. 1, pages 407–414, 1993. 70, 76

[Solomos *et al.* 2008] G. Solomos, A. Pinto and S. Dimova. *A Review of the seismic hazard zonation in national building codes in the context of Eurocode 8*. Tech. Report EUR 23563 EN-2008, JRC European Commission, 2008. 122

[Sonnenberg & Beck 1993] F.A. Sonnenberg and J.R. Beck. *Markov models in medical decision making: a pratical guide*. Medical Decision Making, vol. 13, pages 322–338, 1993. 31

[Stein & Dattero 1984] W.E. Stein and R. Dattero. *A new discrete Weibull distribution*. IEEE Transactions on Reliability, vol. 33, no. 2, pages 196–197, 1984. 19

[Stewart & Netherton 2008] G. Stewart and D. Netherton. *Security risks and probabilistic risk assessment of glazing subject to explosive blast loading*. Reliability Engineering and System Safety, vol. 93, no. 4, pages 627–638, 2008. 122

[Straub & Der Kiureghian 2008] D. Straub and A. Der Kiureghian. *Improved seismic fragility modeling from empirical data*. Structural Safety, vol. 30, no. 4, pages 320–336, 2008. 122, 124

[Strid *et al.* 2010] I. Strid, P. Giordani and R. Kohn. *Adaptive hybrid Metropolis-Hastings samplers for DSGE models*. Tech. Report Working Paper Series in Economics and Finance No. 724, Stockholm School of Economics, 2010. 41

[Sudret *et al.* 2011] B. Sudret, G. Blatman and M. Berveiller. *Surfaces de réponse par chaos polynomial*. In J. Baroth, F. Schoefs and D. Breysse (Eds.), Fiabilité des ouvrages, pages 195–217. Hermès, 2011. 60

[Sudret 2007] B. Sudret. *Uncertainty propagation and sensitivity analysis in mechanical models. Contributions to structural reliability and stochastic spectral methods*. HDR Qualification Thesis, University Blaise Pascal Clermont 2, 2007. 69, 77

[Sudret 2008] B. Sudret. *Global sensitivity analysis using polynomial chaos expansions*. Reliability Engineering & System Safety, vol. 93, no. 7, pages 964–979, 2008. 77

[Sundberg 2010] M. Sundberg. *Cultures of simulations vs. cultures of calculations? The development of simulation practices in meteorology and astrophysics*. Studies in History and Philosophy of Modern Physics, vol. 41, pages 273–281, 2010. 61

[Theisen & Glesner 1998] M. Theisen and M. Glesner. *The Stimulation of the Water Balance in a Landfill with the Help of an Equation Based and a Fuzzy System Based Method*. In EUFIT '98 - 6th European Congress on Intelligent Techniques and Soft Computing, pages 1656–1660, Aachen (Germany), September 1998. 91

[Tran *et al.* 2002] L.T. Tran, M.A. Ridgley, L. Duckstein and R. Sutherland. *Application of fuzzy logic-based modeling to improve the performance of the Revised Universal Soil Loss Equation*. CATENA, vol. 47, no. 3, pages 203–226, 2002. 90

[Tsai & Wang 2011] W.C. Tsai and C.H. Wang. *Economic optimization for an off-line inspection, disposition and rework model*. Computers & Industrial Engineering, vol. 61, no. 3, pages 891 – 896, 2011. 20

[Turkle 1995] S. Turkle. Life on the screen. Identity in the age of the Internet. Touchstone, 1995. 61

[Tuyl *et al.* 2009] P. Tuyl, R. Gerlach and K. Mengersen. *Posterior predictive arguments in favor of the Bayes-Laplace prior as the consensus prior for binomial and multinomial parameters*. Bayesian Analysis, vol. 4, pages 151–158, 2009. 33

[Urakabe *et al.* 1975] S. Urakabe, Y. Orita, Y. Fukuhara, A. Ando, N. Ueda, M. Inque, T. Furukawa and H. Abe. *Prognosis of chronic glomerulonephritis in adult patients estimated on the basis of the Markov process*. Clinical Nephrology, vol. 3, pages 48–53, 1975. 46

[Uygunoğlu & Keçebaş 2011] T. Uygunoğlu and A. Keçebaş. *LCC analysis for energy-saving in residential buildings with different types of construction masonry blocks*. Energy and Buildings, vol. 43, no. 9, pages 2077–2085, 2011. 135

[Vallaghé *et al.* 2011] S. Vallaghé, A. Le Hyaric, M. Fouquembergh and C. Prud'homme. *A successive constraint method with minimal offline constraints for lower bounds of parametric coercivity constant*. Tech. report, Laboratoire Jean Kuntzmann, 2011. 85

[Varian 1974] H. Varian. *A Bayesian Approach to Real Estate Assessment*. In S. Fienberg and A. Zellner (Eds.), Studies in Bayesian Econometrics and Statistics, pages 195–208. North Holland, 1974. 110

[Vezolle & Berthou 2008] P. Vezolle and J.Y. Berthou. *EDF R&D and IBM collaborations on Complex CFD applications using Blue Gene L and P*. In 14th Annual Meeting of ScicomP, the IBM HPC Systems Scientific Computing User Group, Poughkeepsie (NY, USA), May 2008. 65

[Vihola 2010] M. Vihola. *Grapham: Graphical models with adaptive random walk Metropolis algorithms*. Computational Statistics and Data Analysis, vol. 54, pages 49–54, 2010. 39

[Vincke 1992] P. Vincke. Multicriteria Decision-Aid. Wiley, 1992. 133

[Vogel 1987] R.M. Vogel. *Reliability indices for water supply systems*. Journal of Water Resources Planning and Management, vol. 113, pages 563–579, 1987. 31

[Vucinic *et al.* 2000] D. Vucinic, J. Favaro, B. Sünder, I. Jenkinson, G. Tanzini, B.K. Hazarika, M.R. d'Alcalà, D. Vicinanza, R. Greco and A. Pasanisi. *Fast and convenient access to fluid dynamics data via the World Wide Web*. In ECCOMAS 2000 Conference, Barcelona (Spain), September 2000. 8, 169

[Vucinic 2007] D. Vucinic. *Development of a Scientific Visualization System*. PhD thesis, Vrije Universiteit Brussel, 2007. 8

[Wald 1939] A. Wald. *Contributions to the Theory of Statistical Estimation and Testing Hypotheses*. The Annals of Mathematical Statistics, vol. 10, no. 4, pages 299–326, 1939. 109

[Wang & Sheu 2001] C.H. Wang and S.H. Sheu. *Simultaneous determination of the optimal production-inventory and product inspection policies for a deteriorating production system*. Computers & Operations Research, vol. 28, no. 11, pages 1093 – 1110, 2001. 20

[Wang & Sheu 2003] C.H. Wang and S.H. Sheu. *Optimal lot sizing for products sold under free-repair warranty*. European Journal of Operational Research, vol. 149, no. 1, pages 131 – 141, 2003. 20

References

[Wang & Simpson 2004] G.G. Wang and T.W. Simpson. *Fuzzy Clustering Based Hierarchical Metamodeling for Design Space Reduction and Optimization*. Engineering Optimization, vol. 36, no. 3, pages 313–335, 2004. 91

[Wang *et al.* 2004] W. Wang, R. Zmeureanu and H. Rivard. *Applying multi-objective genetic algorithms in green building design optimization*. Building and Environment, no. 40, pages 1512–1525, 2004. 140

[Wang *et al.* 2009] W.Y. Wang, S.H. Sheu, Y.C. Chen and D.J. Horng. *Economic optimization of off-line inspection with rework consideration*. European Journal of Operational Research, vol. 194, no. 3, pages 807 – 813, 2009. 20

[Weichert 1980] D.H. Weichert. *Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes*. Bulletin of the Seismological Society of America, vol. 70, pages 1337–1356, 1980. 128

[Wein & Wu 2001] L.M. Wein and J.T. Wu. *Estimation of replicative senescence via a population dynamics model of cells in culture*. Experimental Gerontology, vol. 36, no. 1, pages 79 – 88, 2001. 20

[Wiener 1938] N. Wiener. *The homogeneous chaos*. American Journal of Mathematics, vol. 60, no. 4, pages 897–936, 1938. 76

[Woldt *et al.* 1997] W. Woldt, C. Dou and I. Bogardi. *Simulation of Solute Transport in the Unsaturated Zone using Fuzzy Rule-Based Approach*. In EUFIT '97 - 5th European Congress on Intelligent Techniques and Soft Computing, pages 2131–2140, Aachen (Germany), September 1997. 91

[Wong & Wiart 2005] M.F. Wong and J. Wiart. *Modelling of electromagnetic wave interactions with the human body*. Comptes Rendus Physique, vol. 6, no. 6, pages 585–594, 2005. 63

[Wu & Carroll 1988] M. Wu and R. Carroll. *Estimation and comparison of changes in presence of informative right censoring by modelling the censoring process*. Biometrics, vol. 44, pages 175–188, 1988. 36

[Xie *et al.* 2002] M. Xie, O. Gaudoin and C. Bracquemond. *Redefining failure rate function for discrete distributions*. Journal of Reliability, Quality and Safety Engineering, vol. 9, no. 3, pages 275–285, 2002. 16, 21

[Ye *et al.* 1993] G.W. Ye, G. Yagawa and S. Yoshimura. *Probabilistic fracture mechanics analysis based on three-dimensional J-integral database*. Engineering Fracture Mechanics, vol. 44, no. 6, pages 887–893, 1993. 73

[Zeleny 1973] M. Zeleny. *Compromise Programming*. In J. Cochrane and M. Zeleny (Eds.), Multiple Criteria Decision Making, pages 262–301. University of South Carolina Press, 1973. 134

[Zhang & Dai 2007] L. Zhang and S. Dai. *Application of Markov model to environmental fate of phenanthrene in Lanzhou Reach of Yellow River*. Chemosphere, vol. 67, pages 1296–1299, 2007. 31

[Zhu *et al.* 2012] Y. Zhu, Y. Tao and R. Rayegan. *A comparison of deterministic and probabilistic life cycle cost analyses of ground source heat pump (GSHP) applications in hot and humid climate*. Energy and Buildings, vol. 55, pages 312–321, 2012. 135

[Zio 2007] E. Zio. An Introduction to the Basics of Reliability and Risk Analysis. World Scientific Publishing Company Inc., 2007. 15