CRANFIELD UNIVERSITY

XIN CHEN

ENABLERS FOR UNCERTAINTY QUANTIFICATION AND MANAGEMENT IN EARLY STAGE COMPUTATIONAL DESIGN. AN AIRCRAFT PERSPECTIVE

SCHOOL OF AEROSPACE, TRANSPORT AND MANUFACTURING

PhD Thesis
Academic Year: 2014-2017

Supervisor: Prof. Marin D. Guenov, Dr Arturo Molina-Cristóbal
October 2017

CRANFIELD UNIVERSITY

SCHOOL OF AEROSPACE, TRANSPORT AND MANUFACTURING

PhD Thesis

Academic Year 2014 - 2017

XIN CHEN

Enablers for Uncertainty Quantification and Management in Early Stage Computational Design. An Aircraft Perspective

Supervisor: Prof. Marin D. Guenov, Dr Arturo Molina-Cristóbal

October 2017

Abstract

Presented in this thesis are novel methods for uncertainty quantification and management (UQ&M) in computational engineering design. The research has been motivated by the industrial need for improved UQ&M techniques, particularly in response to the rapid development of the model-based approach and its application to the (early) design of complex products such as aircraft. Existing work has already addressed a number of theoretical and computational challenges, especially regarding uncertainty propagation. In this research, the contributions to knowledge are within the wider UQ&M area.

The first contribution is related to requirements for an improved margin management policy, extracted from the FP7 European project, TOICA (Thermal Overall Integrated Conception of Aircraft). Margins are traditional means to mitigate the effect of uncertainty. They are relatively better understood and less intrusive in current design practice, compared with statistical approaches. The challenge tackled in this research has been to integrate uncertainty analysis with deterministic margin allocations, and to provide a method for exploration and trade-off studies. The proposed method incorporates sensitivity analysis, uncertainty propagation, and the set-based design paradigm. The resulting framework enables the designer to conduct systematic and interactive trade-offs between margins, performances and risks. Design case studies have been used to demonstrate the proposed method, which was partially evaluated in the TOICA project.

The second contribution addresses the industrial need to properly ‘allocate’ uncertainty during the design process. The problem is to estimate how much uncertainty could be tolerated from different sources, given the acceptable level of uncertainty associated with the system outputs. Accordingly, a method for inverse uncertainty propagation has been developed. It is enabled by a fast forward propagation technique and a workflow reversal capability. This part of the research also forms a contribution to the TOICA project, where the proposed method was applied on several test-cases. Its usefulness was evaluated and confirmed through the project review process.

The third contribution relates to the reduction of UQ&M computational cost, which has always been a burden in practice. To address this problem, an efficient sensitivity analysis method is proposed. It is based on the reformulation and approximation of Sobol's indices with a quadrature technique. The objective is to reduce the number of model evaluations. The usefulness of the proposed method has been demonstrated by means of analytical and practical test-cases. Despite some limitations for several specific highly non-linear cases, the tests confirmed significant improvement in computational efficiency for high dimensional problems, compared with traditional methods.

In conclusion, this research has led to novel UQ&M tools and techniques, for improved decision making in computational engineering design. The usefulness of these methods with regard to efficiency and interactivity has been demonstrated through relevant test-cases and qualitative evaluation by (industrial) experts.

Finally, it is argued that future work in this field should involve research and development of a comprehensive framework, which is able to accommodate uncertainty, not only with regard to computation, but also from the perspective of (expert) knowledge and assumptions.

Keywords:

Computational Model, Probabilistic Approach, Sensitivity Analysis, Margin Allocation, Inverse Propagation

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors: Prof. Marin D. Guenov and Dr Arturo Molina-Cristóbal, for providing me the opportunity to conduct this research, and for their patient support, guidance, and encouragement, throughout my four years at Cranfield University.

I would also like to thank my colleagues and friends in the Advanced Engineering Design Group: Dr Atif Riaz, Dr Varun Datta, Dr Vitaly Voloshin, Dr Mattia Padulo, Dr Gabriele Luigi Mura, Albert S.J. van Heerden, Yogesh Bile, and Sergio Jimeno Altelarrea, for their helpful suggestions and support.

Thanks to Dr Nan Yu, Dr Duo Zhang, Dr Jiajun Zhang, Dr Zihang Zhu, Zhiyu Xia, Jiawen Xia, and Ming Yu for many reasons.

Finally I would I like to thank my beloved parents, for their love and support throughout my life.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2D	Two Dimensional
3D	Three Dimensional
ANOVA	Analysis of Variance
APU	Auxiliary Power Unit
Avion	Avionics
BPR	Bypass Ratio
Brk	Brakes
CDF	Cumulative Density Function
CFD	Computational Fluid Dynamics
DB	Design Band
DoE	Design of Experiments
DP	Design Point
DS	Design Space
ECS	environmental control system
EPS	Electrical Power System
Equip	Equipment
ESAR	Equivalent Still Air Range
FASM	Fuel per Available Seat Mile
FAST	Fourier Amplitude Sensitivity Test
FCS	Flight control system
FEM	Finite Element Method
FLOPS	Flight Optimization System
FONoise	Flyover Noise
FS	Fuel System
GA	Genetic Algorithm
Gal	Galley
GIT	Generalized Information Theory
GN	Gaussian-Newton (Method)
GOSA	Goal Oriented Sensitivity Analysis
GSA	Global Sensitivity Analysis
GW	Gross Weight
GW	Gross Weight
HPS	Hydraulic Power System
ICP	Iso-Contours Plot
IfEnt	Inflight Entertainment System
IM	Incidence Matrix
Instr	Instruments
IPS	Ice Protection System
KLE	Karhunen–Loève Expansion
Lav	Lavatories

LFL	Landing Field Length
LG	Landing Gears
LoD	Lift over Drag
LSA	Local Sensitivity Analysis
MCS	Monte Carlo Simulation
MP	Margin Combination Point
MS	Margin Space
NASA	National Aeronautics and Space Administration
NOx	Nitrogen Oxide Emissions
PCE	Polynomial Chaos Expansion
PDE	Partial Differential Equation
PCP	Parallel Coordinates Plot
PDF	Probability Density Function
PaS	Passenger Service
PS	Performance Space
PSO	Particle Swarm Optimization
QoI	Quantities of Interest
RBDO	Reliability-Based Design Optimization
RDO	Robust Design Optimization
R&D	Research and Development
SA	Sensitivity Analysis
SIMPCODE	Simple Code for Aircraft Sizing
SLNoise	Sideline Noise
SLST	Sea Level Static Thrust
SPOT	Total Shaft Power Off-Take
std	Standard Deviation
ThrRev	Thrust Reverser System
TOFL	Take-Off Field Length
TOICA	Thermal Overall Integrated Conception of Aircraft
UQ&M	Uncertainty Quantification and Management
URQ	Univariate Reduced Quadrature
Vapp	Approach Velocity

Nomenclature

$(\cdot)^*$	target value of $(\cdot)$
$\boldsymbol{\mathcal{A}}$	matrix of coefficients in linear equations
a_{ij}	coefficients in linear equations
$\boldsymbol{b}$	column of constants in linear equations
$\mathbb{C}$	cost budget in uncertainty budgeting
$C_i(\boldsymbol{dv}, \boldsymbol{p})$	i_{th} constraint function
DS	design space
$\boldsymbol{dv}$	vector of design variables
dv_i	i_{th} design variable
dv_i^{LB}	lower bound of the i_{th} design variable
dv_i^{UB}	upper bound of the i_{th} design variable
DV_i	specified (valid) interval of the i_{th} design variable
$E(y)$	expectation of a general output
$E_{x_{\sim i}}\left(y\vert_{x_i=X_i}\right)$	conditional expectation of a general output, when the i_{th} input is fixed
$E_{URQ}(y)$	expectation of a general output using URQ
$\mathbb{f}$	probability density function
$\mathbb{F}$	cumulative density function
$f(\boldsymbol{x})$	computational workflow
$\mathbb{F}^{-1}$	Inverse cumulative density function of variable
$f_+(\boldsymbol{x_{+i}}), f_{++}(\boldsymbol{x_{++ij}}),$ $f_-(\boldsymbol{x_{-i}})$	intermediate function defined in calculation of Sobol' indices
$f_1^{(\sim i)}, f_2^{(\sim i,j)}, f_3^{(i)},$ $g_1, g_2, g_3, g_4, g_5, g_6$	intermediate function defined in the proposed sensitivity analysis method
F_C	cost function in uncertainty budgeting
$F_l(y_i)$	loss function in Taguchi method corresponding to y_i
F_{obj}	objective function
$F_{Outer}(\boldsymbol{\mathcal{M}_x})$	outer workflow
h_i^+, h_i^-	coefficients in URQ method
i, j, r	counting indices for loops
K	level of the sparse grid
$k_{P_{C_i}}$	factor for estimating probability of constraint satisfaction
$k_{T_{y_i}}$	factor in Taguchi method corresponding to y_i
l	number of output variables
M	order of interference in the FAST method
Mar_z	margin on variable z
Mar_z^{LB}	lower bound of margin on variable z
Mar_z^{UB}	upper bound of margin on variable z
$\mathcal{MG}(\mu_1, \mu_2, \sigma_1, \sigma_2, \lambda_1, \lambda_2)$	mixed-Gaussian distribution defined by: first mean, second mean,

	first standard deviation, second standard deviation, first proportion, and second proportion
MS	margin space
N	number of samples used in Taguchi method or a (quasi) Monte Carlo Simulation
n	number of input variables
$\mathcal{N}(\mu, \sigma)$	normal distribution defined by: mean and standard deviation
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	multivariate normal distribution: defined by means and covariance matrix
n_q	number of collocation point
n_s	number of locations of interest
N_{total}	total number of samples in the proposed sensitivity analysis method
$\boldsymbol{p}$	vector of parameters
$\mathbb{P}$	required probability of satisfying constraint
P	order of the polynomials
p_i	i_{th} parameter variable (part of a model or pre-defined when the design problem is formulated)
$\mathbb{P}\{\cdot\}$	probability of event $\{\cdot\}$
PS	performance space
$\boldsymbol{pv}$	vector of performance variables
pv_i	i_{th} performance variable
PV_i	calculated interval of the i_{th} performance variable
pv_i^{MAX}	maximum value of the i_{th} performance variable
pv_i^{MIN}	minimum value of the i_{th} performance variable
Q	order of quadrature
$q(\boldsymbol{s}, t, \boldsymbol{\xi_q})$	stochastic heat flux in space and time
q_i	deterministic heat flux corresponding to the i_{th} collocation point
rv_{y_i}	random calibration variable corresponding to the y_i
$\boldsymbol{s}$	vector indicating the location
$\boldsymbol{S_1}, \boldsymbol{S_2}$	sampling sets in Sobol' Pick and Freeze method
$\boldsymbol{s_i}$	vector indicating the i_{th} location
S_i	first order Sobol' index corresponding to the i_{th} input variable (regarding a single output)
S_i^T	total order Sobol' index corresponding to the i_{th} input variable (for a single output)
S_{ij}	second order Sobol' index corresponding to the combination (interaction) of the i_{th} and j_{th} input variables (for a single output)
t	time
$\mathcal{T}$	temperature
t	number of constraints
tol	tolerance
$\mathcal{T}ri(LB, UB, MV)$	triangular distribution defined by: lower bound, upper bound and middle vertex

u	number of design variables
U	intermediate variable in calculation of Sobol' indices
$\mathcal{U}(LB, UB)$	uniform distribution defined by: lower bound, upper bound
v	number of parameters
$V(y)$	variance of a general output
$V_{x_{\sim i}}\left(y\vert_{x_i=X_i}\right)$	conditional variance of a general output, when the i_{th} input is fixed
$V_{x_{\sim i,j}}\left(y\vert_{x_i=X_i,x_j=X_j}\right)$	conditional variance of a general output, when the i_{th} and j_{th} inputs are fixed
$V_{x_i}\left(y\vert_{\boldsymbol{x}_{\sim \boldsymbol{i}}=\boldsymbol{X}_{\sim \boldsymbol{i}}}\right)$	conditional variance of a general output, when all except the i_{th} input is fixed
$V_{URQ}(y)$	variance of a general output using URQ
w	weight factor in an objective function
w	number of margins
$W_0, W_i^+, W_i^-, W_i^\pm$	weight factors in URQ method
$W_0^{(i)}, W_0^{(i,j)}, W_0^{(\sim i)}, W_0^{+i}, W_0^{++ij}, W_0^{-+i}$	modified URQ weight factors in the proposed sensitivity analysis method
$\boldsymbol{x}$	vector of inputs
$\boldsymbol{x_{+i}}, \boldsymbol{x_{++ij}}, \boldsymbol{x_{-i}}$	vectors defined in calculation of Sobol' indices
x_i	i_{th} input (independent) variable
X_i	value of the i_{th} input variable
$x_i^{(r)}$	r_{th} sample of x_i from sampling set $\boldsymbol{S_1}$
$x_i^{(r)}$,	r_{th} sample of x_i from sampling set $\boldsymbol{S_2}$
$\boldsymbol{x_i^+}, \boldsymbol{x_i^-}$	vectors of inputs in URQ method
$\Delta\boldsymbol{x}$	vector of $\Delta x_i, i = 1,2,3, \dots, n$
Δx_i	variation of x_i
Δx_i^{LB}	lower bound of Δx_i
$\boldsymbol{y}$	vector of outputs
y_i	i_{th} output (dependent) variable
$y_i^{(j)}$	j_{th} sample of y_i
y_{i_R}	randomized i_{th} output (dependent) variable
z	general variable (input or output)
$\langle z^i \rangle$	i_{th} statistical moment
z_C	conservative estimation of variable z
z_N	nominal value of variable z
z_P	percentile of variable z
α	highest order of input statistical moment under study
β	highest order of output statistical moment under study
γ_z	skewness of variable z
Γ_z	kurtosis of variable z
η_{y_i}	signal to noise ratio in Taguchi method corresponding to y_i

θ_1, θ_2	coefficients in the Ishigami Function
$\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i)}, \boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i,j)}, \boldsymbol{x}_{\boldsymbol{r}}{}^{+}{}^{(\sim i)}, \boldsymbol{x}_{\boldsymbol{r}}{}^{+}{}^{(\sim i,j)}, \boldsymbol{x}_{\boldsymbol{r}}{}^{-}{}^{(\sim i)}, \boldsymbol{x}_{\boldsymbol{r}}{}^{-}{}^{(\sim i,j)}$	modified URQ vectors in the proposed sensitivity analysis method
μ_z	mean of variable z
$\boldsymbol{\xi}_{\boldsymbol{q}}$	vector of random variables in KLE
$\boldsymbol{\xi}_{\boldsymbol{q}}^{\boldsymbol{i}}$	i_{th} realization of $\boldsymbol{\xi}_{\boldsymbol{q}}$
ρ_1, ρ_2, ρ_3	coefficients in the Sobol' Function
$\sigma_{y_j}^{UB}$	upper bound of σ_{y_j}in uncertainty budgeting
σ_z	standard deviation of variable z
$\Phi(z)$	cumulative density function of standard normal distribution
$\Phi^{-1}(z)$	inverse cumulative density function of standard normal distribution
$\psi_i(\boldsymbol{\xi}_{\boldsymbol{q}})$	polynomial interplant of $\boldsymbol{\xi}_{\boldsymbol{q}}$
ω_{max}	maximum frequency in the FAST method

Variables in the test cases are not included.

1 Introduction

The embrace of computation in modern engineering design has largely reshaped the paradigm of synthesis and analysis. Back to the era of pioneering designers, such as the Wright brothers, it was extremely difficult to predict the performance of a design solution, due to limited knowledge and experience at that time. In fact, the Wright brothers had to build their own wind tunnel, so that proper wing geometry could be selected, based on experiment results (Keane and Nair, 2005). As knowledge was gained in relevant disciplines (such as structures and aerodynamics), several analytical and systematic aircraft design approaches were developed (Cherry and Croshere, 1948; Driggs, 1950), based on explicit mathematical relationships between major design variables and performance. Following this pattern, research was conducted to setup design problems as numerical optimizations (Cramer, 1995; Haupt, 1977; Hicks and Henne, 1977). As a result, the distinction between synthesis and analysis has been blurred. This is because the corresponding design knowledge (for synthesis and analysis) is formulated as computational models, while the latter can be combined and organised in one comprehensive *Computational Workflow* (Balachandran and Guenov, 2010). In such a model-based design approach, products can be virtually 'created' and 'assessed'. There is no doubt that physical prototypes are still necessary for design decision making. The benefit of using computational models is reflected on the reduction of development time and cost. Furthermore, it allows a designer to explore a design space and search for potentially promising solutions (Balachandran and Guenov, 2010; Guenov, Nunez and Gondhalekar, 2011; Nunez et al., 2009, 2012; Riaz, Guenov and Molina-Cristóbal, 2017), before any physical prototypes are produced.

Resulting from the extensive application of computational design, is a growing need for capabilities to handle uncertainty. Uncertainty is normally inevitable, due to limited knowledge and uncontrollable factors. Thus the computational results may deviate from the reality. Consequentially, a theoretical optimal solution may turn out to be sub-optimal or even infeasible at all. This effect can lead to excessive rework during a design process, especially for complex and highly integrated systems. For instance, if the stress in one structural component is underestimated, considerable amount of modification might be required, not only for this specific component, but also for the

entire aircraft as well. In the worst cases, catastrophic failures could happen in service operation, if this type of deviation remains undetected.

1.1 Context

Design under the presence of uncertainty is not a brand new topic. Traditionally, concerns of uncertainty are implicitly embedded in decision making. For instance, one approach can be to design for the worst-case scenarios, so that unknown or uncontrollable factors can be accommodated. In practice, such a manner can be problematic. Firstly, the worst-case scenarios are difficult to identify in reality (Zang et al., 2002), and are normally based on assumptions (Miles Jr., 1993). These assumptions themselves can be uncertain as well. Secondly, the worst-case scenarios will lead to conservative design solutions, which are likely to be uncompetitive regarding the overall performance. These limitations have led to the development of explicit *Uncertainty Quantification and Management* (UQ&M) as an independent discipline. On the other hand, UQ&M provides a systematic approach to estimating the amount of uncertainty and its consequence, so that better decisions can be made with the uncertain results of design computation.

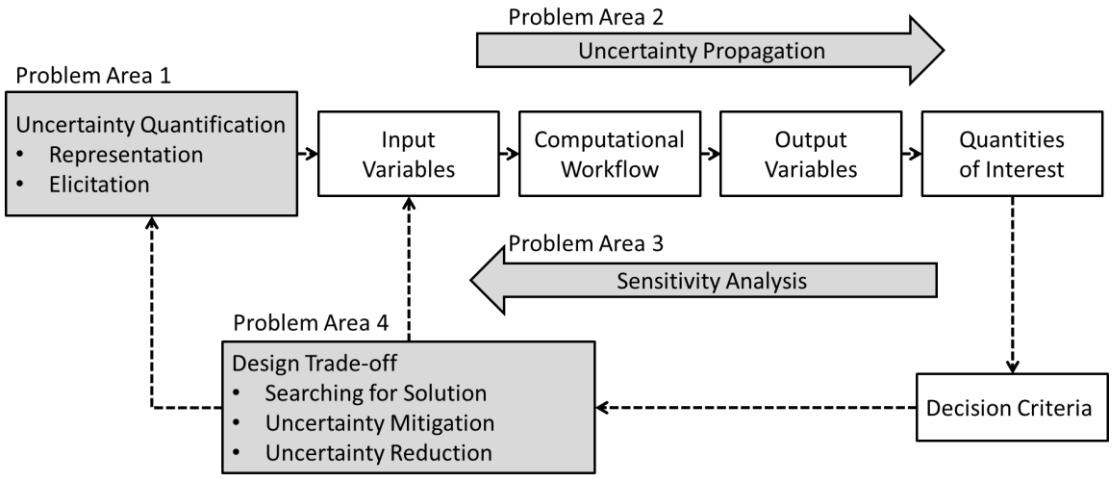

Figure 1-1. A common framework for UQ&M (adapted from (Mangeant, 2011; Rocquigny, Devictor and Tarantola, 2008))

Illustrated in Figure 1-1 is a common framework for UQ&M within the computational context (Mangeant, 2011; Rocquigny, Devictor and Tarantola, 2008). The white blocks are key elements of problem formulation, which include input and output variables, a

computational workflow, *Quantities of Interest* (QoI), and decision criteria. A QoI is an indicator of the uncertainty associated an output variable (for instance a statistical moment or a probability), while a decision criterion is a threshold of accepting/rejecting a specific design solution (Rocquigny, Devictor and Tarantola, 2008). More details of the problem formulation are presented in Section 2.2 and 2.5 as a background of this research.

The grey blocks in Figure 1-1 are the major UQ&M problem areas, as listed below:

Uncertainty Quantification: Uncertainty is an abstract concept and is different from general physical quantities (such as mass, height, etc.). To start with, the sources of uncertainty should be represented with proper mathematical forms. Currently, uncertainty quantification can be conducted following either the probabilistic approach or non-probabilistic approach. In the former, the variations of uncertain variables are represented as probability distributions, and described as *Probability Density Functions* (PDF). The non-probabilistic approaches include: *Possibility Theory* (Zadeh, 1999), *Evidence Theory* (Dempster, 1967; Shafer, 1976; Zadeh, 1986) etc. Within the scope of this research, the probabilistic approach is adopted for uncertainty quantification.

Apart from choosing a mathematical form, another aspect of uncertainty quantification is to decide the relevant details, for instance, the type and parameters of a probabilistic distribution. If historical data are available, these details can be obtained with a statistical regression. Otherwise, expert elicitation can be applied to construct the probabilistic distributions, based on personal knowledge and experience (Ayyub, 2001; Cooke, 1991; O'Hagan et al., 2006). This part of the problem is beyond the scope of this research, in which we assume that probabilistic distributions of uncertain sources are available for further analysis.

Uncertainty Propagation: Given the sources of uncertainty, the next step is to model the "non-deterministic behaviour" (Padulo, 2009) of the output variables, and to calculate the quantities of interest. This process is referred as uncertainty propagation.

Extensive research has been conducted within this problem area, with main focuses on the development and implementation of various numerical techniques. Among the available techniques, *Monte Carlo Simulation* (MCS) has been widely used for its

simplicity and effectiveness. It utilizes repeated sampling of the uncertain input variables, and approximates the output distributions (and QoI'S) based on statistical histograms. The major shortcoming of MCS is the computational cost, which can be prohibitive for high-fidelity models. Alternative techniques are based on *Gaussian Quadratures* (Evans, 1972; Seo and Kwak, 2002), *Polynomial Chaos* (Eldred and Burkardt, 2009; Xiu et al., 2002; Xiu and Karniadakis, 2002), *Taylor Series Expansion* (Du and Chen, 2002). The reader is referred to (Lee and Chen, 2009; Mares et al., 1853; Padulo, Campobasso and Guenov, 2007) for more comprehensive reviews and comparisons.

Sensitivity Analysis: The terminology of *Sensitivity Analysis* may refer to various concepts within different research communities. A general definition from (Smith, 2013) is "quantifying the relative contributions due to individual parameters or inputs and determining how variations in parameters affect measured responses". In the context of UQ&M, sensitivity analysis is utilized to rank the contributions from different sources of uncertainty, regarding their impact over the uncertainty of output variables (Saltelli et al., 2004).

For large scale design problems, it is normally impractical to consider all the uncertain factors, especially at the outset. Sensitivity analysis can be applied to identify the most influential factors, so that the problem dimensions can be reduced by freezing insignificant factors. In the reduced problem, priorities can be set for further analysis (e.g., to allocate more time and resource on the most influential factors).

Trade-off under uncertainty: In the original figure from (Rocquigny, Devictor and Tarantola, 2008), this part is referred to as "Feedback", which includes actions for improvement (e.g. modifying the input variables, using more accurate models, collecting more experiment data, etc. (Rocquigny, Devictor and Tarantola, 2008)). In the context of engineering design, we narrow down this problem area as conducting design trade-offs. That is to make decisions with information obtained from quantification, propagation, and sensitivity analysis.

Several aspects can be considered within this problem area. The first one is to search for promising design solutions which are less affected by uncertain variables. The second aspect is mitigation of uncertainty (McManus and Hastings, 2005), which include

margin, safety factor, tolerance, redundancy, etc. The last aspect is to reduce uncertainty from the sources, for instance, to calibrate a computational model with experiment results (Kennedy and O'Hagan, 2001).

1.2 Motivations

For this research, trade-off under uncertainty and sensitivity analysis are the particular problem areas of interest. Specifically, the FP7 European project, *Thermal Overall Integrated Conception of Aircraft*[1] (TOICA) (Rouvreau, S., Mangeant, F., and Arbez, 2015), has specified some of the research challenges, which motivated this work.

The first motivation is related to requirements extracted from the TOICA project, for an improved margin management policy (Rouvreau, S., Mangeant, F., and Arbez, 2015). Compared with statistical design approaches, margins are relatively better understood and easier to implement. This enables the designer to focus more on design decision making in his/her own expertise (e.g. aerodynamics, thermodynamics, structure, etc.), rather than be heavily involved in statistical analysis. Furthermore, margins have already been applied extensively in industrial practice, which makes them less intrusive compared with statistical design approaches. Last but not least, margins still play import rules to handle "unknown-unknowns" (McManus and Hastings, 2005), especially at the early stage of design process. The shortcomings are also clear. Firstly, a margin lumps all the uncertainty into one piece, which goes against the need to treat uncertainty explicitly. Secondly, the allocation of margins is largely restricted by historic data and designer's experience. For novel technologies, it is more difficult to make decisions upon margins. Last but not least is the lack of systematic approach to assess the influence of margins on other margins, and on system performances.

The second motivation is the industrial need to manage sources of uncertainty during the design process. Uncertainty is inevitable in design computation. Nevertheless, the amount of uncertainty can be reduced by conducting more analysis, for instance, by using physical experiments and/or models of higher fidelity. Obviously, those activities will lead to additional time and/or computational resources, therefore a trade-off should be sought, between the cost and the accepted amount of uncertainty. This trade-off is referred as the problem of *Uncertainty Allocation*. Specifically, a sub-problem is to estimate how much uncertainty could be tolerated from different sources, given the acceptable amount of uncertainty associated with the system outputs. This sub-problem

[1] Project webpage: http://www.toica-fp7.eu/

is referred to as *Inverse Uncertainty Propagation* in this thesis. Compared with forward propagation, relatively less has been researched on the inversed process, which requires novel enablers to provide the relevant capability.

The third motivation is the requirement for an efficient method of sensitivity analysis. This is not directly proposed by the TOICA project, but from the general need to reduce UQ&M computational cost. Compared with the deterministic design approach, implementation of UQ&M requires additional calculations to provide sensible information about the uncertainty. Therefore the resulting computational cost has always been a burden in industrial practice, especially for large scale problems such as design of aerospace systems (Zang et al., 2002). Within this context, the challenge is to improve the current numerical techniques.

1.3 Aim and Objectives

The aim of this PhD research is to develop novel enablers, to improve UQ&M practice in the context of model-based design of complex systems such as aircraft.

The aim is supported by three objectives, corresponding to the aforementioned motivations:

- To develop a method for margin allocation and trade-off, which incorporates the UQ&M techniques, and facilitates systematic and assessment on the interaction between margins, performances, and probabilities of constraint satisfaction.
- To develop a method for inverse uncertainty propagation, as an enabler of the uncertainty allocation problem.
- To develop a method for efficient sensitivity analysis, which has lower computational cost compared with existing techniques, and will ultimately result in a more efficient UQ&M process.

1.4 Research Methodology

In a typical research on natural or social science, the research methodology may include proposing hypnosis, conducting experiment/simulation, collecting data, analysis, etc. On the other hand, the research presented in this thesis can be categorised under the so-

called “Science of Design” (Cross, 1993), where the methods themselves are under study, with the aim is to make them more efficient and effective.

Given the aforementioned motivations and objectives, the research questions are:

- What capabilities should be provided by the methods under development?
- How to realise these capabilities?

The answer to the first question is based on literature review, which identifies the gaps between the existing methods and the industrial need (from the TOICA project). Answering the second question is the actual development of these methods. As the question is ill-defined (Simon, 1973), the answer is not unique, and the path to this answer is more of a creation (through try and error) rather than fixed procedures. Specifically in this research, it involves integration and modification of existing approaches and tools to achieve the targeting capabilities.

The evaluation of these methods involves two levels. The lower level is regarding the correctness of the proposed mathematical equations and algorithms, which can be analytically verified or validated through numerical test-cases. This type of evaluation can be found for the second and third objectives. The higher level is regarding the usefulness of the proposed approaches and concepts, especially in the case of objective one. In the corresponding evaluation, the method was applied to a realistic test-case and demonstrated at one of the TOICA project milestone meetings. The feedback from the practicing designers and airframe systems architects was used as a qualitative assessment of the proposed method.

1.5 Thesis Overview

The remaining part of the thesis is structured as follows. Chapter 2 provides the background information of this PhD research. Specifically, the model-based design approach is explained briefly as a general context, along with essential definitions and existing enablers to support the methodology chapters. In Chapter 3, a literature review is presented regarding the specific problem areas of interest (motivations). Chapters 4, 5, and 6 present the novel methods developed in this research. They are, respectively, a method for margin allocation and trade-offs, a method for inverse uncertainty propagation, and a method for efficient sensitivity analysis. In each of these three

chapters, the specific problem is defined and formulated. Also, some additional literature review for every specific development is presented separately in the corresponding chapters. The proposed methods are evaluated with different test-cases, for which the results and discussions could be found at the end of these chapters. Chapter 7 is the overall conclusion of the PhD research, in which the research is summarised. Contribution to knowledge is stated, along with limitations of the current development. At the end of this chapter, a roadmap is represented for future work, regarding the broader picture of UQ&M research.

2 Background

2.1 Introduction

This chapter presents concepts and terminologies which are adopted in this PhD research. In Section 2.2, mathematical definitions are introduced for a general computational workflow and the associated variables. Within this context, the sources of uncertainty are briefly discussed in Section 2.3. The probabilistic approach to quantifying these uncertainties is briefly explained in Section 2.4. Section 2.5 presents two general formulations for uncertainty-based design problems, which include the Robust Design Optimization (Section 2.5.1) and Reliability Base Design Optimization (Section 2.5.2). In Section 2.6, an in-house software named AirCADia is introduced. This software has provided a computational environment for implementations of the proposed methods in Chapter 4, 5, and 6.

2.2 Computational Workflow

In this thesis, a model refers to a mathematical relationship between some input and output variables. At early design stages, e.g. in an aircraft sizing problem, models can be as simple as single-line algebraic equations, look-up tables, or design graphs. Later on, more complex models may take the form of numerical simulations, such as *Finite Element Method* (FEM) models and *Computational Fluid Dynamic* (CFD) models.

In design computation, an output of one model can be used as an input to another. In this manner, a computational workflow can be defined as an assembly of models connected by variables (Balachandran and Guenov, 2010; Nunez et al., 2012). Such a computational workflow is noted in the form of a function:

$$\boldsymbol{y} = f(\boldsymbol{x}), \tag{2-1}$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ refer to the vectors of n inputs and l outputs, respectively.

$$\boldsymbol{x} = [x_1, x_2, x_3 \dots, x_n] \in R^n \tag{2-2}$$

$$\boldsymbol{y} = [y_1, y_2, y_3 \dots, y_l] \in R^l \tag{2-3}$$

In this notation, $f(\boldsymbol{x})$ serves only for discussion purposes, since its analytical expression might be implicit, due to the models within the workflow (i.e. CFD, FEM or black-boxes, for which analytical equations are not available). The input/output variables are assumed to be scalars. In each vector, the elements are general and can be arranged in any orders, unless there is a need to group them due to the specific physical meanings (as will be demonstrated later in equation (2-4)).

Finally it should be stated that in this research, the input variables are assumed to be independent (If a correlation exists, it will take the form either as a model or as a constraint).

An illustrative example is shown in Figure 2-1 (Guenov et al., 2017), which represents a simplified computational workflow for aircraft sizing (named as SIMPCODE). This workflow is constructed with five models from aircraft design text books (Jenkinson, Simpkin and Rhodes, 1999; Raymer, 2012). The input and output variables are indicated as green and red ovals, respectively; while the models are represented as blocks. Full details of these variables are listed Table 2-1.

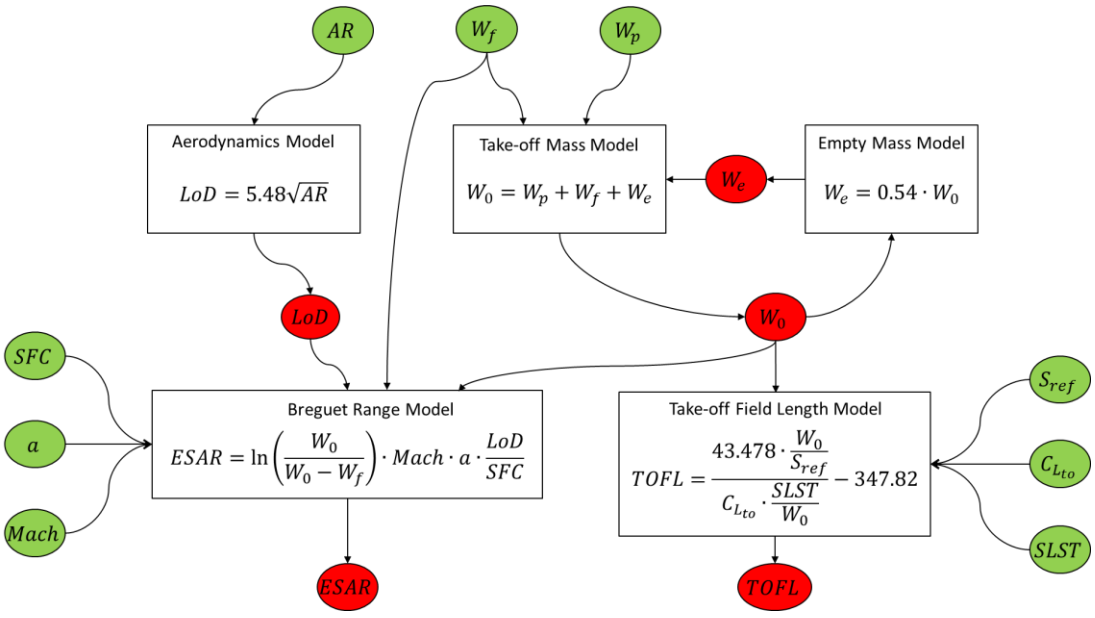

Figure 2-1. an illustrative example of an aircraft sizing workflow

Table 2-1. List of input and output variables

	Variable	Notation	Unit
Inputs	Aspect Ratio	AR	
	Fuel Weight	W_f	lb
	Payload Weight	W_p	lb
	Specific Fuel Consumption	SFC	$lb/hr/lb$
	Speed of Sound	a	$knot$
	Mach Number	$Mach$	
	Wing Reference Area	S_{ref}	ft^2
	Take-off Lift Coefficient	$C_{L_{to}}$	
	Sea Level Static Thrust	$SLST$	lb
Outputs	Empty Weight	W_e	lb
	Lift over Drag	LoD	
	Total Weight	W_0	lb
	Equivalent Still Air Range	$ESAR$	$n.m$
	Take-off Field Length	$TOFL$	ft

In some cases, the input variables can be further categorised as design variables and parameters. The former are used to define a specific solution, and their values are controlled by the designer. On the other hand, parameters are used to define factors which are beyond the design decision making, such as pre-specified operational conditions, material properties, and so on. Some of the parameters can also be embedded in computational models, for instance, the empirical coefficient "5.48" in the "Aerodynamics Model" (Figure 2-1). The values of parameters are normally fixed during the design process, but can be modified if the problem setup is changed (e.g. change of design cruise altitude due to modification of mission profiles).

It should be noted that, the distinction between design variables and parameters is not absolute, but dependent on the specific problem formulations. In this thesis, if the focus is on the mathematical aspect of a function (e.g. in Section 6), the notation of $\boldsymbol{x}$ will be used, which refers to general input variables. In this case, the ordering of variables in the vector $\boldsymbol{x}$ is arbitrary. If the emphasis is placed on a design problem, the design variables and parameters will be represented by $\boldsymbol{dv}$ and $\boldsymbol{p}$ respectively. Without loss of generality, the design variables are arranged at the front, followed by the parameters:

$$\boldsymbol{x} = (\boldsymbol{dv}, \boldsymbol{p}) \tag{2-4}$$

$$\boldsymbol{dv} = [dv_1, dv_2, dv_3, \dots, dv_u] \in R^u \quad \textbf{(2-5)}$$

$$\boldsymbol{p} = [p_1, p_2, p_3, \dots, p_v] \in R^v \quad \textbf{(2-6)}$$

Within each sub-category, the sequence is arbitrary, and the numbers of inputs, design variables, and parameters has the following relationship:

$$n = u + v \quad \textbf{(2-7)}$$

2.3 Sources of Uncertainty

Within a computational workflow, uncertainty is mainly introduced from three sources: models, unfixed design variables, and external factors.

- Uncertainty in models is due to lack of knowledge. Models are mathematical depictions of systems or phenomena, but they do not necessarily capture all the relevant information. Some models are based on empirical data, which are not direct representations of the physics behind. For instance, at conceptual stage, the weight of a wing can be defined as a portion of the total aircraft weight, based on regression analysis of previous aircrafts. However each individual design is different, thus the prediction from such an empirical model can be diverted from the reality. This is especially the case when designing aircrafts with unconventional configurations, because most of the existing data are no longer applicable as references. For physical-based models, this diversion between calculation and the reality may also exist. Many physical phenomena have not been fully understood yet (e.g. turbulence), therefore the corresponding models may contain assumptions, which are not always valid. Furthermore, in engineering design, simplification can be applied on the original mathematical relationships for practical implementation, which will also affect the accuracy of design computation (Mangeant, 2011).
- Uncertainty in unfixed design variables is due to lack of definition, which refers to "Things about the system in question that have not been decided or specified" (McManus and Hastings, 2005). Although design variables are controlled by the designer (or obtained by certain searching strategies e.g. an optimization process), it is impractical to totally freeze a design solution at early stages,

especially regarding the details. For instance, the location of a wing spar may be roughly defined so that a computational workflow can be executed for initial results. Later on, the designer may revisit this part of the design problem, and modify the spar location, given that new analysis has been conducted.

- Uncertainty in external factors is due to their "inherent variation" (Thunnissen, 2003). This type of variation is not necessarily aleatory, but tends to be unpredictable in engineering practice. External factors include operational conditions, material properties, manufacturing accuracy, etc. In a computational workflow, these factors are reflected on values of both parameters and design variables. Although the nominal values of the relevant variables may be known or specified, their variations are beyond the control of the designer, thus introducing uncertainty in the computation. For instance, due to limited manufacture accuracy the dimension of a component (after being manufactured) might be slightly different from its designed value. Note that in this example, the uncertainty is different from the one mentioned in the previous paragraph, although both of them are associated with design variables. In this case, the design variable has already been fixed and there is no uncertainty associated with the decision itself.

In practice, uncertainty can also be introduced due to specific numerical implementations (Aughenbaugh, 2006; Mangeant, 2011; Oberkampf et al., 2002b, 2002a; Thunnissen, 2003) for instance, the way of coding models, storing data, and performing calculations, etc.

Specifically in this thesis, the numerical uncertainty can be introduced due to the fact that, all the computational results will be kept with two digits after the decimal point, except for two cases:

- When the variable itself is very small but highly influential (e.g. a standard deviation on a small coefficient).
- When the accuracy of the method needs to be demonstrated (e.g. numerical test-cases in Section 5.4).

In general, the impact of numerical uncertainty/error is assumed to be negligible, because the context of the current research is within the conceptual design stage, where

most of the models being used are based on simple algebraic equations. In some specific cases, for instance, the computational workflow is based on a set of comprehensive *Partial Differential Equations* (PDE), the numerical uncertainty/error will become significant, and its influence can be represented as part of the model uncertainty, using the method discussed in the following section. The detailed quantification process is assumed to be based on expert elicitation and beyond the scope of this thesis.

2.4 Probabilistic Approach

In this PhD research, the probabilistic approach is adopted for quantification of uncertainty. The variations of uncertain variables are represented as probability distributions, and the latter are described as *Probability Density Functions* (PDF). This setup is straight forward for design variables and parameters. To represent the uncertainty associated with computational models, the method proposed in (Molina-Cristóbal et al., 2014) is adopted.

As illustrated in Figure 2-2, given a set of deterministic input variables, y_i is one of the outputs obtained from a deterministic model in the computational workflow. Due to the model uncertainty (or numerical uncertainty) as mention in Section 2.3, this calculated y_i is likely to be deviated from the true response of the system or phenomenon under study. To address this deviation (which is unknown), y_i is multiplied with a random variable rv_{y_i}, and the final output y_{i_R} becomes random as well:

$$y_{i_R} = rv_{y_i} \cdot y_i \tag{2-8}$$

For each output variable, a different rv_{y_i} will be assigned. All the rv_{y_i}'s are assumed to be scalars. The probability distributions of rv_{y_i}'s can be obtained from historical data (using physical experiments) or expert elicitation (as a subjective estimation of the model discrepancy). Mathematically, those random variables are equivalent to other uncertain parameters, therefore they will be included in the vector $\boldsymbol{p}$, or in a more general case, as elements in the vector $\boldsymbol{x}$. In this thesis, all the PDF's are assumed to be independent and available.

It should be mentioned that a more systematic approach to represent model uncertainty is to further distinguish between model structure uncertainty and model parameter

uncertainty. The former is due to inaccurate model forms and potential missing terms in the equations, while the latter is caused by possible variations of the coefficients in the model (Kiureghian and Ditlevsen, 2009; Renard et al., 2011). In this paradigm, apart from multiplicative random variables, additive rv_{y_i}'s can also be used, especially to account for the model structure uncertainty. As a result, several random variables may be assigned for a single output as calibration factors. The implantation of this systematic approach requires specific domain knowledge and detailed investigation into the models themselves, therefore is beyond the scope the current research.

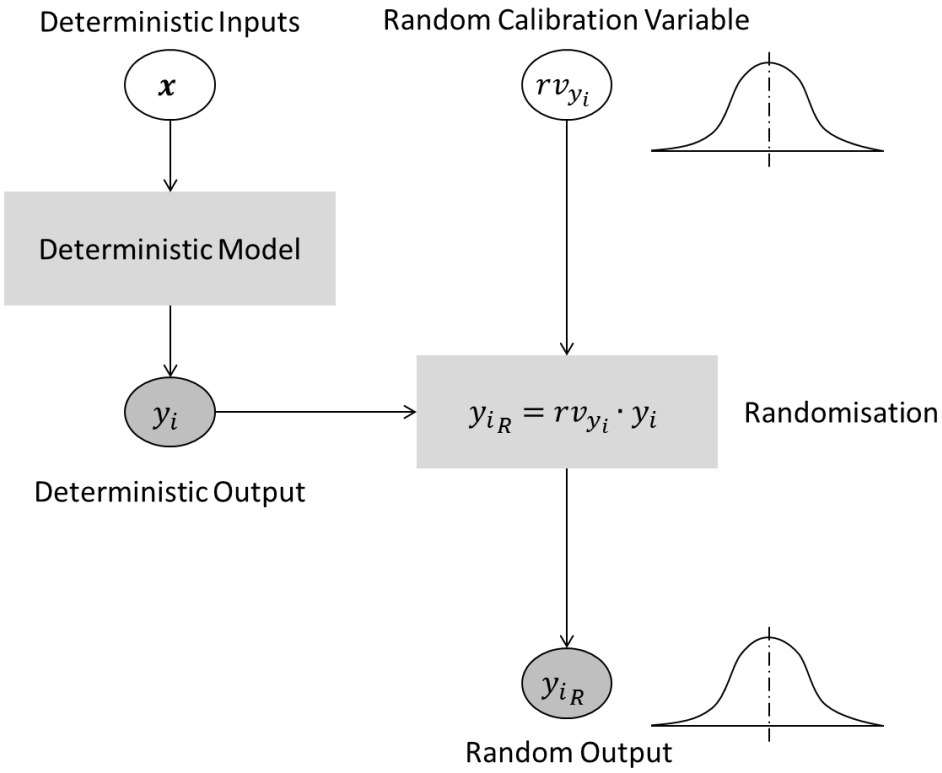

Figure 2-2. Representation of model uncertainty (Adopted from (Molina-Cristóbal et al., 2014))

2.5 Formulation of Uncertainty-Based Design Problem

Within the framework of the probabilistic approach, two paradigms are widely adopted to formulate an uncertainty-based design problem: *Robust Design Optimization* (RDO) and *Reliability-Based Design Optimization* (RBDO). It should be mentioned that, although the terminology "optimization" is adopted, our focus here is not on searching for optimal solutions, but on the common structure to define a problem.

As illustrated in Figure 2-3 (Huyse and Bushnell, 2001), a RDO problem intends to minimise the performance variation due to "everyday fluctuations" of uncertain variables, while a RBDO problem aims to reduce the probabilities of "infrequent but potentially catastrophic events" (Huyse and Bushnell, 2001).

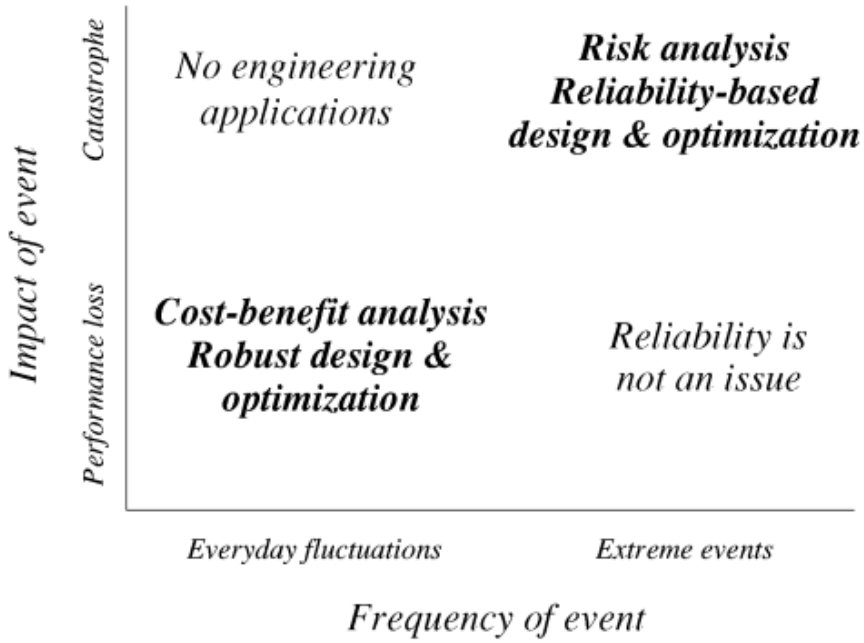

Figure 2-3. Application scenarios of RDO and RBDO (Huyse and Bushnell, 2001)

This distinction is also reflected as different areas of interest in a relevant probabilistic distribution. As shown in Figure 2-4 (Zang et al., 2002), a RDO problem focuses more on the area near the nominal value; while in a RBDO problem, more emphases are placed on the tails.

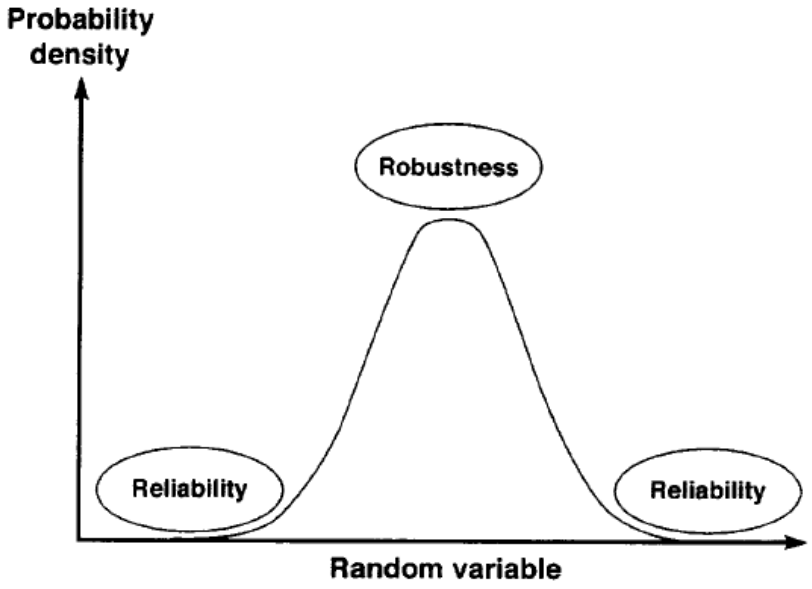

Figure 2-4. Areas of interest of RDO and RBDO (Zang et al., 2002)

2.5.1 Robust Design Optimization

It has been widely accepted (Huyse and Bushnell, 2001; Keane and Nair, 2005; Park et al., 2006; Schuëller and Jensen, 2008; Yao et al., 2011) that RDO is originated from the pioneer development of Taguchi in 1940s. The philosophy of Taguchi is to make the performance of a product (or a process) insensitive to the so-called noise factors (Phadke, 1989; Taguchi, 1986). To measure the quality of a product, a *Loss Function* (Phadke, 1989; Taguchi, 1986) can be defined as:

$$F_l(y_i) = k_{T_{y_i}}(y_i - y_i^*)^2 \quad \textbf{(2-9)}$$

In this function, y_i is one of the output (performance) variables and y_i^* is its target value. The coefficient $k_{T_{y_i}}$ is dependent on the costs of rework or scrap.

Considering the variation of y_i, the expectation of $F_l(y_i)$ (Phadke, 1989) is given by:

$$E[F_l(y_i)] = k_{T_{y_i}}[{\sigma_{y_i}}^2 + \left(\mu_{y_i} - y_i^*\right)^2] \quad \textbf{(2-10)}$$

From this equation, it can be noted that the critical parts are the variance of the performance ${\sigma_{y_i}}^2$, and the gap between the mean and target $\left(\mu_{y_i} - y_i^*\right)^2$. Because there are two objectives, the concept of *Signal to Noise* (S/N) ratio is adopted as an measure of robustness (Phadke, 1989; Taguchi, 1986). If the target is finite and positive, the

scenario is referred to as 'nominal-the-better'. If the target is zero or infinite, the scenarios can be defined as 'smaller-the-better' and 'larger-the-better', respectively. Corresponding to these scenarios, the S/N ratio (Phadke, 1989) is defined as:

$$\text{Smaller-the-better: } \eta_{y_i} = -10\, log_{10}\left[\frac{1}{N}\sum_j^N (y_i^{(j)})^2\right]$$

$$\text{Nominal-the-better: } \eta_{y_i} = 10\, log_{10}\left(\frac{{\mu_{y_i}}^2}{{\sigma_{y_i}}^2}\right) \qquad \textbf{(2-11)}$$

$$\text{Larger-the-better: } \eta_{y_i} = -10\, log_{10}\left[\frac{1}{N}\sum_j^N \frac{1}{(y_i^{(j)})^2}\right]$$

Here the parameter N is the number of samples drawn from the production, and $y_i^{(j)}$ are the j_{th} sample of y_i. With this formulation, the goal becomes searching for a design which achieves highest S/N ratio.

Modern RDO in the probabilistic framework has adopted the setup to "optimise the mean" and "minimise the variance" (Chen, Wiecek and Zhang, 1999). Mathematically, the problem can be reformulated as (Park et al., 2006; Yao et al., 2011):

Find:

$$\boldsymbol{dv} \in R^u \qquad \textbf{(2-12)}$$

To minimize:

$$F_{obj}[\mu_{F_{obj}^D}(\boldsymbol{dv}, \boldsymbol{p}), \sigma_{F_{obj}^D}(\boldsymbol{dv}, \boldsymbol{p})] \qquad \textbf{(2-13)}$$

Subject to:

$$C_i(\boldsymbol{dv}, \boldsymbol{p}) \leq 0, i = 1,2,3, \dots, t \qquad \textbf{(2-14)}$$

In equation (2-13), F_{obj}^D is the objective function of the deterministic problem, $\mu_{F_{obj}^D}$ is the mean value, while $\sigma_{F_{obj}^D}$ is the standard deviation. $F_{obj}[\mu_{F_{obj}^D}(\boldsymbol{dv}, \boldsymbol{p}), \sigma_{F_{obj}^D}(\boldsymbol{dv}, \boldsymbol{p})]$ is a function which combines these two statistical values, for example, as a weighted sum of $\mu_{F_{obj}^D}$ and $\sigma_{F_{obj}^D}$. $C_i(\boldsymbol{dv}, \boldsymbol{p}) \leq 0$ is the i_{th} constraint.

2.5.2 Reliability-Based Design Optimization

The RBDO focuses more on the constraint satisfaction. A typical RBDO problem can be formulated as:

Find:

$$\boldsymbol{dv} \in R^u \tag{2-15}$$

To minimize:

$$F_{obj}(\boldsymbol{dv}, \boldsymbol{p}) \tag{2-16}$$

Subject to:

$$\mathbb{P}\{C_i(\boldsymbol{dv}, \boldsymbol{p}) \leq 0\} \geq \mathbb{p}_{C_i}, i = 1,2,3, \dots, t \tag{2-17}$$

Here$\mathbb{P}\{C_i(\boldsymbol{dv}, \boldsymbol{p}) \leq 0\}$ is the probability of satisfying constraint $C_i(\boldsymbol{dv}, \boldsymbol{p}) \leq 0$. This is also referred to as "robustness of constraints", and is an overlapping area between RDO and RBDO (Park et al., 2006). $\mathbb{p}_{C_i}$ is a threshold value set by the designer for the minimum probability to be achieved.

In design practice, equation (2-17) can be rewritten as a combination of the mean and standard deviation (Du and Chen, 2002; Padulo, 2009; Park et al., 2006; Parkinson, Sorensen and Pourhassan, 1993):

$$\mu_{C_i}(\boldsymbol{dv}, \boldsymbol{p}) + k_{P_{C_i}}\sigma_{C_i}(\boldsymbol{dv}, \boldsymbol{p}) \leq 0, i = 1,2,3, \dots, t \tag{2-18}$$

In this equation, the constant $k_{P_{C_i}}$ is dependent on to the value of $\mathbb{p}_{C_i}$ in equation (2-17). In (Du and Chen, 2002; Parkinson, Sorensen and Pourhassan, 1993), the constraint is assumed to follow a normal distribution, which provides the relationship:

$$k_{P_{C_i}} = \Phi^{-1}(\mathbb{p}_{C_i}) \tag{2-19}$$

Here $\Phi^{-1}(\mathbb{p}_{C_i})$ is the inverse CDF of a standard normal distribution.

In (Padulo and Guenov, 2011), more general assumptions are considered, and the resulting relationships are shown in Table 2-2.

Table 2-2. Coefficient $k_{P_{C_i}}$ as a function of the probability $\mathbb{p}_{C_i}$ (Padulo and Guenov, 2011)

Distribution Assumptions	$k_{P_{C_i}}(\mathbb{p}_{C_i})$	Validity
None	$k_{P_{C_i}} = \sqrt{\frac{\mathbb{p}_{C_i}}{1-\mathbb{p}_{C_i}}}$	$0 \leq \mathbb{p}_{C_i} \leq 1$
Symmetry	$k_{P_{C_i}} = \sqrt{\frac{1}{2(1-\mathbb{p}_{C_i})}}$	$\mathbb{p}_{C_i} \geq \frac{1}{2}$
Unimodality	$k_{P_{C_i}} = \sqrt{\frac{9\mathbb{p}_{C_i}-5}{9(1-\mathbb{p}_{C_i})}}$	$\mathbb{p}_{C_i} \geq \frac{5}{6}$
	$k_{P_{C_i}} = \sqrt{\frac{3\mathbb{p}_{C_i}}{(4-3\mathbb{p}_{C_i})}}$	$\mathbb{p}_{C_i} < \frac{5}{6}$
Symmetry+Unimodality	$k_{P_{C_i}} = \sqrt{\frac{2}{9(1-\mathbb{p}_{C_i})}}$	$\mathbb{p}_{C_i} \geq \frac{1}{2}$

2.6 AirCADia Software

AirCADia is an in-house model-based design tool, developed by the Advanced Engineering Design Group in Cranfield University (Guenov et al., 2014a, 2014b).

In AirCADia, computational models can be created within the software, or imported as .dll plug-ins (e.g. MATLAB, FMI (Blochwitz et al., 2011), etc.). The software is capable of assembling the models as computatioal workflows automatically. It also contains inbuilt enablers for performing optimization, uncertainty analysis, as well as visualization of the results.

In this research, the case studies in Chapter 4 and 5 are implemented with AirCADia. Specifically, some visualization techniques (as presented in Section 4.2.2 and 4.2.3) are utilised in the proposed method for margin allocation in Chapter 4. A workflow reversal capability (as presented in Section 5.2.2) is utilised in the proposed method for inverse uncertainty propagation in Chapter 5.

3 Literature Review

3.1 Introduction

The general UQ&M problem areas were briefly reviewed in Section 1.1. This chapter focuses on three specific topics, aligned with the motivations of this research. In Section 3.2, methods for margin allocation and trade-off are presented, specifically within the UQ&M context. Section 3.3 presents research on uncertainty allocation. Because the problem itself has not been fully defined, the literature is separated in several different research communities. The development of sensitivity analysis is presented in Section 3.4, specifically regarding the so-call *Variance-Based Sensitivity Indices* or *Sobol' Indices*. This section includes the general mathematical formulation of these sensitivity indices, the conventional approach of calculation, and existing research aiming to reduce the relevant computational cost. Finally, the conclusions of this literature review are given in Section 3.5.

3.2 Margin Allocation and Trade-off

Margins are widely applied as mitigation of uncertainty. They could be allocated on technical parameters (e.g. mass, power consumption, etc.), or from the perspective of project management (e.g. on budgets, schedules, etc.). The philosophy behind is that the system (or the process) is designed "to be more capable" than "necessary" (McManus and Hastings, 2005), so that possible design changes or over-optimistic estimations could be accommodated.

3.2.1 Definition of Margin

In general, a margin can be defined as a "reserve" on a relevant quantity. In this thesis, only scalar quantities are considered. For certain purposes, the designer may be interested in vectors or fields of scalars/vectors (e.g. lift distribution over the wing surface). This is beyond the scope of this research.

Within the computational context, consider z as a general scalar variable, which can be either an input or an output. The value of z is uncertain, but its nominal value z_N is known. A margin can then be defined as a distance as shown in the following equations:

$$Mar_z = z_C - z_N \quad \text{Lower the better}$$
$$Mar_z = z_N - z_C \quad \text{Higher the better,} \tag{3-1}$$

or as a percentage:

$$Mar_z = [(z_C - z_N)/\, z_N] \cdot 100\% \quad \text{Lower the better}$$
$$Mar_z = [(z_N - z_C)/\, z_N] \cdot 100\% \quad \text{Higher the better} \tag{3-2}$$

Here z_C is a conservative estimation of z (on the pessimistic side) (Thunnissen, 2004), or a threshold value (e.g. a constraint) (Helton, 2009). It should be noted that, in this thesis, the variable z is considered to either "lower the better" or "higher the better". In the case of "nominal the better", two margins may be allocated simultaneously on both side of z_N. This is not demonstrated in this thesis.

In engineering design, margins can be classified into two categories: safety margins and design margins. The former could be regarded as synonym of safety factors; once decided, they remain fixed during the design process. The choices of this type margins are normally based on regulations, with less freedom for decision making. On the other hand, the design margins may vary "throughout the development" (Thunnissen, 2005), and depend more on the designers' knowledge and experience. The magnitude of a margin indicates a deterministic estimation of uncertainty. Traditionally this is "an afterthought" and lumps all the uncertainties "into one value with little or no analysis" (Thunnissen, 2005). For instance, given the nominal value (z_N), the true value of z is unknown (or keeps on changing) within a range. Ideally the margin should be chosen to cover this possible variation (z is considered to be 'lower the better'), as illustrated in Figure 3-1 (adapted from (Pilch, Trucano and Helton, 2011)). In practice, due to subjectivity, the margin can be over- or under-estimated, which leads to over-conservative (thus non-competitive) or over-radical (thus risky) design solutions.

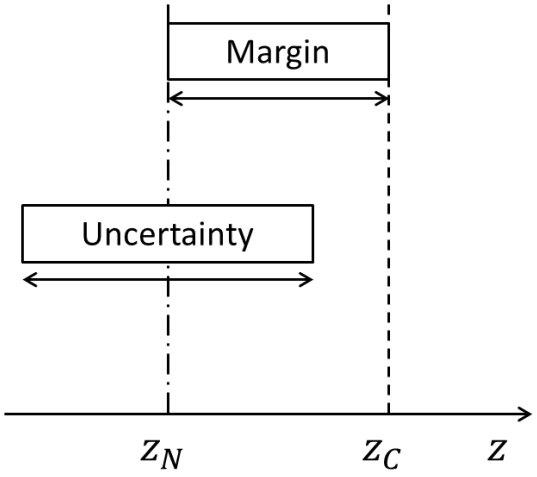

Figure 3-1. Margin as a deterministic estimation of uncertainty (adapted from (Pilch, Trucano and Helton, 2011))

3.2.2 Probabilistic Formulation of Margins

Assuming that the uncertain variable (z) follows a probability distribution, and that the PDF ($\mathbb{f}_z$) is available, then the margin (Mar_z) can be formulated based on a percentile (z_P) of the relevant distribution (Thunnissen, 2004, 2005; Thunnissen and Tsuyuki, 2004):

$$Mar_z = z_P - z_N \tag{3-3}$$

In the form of a percentage, equation (3-3) becomes:

$$Mar_z = \left[\frac{z_P - z_N}{z_N}\right] * 100\% \tag{3-4}$$

The probability of not exceeding this percentile (z_P) is given by:

$$\mathbb{P}\{z \leq z_P\} = \mathbb{F}_z(z_P) = \int_{-\infty}^{z_P} \mathbb{f}_z(z)dz \tag{3-5}$$

Here $\mathbb{F}_z(z_P)$ is the CDF of z_P. In this way, an analytical relationship can be established between the magnitude of a margin and the corresponding probability of maintaining this margin, as illustrated in Figure 3-2 (if z_P is considered as a constraint, this becomes the probability of constraint satisfaction).

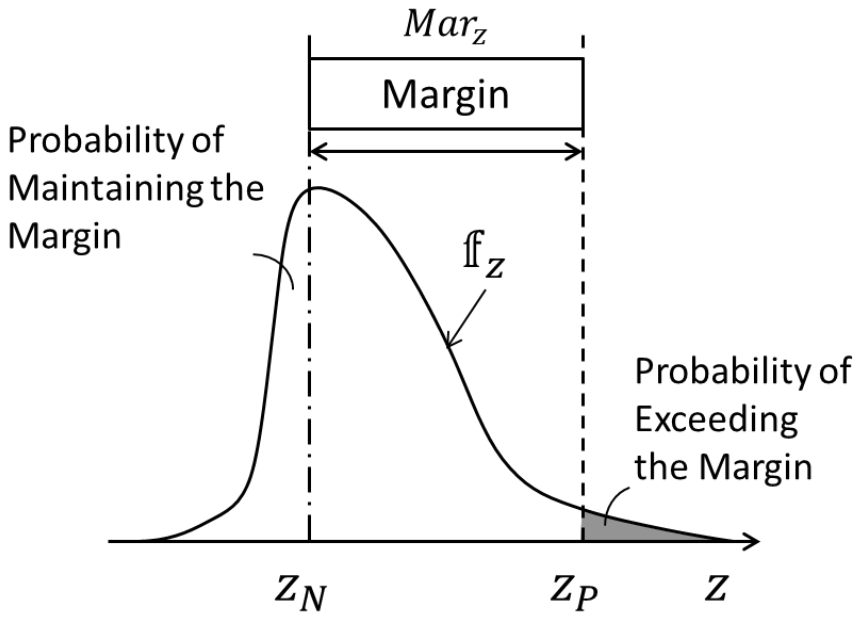

Figure 3-2. Relationship between margin and uncertainty

3.2.3 Margin Allocation

In practice, two possible design scenarios could be adopted, as illustrated in Figure 3-3.

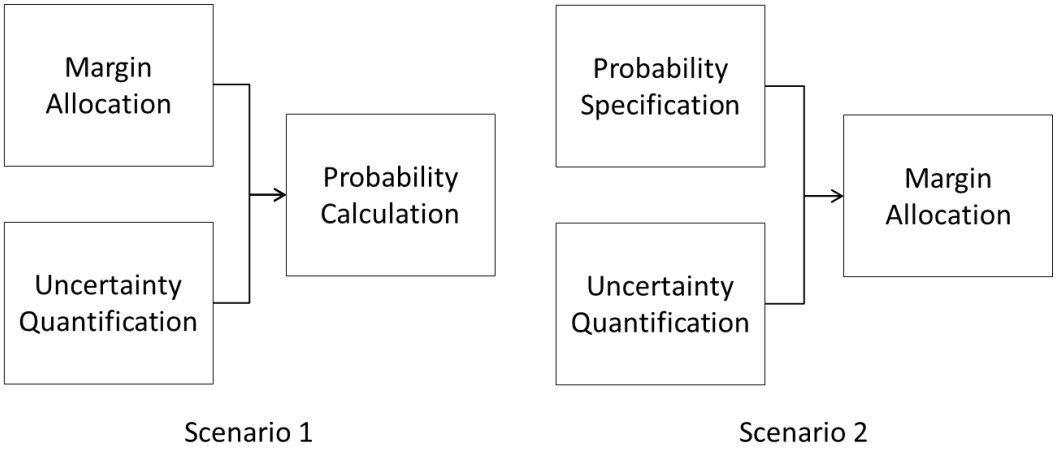

Figure 3-3. Possible design scenarios for margin allocation

Scenario 1: This scenario is adopted in (Birman, 2013; Cooke et al., 2015; Guenov, Nunez and Gondhalekar, 2011; Zang et al., 2015), where a margin is allocated *a priori*, based on experience or on historical data. By using this specified margin, as well as the results from uncertainty quantification, the designer would be able to assess the probability of maintaining such a margin (or probability of constraint satisfaction):

$$\mathbb{P}\{z \leq z_N + Mar_z\} = \mathbb{F}_z(z_N + Mar_z) = \int_{-\infty}^{z_N + Mar_z} \mathbb{f}_z(z)dz \quad \textbf{(3-6)}$$

Scenario 2: This scenario is adopted in (Thunnissen, 2004, 2005; Thunnissen and Tsuyuki, 2004), where uncertainty quantification and propagation are conducted *a priori* to obtain the relevant PDF's. Then the designer may specify a probability $\mathbb{p}_z$, which he/she wants to achieve. Based on this probability, the corresponding percentile is calculated as:

$$z_P = \mathbb{F}_z^{-1}(\mathbb{p}_z) \quad \textbf{(3-7)}$$

Substituting equation (3-7) into equation (3-3), the margin is given as:

$$Mar_z = \mathbb{F}_z^{-1}(\mathbb{p}_z) - z_N \quad \textbf{(3-8)}$$

Here $\mathbb{F}_z^{-1}(\mathbb{p}_z)$ is the inversed CDF of z.

In practice, $\mathbb{f}_z$, $\mathbb{F}_z$, and $\mathbb{F}_z^{-1}$ are difficult to be obtained analytically, therefore numerical methods are normally adopted (e.g. MCS in (Thunnissen, 2004, 2005; Thunnissen and Tsuyuki, 2004; Yuan et al., 2016), method of moments in (Birman, 2013), URQ in (Guenov, Nunez and Gondhalekar, 2011), and Bayesian Belief Networks in (Cooke et al., 2015; Zang et al., 2015)).

3.2.4 Margin Trade-off

Most of the research (Cooke et al., 2015; Guenov, Nunez and Gondhalekar, 2011; Thunnissen, 2004, 2005; Thunnissen and Tsuyuki, 2004; Yuan et al., 2016) has addressed the trade-off between a margin and a probability of constraint satisfaction. In (Yuan et al., 2016; Zang et al., 2015), this trade-off is implicitly embedded in an optimization process of the margins.

In (Cooke et al., 2015), a set based design approach is used to down select of multiple margin combinations. In this method, the trade-off between margins and performances are partially addressed, however the relationship between margin and performance is not explicitly captured.

3.2.5 Discussion

According to the literature review, several limitations can be identified regarding the existing methods for margin management:

Firstly, the existing research only considers the margins to accommodate model uncertainties, while uncertainties associated with input variables and constraints are not accounted for (e.g., to accommodate changes in design requirements).

Secondly, the aforementioned research focuses on combing margins with uncertainty quantification and propagation, while the aspect of trade-offs is relatively less-developed, especially regarding the trade-off (interaction) between different margins.

3.3 Uncertainty Allocation

Uncertainty allocation is not a conventional problem, compared with other topics under UQ&M (such as uncertainty propagation, sensitivity analysis, RDO, RBDO, etc.). A similar problem, within the context of design methodology, is the *Uncertainty Budgeting*, as presented in Section 3.3.1. Within the manufacturing research communities, relevant topics include *Tolerance Design* and *Tolerance Allocation*, which are presented in Section 3.3.2 and 3.3.3, respectively. Section 3.3.4 presents several methods for *Inverse Uncertainty Propagation*, which is a specific sub-problem of interest, under uncertainty allocation. The relevant research is mainly from the modelling and simulation communities.

3.3.1 Uncertainty Budgeting

The concept of uncertainty budgeting is to "spend" uncertainty as a "currency" (Opgenoord, 2016; Opgenoord and Willcox, 2016). The method assumes that a specific amount of uncertainty is accepted in the design process, and more uncertainty should be "spend" on the factors which are less influential to the output performances.

Specifically, the "budgets" of uncertainty include three aspects: probabilities of exceeding constraints ($1 - \mathbb{p}_{C_i}$), upper bounds of performance standard deviations ($\sigma_{y_i}^{UB}$), and costs of reducing uncertainty ($\mathbb{C}$). These budgets are predefined by a designer, and the problem is formulated as an optimization.

Find:

$$\boldsymbol{\mu_x}, \boldsymbol{\sigma_x} \in R^n \tag{3-9}$$

To minimise:

$$F_{obj}(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x}) = \sum_{i=1}^{t} w_{1i} \frac{1 - \mathbb{P}\{C_i(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x}) \leq 0\}}{1 - \mathbb{p}_{C_i}} + \sum_{j=1}^{l} w_{2j} \frac{\sigma_{y_j}(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x})}{\sigma_{y_j}^{UB}} + w_3 \frac{F_C(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x})}{\mathbb{C}} \tag{3-10}$$

Subject to:

$$\mathbb{P}\{C_i(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x}) \leq 0\} \geq \mathbb{p}_{C_i}, i = 1,2,3, \dots, t \tag{3-11}$$

$$\sigma_{y_j}(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x}) - \sigma_{y_j}^{UB} \leq 0, j = 1,2,3, \dots, l$$

$$F_C(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x}) - \mathbb{C} \leq 0$$

In this formulation, $\boldsymbol{\mu_x}$ and $\boldsymbol{\sigma_x}$ are the vectors of input means and standard deviations. The output standard deviation σ_{y_i} is defined as a function of $\boldsymbol{\mu_x}$ and $\boldsymbol{\sigma_x}$. This calculation is enabled by constructing a *High-Dimensional Model Representation*-based (HDMR) surrogate. On the other hand, to obtain $\mathbb{P}\{C_i(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x})$, a base-line MCS is first conducted on the surrogate model, which produces an initial probability $\mathbb{P}_0$. Then partial derivatives are used to linearize and approximate the change of this probability due to $\boldsymbol{\mu_x}$ and $\boldsymbol{\sigma_x}$, respectively. The cost model $F_C(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x})$ needs to be specified by the designer. w's are the weight factors. In the optimization process, sensitivity analysis is used to identify which input variables should be updated.

3.3.2 Tolerance Design

The tolerance design is part of the Taguchi method (Phadke, 1989; Taguchi, 1986) as mentioned in Section 2.5.1. Unlike the RDO problem, tolerance design is conducted after the values of design variables are fixed. Instead of exploring the design space, restrictions are directly placed on the uncertain factors. Within the context of the Taguchi method, such restrictions are realized by improving manufacture accuracy or deploying parts/components of higher quality. Since the product cost will be increased by the aforementioned activities, a trade-off shall be sought between robustness and the cost of upgrading parts/components.

In the Taguchi method, this process employs *Design of Experiments* (DoE) to assess the performance variances. *Analysis of Variance* (ANOVA) is used to identify the most influential noise factors. The aforementioned trade-off is conducted by evaluating different options to reduce the variances of the most influential uncertain factors (e.g. by choosing from a set of off-the-shelf components with different prices and qualities). The objective is to minimise the overall cost, which is calculated by adding the price of the chosen component and the quality loss of the product (as defined in equation (2-6)). It should be mentioned that, the "performance variation" in the Taguchi method is mathematically different from the "statistical variance" in the context of probability theory.

3.3.3 Tolerance Allocation

Tolerance allocation is slightly different from the concept of Taguchi's tolerance design. In the context of the tolerance allocation, a "tolerance" specifically refers to the manufacturing accuracy of component dimensions (while Taguchi's tolerance can also refer to other physical quantities, such as electrical resistance, ingredient concentration, and so on).

Methods for tolerance allocation are normally based on local sensitivity analysis (partial derivatives) to obtain an explicit relationship between component tolerances and overall assemble tolerance (Chase, 1999). To choose the optimal tolerance for each component, optimization process can be used, as presented in Lööf et al (Lööf, Hermansson and Söderberg, 2007; Lööf and Söderberg, 2011).

3.3.4 Inverse Uncertainty Propagation

The inverse uncertainty propagation refers to estimating the uncertainty of input variables, given the known or pre-defined uncertainty associated with system (model) outputs. Such a problem is also referred to as *Backward Uncertainty Propagation* in (Congedo et al., 2012), *Stochastic Inverse Problem* in (Narayanan and Zabaras, 2004; Zabaras and Ganapathysubramanian, 2008), and *Uncertainty Identification* in (Fonseca et al., 2005).

It should be mentioned that distinctions need to be made between the inverse uncertainty propagation and several similar, but not identical problems in the literature.

For instance, the *Inverse Uncertainty Quantification* (Litvinenko and Matthies, 2013), refers to estimating the unknown parameters of a specific probability distribution, based on a set of samples drawn from this distribution. In (Jiang, Liu and Han, 2008; Ngnepieba and Hussaini, 2007; Wang and Zabaras, 2004), the problem is to estimate the unknown parameters of a specific model, given the uncertain outputs of the model. These model parameters themselves are deterministic, but the estimations will follow distributions, due to other sources of uncertainty (e.g. measurement uncertainty). In (Abusam, Keesman and van Straten, 2003), the term *Backward Uncertainty Propagation* (different from (Congedo et al., 2012)) refers to identification of regions of the input space, corresponding to the regions of interest within an output distribution. These problems (Abusam, Keesman and van Straten, 2003; Jiang, Liu and Han, 2008; Litvinenko and Matthies, 2013; Ngnepieba and Hussaini, 2007; Wang and Zabaras, 2004) are not within the scope of the current research. The reader is referred to (Tarantola, 2005), for a discussion of *Inverse Problems* within a more general context.

Within our scope, three approaches for inverse uncertainty propagation are identified:

The first approach requires the explicit relationship between the input and output statistical moments, so that when the output uncertainty is specified, the input uncertainty can be calculated accordingly. However, this relationship is normally inaccessible in most practical applications. Therefore a method has been proposed in (Baumgärtel et al., 2014) to approximates this relationship based on a Gaussian process as a surrogate (Girard and Murray-Smith, 2005).

The second approach is based on maximum likelihood method (Fonseca et al., 2005). The uncertain input variables are firstly assumed to follow a multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu_x}, \boldsymbol{\Sigma_x})$, where $\boldsymbol{\mu_x}$ and $\boldsymbol{\Sigma_x}$ are the unknown vector of input means and covariance matrix, respectively. Given a set of data from model calculation or observation, the likelihood function of $\boldsymbol{\mu_x}$ and $\boldsymbol{\Sigma_x}$ can be constructed by using perturbation method (based on first order Taylor expansion) or Monte-Carlo simulation. The problem is then transformed into an optimization to maximize the likelihood function. The optimal solutions are the most likely $\boldsymbol{\mu_x}$ and $\boldsymbol{\Sigma_x}$, which recover the unknown input distribution.

The third approach is to minimise the difference between the actual and target output distributions, by optimizing the input ones (Congedo et al., 2012; Narayanan and Zabaras, 2004; Zabaras and Ganapathysubramanian, 2008).

In (Zabaras and Ganapathysubramanian, 2008) and some early development in (Narayanan and Zabaras, 2004), thermal conduction within a heat sink is considered. In the forward problem, a stochastic temperature profile is solved, given the stochastic heat flux from the boundary of the region under study. The heat flux is represented as $q(\boldsymbol{s}, t, \boldsymbol{\xi}_q)$, where $\boldsymbol{s}$ is the vector to define the location in the space domain, t is the time, and $\boldsymbol{\xi}_q$ is a finite-length vector of random variables to replace the original probability space of q, using the *Karhunen–Loève Expansion* (KLE) (Ghanem and Spanos, 1991). Then a collocation approach is applied to approximate the heat flux as a polynomial interpolant:

$$q(\boldsymbol{s}, t, \boldsymbol{\xi}_q) = \sum_{i=1}^{n_q} q(\boldsymbol{s}, t, \boldsymbol{\xi}_q^{\boldsymbol{i}})\psi_i(\boldsymbol{\xi}_q), \tag{3-12}$$

where n_q is the number of collocation points, $\boldsymbol{\xi}_q^{\boldsymbol{i}}$ is one realization of $\boldsymbol{\xi}_q$ (at one collocation point) over the stochastic space, and ψ_i is the corresponding polynomial. In this equation, $q(\boldsymbol{s}, t, \boldsymbol{\xi}_q^{\boldsymbol{i}})$ is a deterministic function, which will be noted as q_i hereafter. Based on the known q_i's, the stochastic temperature in space and time can be deducted.

The inverse problem is to re-construct/design the heat flux, to maintain an observed/specified temperature profile. Assuming that the temperatures at n_s locations $(\boldsymbol{s_i}, i = 1,2,3, \dots, n_s)$ in space are of interest, an optimization is used to find $q_i, i = 1,2,3, \dots, n_q$ which will minimize:

$$F_{obj}\left(q_1, q_2, q_3, \dots, q_{n_q}\right) = \frac{1}{2}\sum_{i=1}^{n_s}\int_{t=0}^{t_{max}}\left[\sum_{j=1}^{\beta} w_j\left(\langle \mathcal{T}^j(\boldsymbol{s_i}, t)\rangle - \langle \mathcal{T}^j(\boldsymbol{s_i}, t)\rangle^*\right)^2\right]dt \tag{3-13}$$

or

$$F_{obj}\left(q_1, q_2, q_3, \dots, q_{n_q}\right) = \frac{1}{2}\sum_{i=1}^{n_s}\int_{t=0}^{t_{max}}\int_0^1 [\mathbb{F}_{\mathcal{T}}^{-1}(\boldsymbol{s_i}, t, \mathbb{p}) - \mathbb{F}_{\mathcal{T}}^{-1^*}(\boldsymbol{s_i}, t, \mathbb{p})]\, d\mathbb{p} \tag{3-14}$$

In equation (3-13), $\langle \mathcal{T}^j(\boldsymbol{s_i}, t)\rangle$ and $\langle \mathcal{T}^{j^*}(\boldsymbol{s_i}, t)\rangle$ are respectively the actual and target j_{th} moment of the temperature at location $\boldsymbol{s_i}$ and time t. w_j is a weight factor for the j_{th} moment. This objective function will be used when the information is available only regarding the statistical moments. If the shape of the target probability distribution is known, the objective function defined in equation (3-14) will be used, where $\mathbb{F}_{\mathcal{T}}^{-1}(\boldsymbol{s_i}, t, \mathbb{p})$ and $\mathbb{F}_{\mathcal{T}}^{-1^*}(\boldsymbol{s_i}, t, \mathbb{p})$ are respectively the actual and target inverse CDF of the temperature at location $\boldsymbol{s_i}$ and time t, and $\mathbb{p}$ is the probability as an input to the inverse CDF. For both objective functions, the optimization is solved using a gradient-based method, the reader is refer to (Narayanan and Zabaras, 2004; Zabaras and Ganapathysubramanian, 2008) for details of calculating the directional derivatives of these objective functions.

Presented in (Congedo et al., 2012) is a study of rarefaction shock waves. The sources of uncertainty include the initial flow conditions and the thermodynamic models being used. Consequently, the stochastic output is the Mach number at a specific location of interest. In this research, inverse propagation is used to estimate the maximum amount of uncertainty from different sources, which can ensure the Mach number to be higher than one (as a condition for the occurrence of rarefaction shock waves). This problem is proposed as,

To find:

$$\Delta \boldsymbol{x} = [\Delta x_1, \Delta x_2, \Delta x_3, \dots, \Delta x_n], \tag{3-15}$$

which minimize:

$$F_{obj_1}(\Delta \boldsymbol{x}) = \left|\langle y_j^r\rangle - \langle y_j^r\rangle^*\right|, \tag{3-16}$$

and maximize:

$$F_{obj_2}(\Delta \boldsymbol{x}) = \|\Delta \boldsymbol{x}\|, \tag{3-17}$$

subject to

$$\Delta x_i \geq \Delta x_i^{LB}, i = 1,2,3, \dots, n \tag{3-18}$$

Here Δx_i is the variation of the i_{th} uncertain input (assumed to be uniformly distributed). Δx_i^{LB} is the lower bound of Δx_i, due to physical limitations. $\langle y_j^r \rangle$ and $\langle y_j^r \rangle^*$ are respectively the actual and target r_{th} moment of the j_{th} output (in this case the Mach number is the only output being considered).

In the implementation, the link between $\Delta \boldsymbol{x}$ and $\langle y_j^r \rangle$ was obtained with polynomial chaos expansions (as a forward uncertainty propagation), and the optimization was solved using a *Genetic Algorithm* (GA). To reduce the computational cost, an artificial neural network was constructed simultaneously, to replace the original CFD model, using existing model evaluations. The second objective defined in equation (3-16) was achieved by gradually reducing each Δx_i from its initial value, until the first objective function was converged. Specifically, an iterative scheme was proposed, in which the solver started with updating only the most influential uncertain input variable. If this variable reached its lower bound, one additional variable will be taken into account.

It can be seen that, the objective functions defined in (Zabaras and Ganapathysubramanian, 2008) and (Congedo et al., 2012) are closely related. The former is more general by using both the moments (equation (3-13)) and the PDF's (equation (3-14)) of the output distributions, while the in the latter, only the moments are used (equation (3-16)). It should also be noted that, in (Zabaras and Ganapathysubramanian, 2008), the optimization is setup to find an set of collocation points (further deduction is needed to obtain the input uncertainty), while in (Congedo et al., 2012), the input uncertainty are directly manipulated by the solver. Other differences are mainly regarding the solvers (gradient-based versus GA) and stochastic expansions (KL versus PCE) used in these two methods.

3.3.5 Discussion

The relationship between the topics in this section is summarized in Figure 3-4, where four domains are defined by distinctions on two dimensions.

On the vertical axis, the distinction is "what is controlled by the designer". For instance, in the classical RDO and RBDO, the amount of uncertainty remains fixed and the process is to search in the design space for promising solutions, which fulfil requirements on robustness or reliability. On the other hand, in an uncertainty allocation

or an uncertainty budgeting problem, the focus is to directly control the amount of uncertainty, regarding the trade-off between some quantities of interest and the cost of reducing uncertainty.

On the horizontal axis, a distinction is "where the method is used". From this perspective, uncertainty allocation should be considered as a design method, while tolerance allocation is a method for manufacturing process. The Taguchi method (which includes parameter design and tolerance design) has covered all the four domains, while uncertainty budgeting covers the two domains on the left (because in the relevant method (Opgenoord, 2016; Opgenoord and Willcox, 2016), both the means and standard deviations are controlled by the designer).

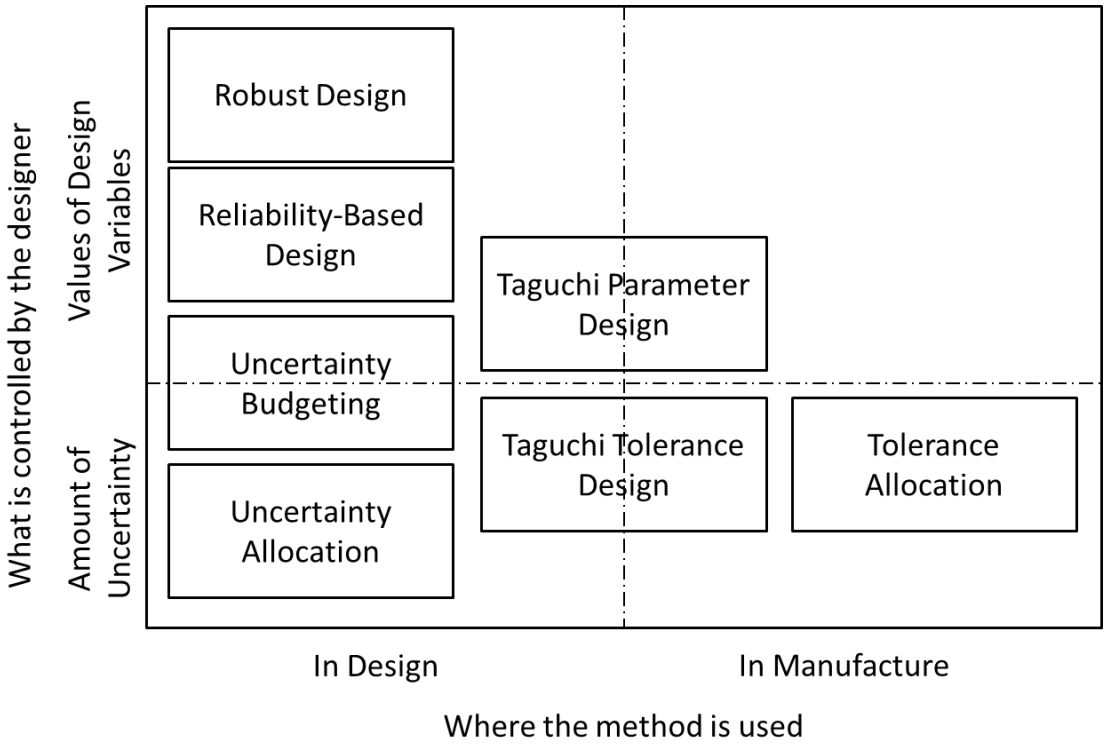

Figure 3-4. Domains of several related topics

From mathematical perspective, uncertainty budgeting, tolerance design, and tolerance allocation are proposed as forward problems. While in uncertainty allocation, an inversed approach is sought, and the latter is enabled by the inverse uncertainty propagation (which is not shown in Figure 3-4). The benefit of using the inversed approach is that the designer can ask 'what-if' type of questions, by interactively specifying the output uncertainty. While in the forward approaches, more effort is spent

on searching for the final solution in one go, rather than on the exploration process. At the early design stages, the latter could be of greater importance, as it brings the design knowledge of what can be potentially achieved, regarding the reduction of uncertainty.

Regarding the existing methods (Baumgärtel et al., 2014; Congedo et al., 2012; Fonseca et al., 2005; Narayanan and Zabaras, 2004; Zabaras and Ganapathysubramanian, 2008) for inverse propagation, a general limitation is the difficulty to set up the problem. For instance, a Gaussian process/collocation expansion/likelihood functions/cost functions need to be constructed for the specific case study, which requires extensive work before the inverse propagation can be conducted. Another aspect is the computational cost, which will become a burden for high dimensional problems.

3.4 Sensitivity Analysis

Various techniques have been developed for sensitivity analysis. In general, these techniques can be classified as: *Local Sensitivity Analysis* (LSA) and *Global Sensitivity Analysis* (GSA). LSA is based on the gradient at a given point of an input space, therefore the result is sensible only within the neighbourhood of the selected point (Smith, 2013). On the other hand, the GSA is able to analyse all the factors simultaneously, and captures the average effect over the entire region of interest (Saltelli et al., 2004, 2008). The GSA can be conducted with *Derivative-Based Methods* (Lamboni et al., 2013; Sobol' and Kucherenko, 2009) and *Variance-Based Methods* (also referred to as the *Sobol' Indices*). In this research, focus has been placed on the latter. For systematic review on other LSA and GSA techniques, the reader is referred to references (Saltelli et al., 2004, 2008; Smith, 2013).

3.4.1 Mathematical Definition

The development of variance-based sensitivity analysis dates back to 1970s, when the *Fourier Amplitude Sensitivity Test* (FAST) is proposed by (Cukier, 1973; Cukier, Levine and Shuler, 1978; Cukier, Schaibly and Shuler, 1975; Schaibly and Shuler, 1973). In this method, the Fourier Transformation and searching curves were used to decompose the output variances. Similar problems were also referred to as *Importance Measure* by (Hora and Iman, 1986; Iman and Hora, 1990; Ishigami and Homma, 1990; Saltelli, Andres and Homma, 1993; Saltelli and Homma, 1992); or *Top/Bottom*

Marginal Variance by (Jansen, Rossing and Daamen, 1994). In parallel, Sobol' adopted the so-called ANOVA-representation to decompose a function, so that the portions of total variance caused by different factors can be formulated separately (Sobol', 1990, 1993, 2003; Sobol' and Levitan, 1999; Sobol′, 2001). It was later pointed out by Saltelli that all these methods calculate an equivalent statistical quantity (Saltelli, Tarantola and Chan, 1999), and that with this regard, the Sobol's approach is the most general one (Homma and Saltelli, 1996).

Recently, a further generalization of these indices has been proposed by (Fort, Klein and Rachdi, 2016). In their research, a sensitivity index can be defined with a contrast function (Rachdi, 2011), and the latter is corresponding to a specific feature of a probability distribution. For instance, the classic Sobol' indices are defined with the contrast function that links to the output variance. In the general case, by constructing different contrast functions, the sensitivity indices can be extended to other statistical quantities as well, (e.g. the quantile of a distribution, as demonstrated in (Fort, Klein and Rachdi, 2016), and further developed in (Browne et al., 2017; Maume-Deschamps and Niang, 2018; Niang, 2016)). This generalized approach is referred to as the *Goal Oriented Sensitivity Analysis* (GOSA), which is able to provide further information beyond the classic sensitivity analysis.

Another generalization of the Sobol' indices is to accommodate the correlation between input variables. In the original definition (Sobol', 1990, 1993, 2003; Sobol' and Levitan, 1999; Sobol′, 2001), the sources of uncertainty (uncertain inputs) are assumed to be independent. However, this assumption cannot be maintained in many practical applications (Kucherenko, Tarantola and Annoni, 2012). To overcome this limitation, research has been conducted to extend the original method with: Gram–Schmidt orthogonalization (Bedford, 1998; Mara and Tarantola, 2012), sampling reordering (Xu and Gertner, 2007), regression-based decomposition (Xu and Gertner, 2008), and new definition of the indices (Kucherenko, Klymenko and Shah, 2017; Kucherenko, Tarantola and Annoni, 2012).

In this research, the focus is on the classic definition of the Sobol' Indices, where all the input variables are assumed to be independent. Regarding GOSA and dependent input variables, the reader is referred to the aforementioned papers for further information.

3.4.2 Formulation of Classic Sobol' Indices

The mathematical formulation of a variance-based sensitivity index is given as follows. Considering the workflow defined in equation (2-1), and for discussion purpose, instead of using a vector of outputs $\boldsymbol{y}$, only one output y is considered. Firstly one of the input variables x_i is fixed as a constant X_i, the conditional variance of output y could be calculated as,

$$V_{x_{\sim i}}\left(y|_{x_i=X_i}\right) = \iint \dots \int \left[f(x_1, x_2, \dots, x_{i-1}, X_i, x_{i+1}, \dots, x_n) - E_{x_{\sim i}}\left(y|_{x_i=X_i}\right)\right]^2 \prod_{\substack{j=1 \\ j\neq i}}^{n} \mathbb{f}_{x_j}(x_j) dx_j \quad \textbf{(3-19)}$$

Here the subscript $x_{\sim i}$ indicates that the variance/expectation is due to all inputs except the i_{th} one. In this equation, X_i is considered as a constant. However by choosing different values of X_i, the value of $V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)$ will be changed accordingly. Therefore to globally analyse the impact from x_i, the expectation of $V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)$ needs to be calculated, with regard to the randomness of X_i (which is the equivalent to the randomness of x_i, as X_i is a realization of x_i.

$$E_{X_i}\left(V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)\right) = \int \left(V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)\right) \mathbb{f}_{x_j}(X_i) dX_i \quad \textbf{(3-20)}$$

Since $E_{X_i}\left(V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)\right)$ is the expected output variance with contributions from all but the i_{th} input, its difference with the total output variance $V(y)$, could then be taken as the contribution from x_i only. Normalized by the total variance, the *First Order Sensitivity Index* for x_i is defined as,

$$S_i = \frac{V(y) - E_{X_i}\left(V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)\right)}{V(y)} \quad \textbf{(3-21)}$$

Apart from the first order indices S_i, similar formulations can also be used to deduce the *Higher Order Index* S_{ij} and *Total Effect Index* S_i^T, which will be further discussed in Section 6.2.2 and 6.2.3, respectively. It should be noted that, in the general case of multiple outputs, a different set of Sobol' indices (S_i, S_{ij}, and S_i^T) will be obtained for each variable in the vector $\boldsymbol{y} = [y_1, y_2, y_3 \dots, y_l]$. In the current discussion, such a

situation will not be demonstrated, therefore no additional notations will be introduced to distinguish Sobol' indices for multi-output functions.

3.4.3 Numerical Calculation

For most of the models used in real applications, analytical solutions of sensitivity indices are not available, therefore numerical methods are used for calculation. Here the most traditional and widely applied Pick-Freeze approach (based on Monte Carlo Simulation) is presented. This approach has been proposed separately by Sobol' (Ishigami and Homma, 1990; Sobol', 1993)

For each index, two sets of samples are needed:

$$\boldsymbol{\mathcal{S}_1} = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_i^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_i^{(2)} & \dots & x_n^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_1^{(r)} & x_2^{(r)} & \dots & x_i^{(r)} & \dots & x_n^{(r)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_1^{(N)} & x_2^{(N)} & \dots & x_i^{(N)} & \dots & x_n^{(N)} \end{bmatrix} \quad \textbf{(3-22)}$$

$$\boldsymbol{\mathcal{S}_2} = \begin{bmatrix} x_1^{(1)\prime} & x_2^{(1)\prime} & \dots & x_i^{(1)\prime} & \dots & x_n^{(1)\prime} \\ x_1^{(2)\prime} & x_2^{(2)\prime} & \dots & x_i^{(2)\prime} & \dots & x_n^{(2)\prime} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_1^{(r)\prime} & x_2^{(r)\prime} & \dots & x_i^{(r)\prime} & \dots & x_n^{(r)\prime} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_1^{(N)\prime} & x_2^{(N)\prime} & \dots & x_i^{(N)\prime} & \dots & x_n^{(N)\prime} \end{bmatrix} \quad \textbf{(3-23)}$$

In these matrixes, N is the number of points in each of the sampling sets. $x_i^{(r)}$ is the r_{th} sample value of the i_{th} variable, from the first sampling set $\boldsymbol{\mathcal{S}_1}$, while $x_i^{(r)\prime}$ is from the second sampling set $\boldsymbol{\mathcal{S}_2}$.

The first order indices can be estimated by (Ishigami and Homma, 1990; Saltelli, 2002; Saltelli et al., 2008):

$$\hat{S}_i = \frac{\hat{U} - \left[\hat{E}(y)\right]^2}{\hat{V}(y)} \quad \textbf{(3-24)}$$

In equation (3-24), $\hat{E}(y)$ and $\hat{V}(y)$ are the estimators of mean and variance, based on either sampling set $\boldsymbol{S_1}$ or $\boldsymbol{S_2}$, or both set $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$ (using a standard MCS). The critical part is to calculate $\widehat{U}$ by:

$$\widehat{U} \approx \frac{1}{N-1}\sum_{r=1}^{N} f\left(x_1^{(r)}, x_2^{(r)}, \dots, x_i^{(r)}, \dots, x_n^{(r)}\right) \cdot f\left(x_1^{(r)\prime}, x_2^{(r)\prime}, \dots x_i^{(r)}, \dots, x_n^{(r)\prime}\right) \quad \textbf{(3-25)}$$

In this equation, $f\left(x_1^{(r)\prime}, x_2^{(r)\prime}, \dots x_i^{(r)}, \dots, x_n^{(r)\prime}\right)$ takes all the input values from sampling set $\boldsymbol{S_2}$, except $x_i^{(r)}$ which is taken from set $\boldsymbol{S_1}$.

While set $\boldsymbol{S_2}$ needs to be re-produced for each of the input variables, set $\boldsymbol{S_1}$ can be reused every time. To obtain all the first order indices, $N(n+1)$ evaluations of the model are need. The value of N is chosen by the user, based on the required accuracy (the error is inversely proportional to $\sqrt{N}$) (Saltelli et al., 2008). If a purely random sampling strategy is used, a representative value of N could be 1000.

The process is similar for total effect and higher order indices. The total effect indices require one more set for their calculation (Saltelli, 2002), which leads to $N(n+2)$ total evaluations. To calculate the higher order indices, each combination of the input variables requires an extra sampling set. For instance, considering up to the second order, the total evaluation for all the first, second order and total effect indices is $N(n+2+\binom{n}{2})$, where $\binom{n}{2}$ is the number of combinations (every pair of input variables).

3.4.4 Improvement

Further research has been conducted to reduce the computational cost of variance-based sensitivity analysis. The focused areas include:

- **Sampling Strategy**: Instead of using a pure random sampling strategy, quasi-random samples can be used to speed up the convergence. For instance, the Sobol' sequences in (Saltelli et al., 2010), and Latin Hyper Cube in (Helton and Davis, 2003). An improved FAST (Tarantola, Gatelli and Mara, 2006) was proposed with the *Random Balance Design* (RBD) strategy (Satterthwaite, 1959). Later on, this sampling strategy was also introduced to the MCS

approach by (Mara and Rakoto Joseph, 2008). In these methods, the same sampling points are reused for sensitivity indices corresponding to all the input variables; therefore the total number of model evaluations is reduced.

- **Formulation of Estimator**: the estimator in equation (3-25) was improved by (Jansen, 1999; Saltelli, 2002; Sobol' et al., 2007) The new estimators can accelerate the convergence of the Monte Carlo approach (leading to a reduced N), therefore reduce the total number of model evaluations.
- **Approximation:** In (Lamboni, 2016), a method was proposed to use quadrature and Latin Hyper Cube for the total effect indices. In (Kucherenko, Klymenko and Shah, 2017), a new formulation of Sobol' indices was proposed which can be used for correlated input variables. In this research, grid quadrature is used for approximation.
- **Using Expansion:** Oakley and O'hagan proposed the Bayesian approach based on Gaussian processes (Oakley and O'Hagan, 2004). In (Blatman and Sudret, 2010; Crestaux, Le Maître and Martinez, 2009; Konakli and Sudret, 2016; Sudret, 2008), the Polynomial Chaos Expansion (PCE) was applied for sensitivity analysis. In this approach, the indices can be calculated from the coefficients of the orthogonal polynomials, while the latter is obtained by evaluating the model at specific sampling points.

3.4.5 Discussion

Regarding the traditional MCS approach, equations (3-26), (3-27), and (3-28) are adopted from (Saltelli, 2002). The cost of the improved MCS with Random Balance Design (Mara and Rakoto Joseph, 2008) is given by equation (3-29). In these equations, N is the number of samples, which will be used for a general MCS. It is chosen by the user, based on the required accuracy (the error is inversely proportional to $\sqrt{N}$) (Saltelli et al., 2008). If a purely random sampling strategy is used, a representative value of N could be 1000. Some of the modified estimators (Jansen, 1999; Saltelli et al., 2010) or quasi-random sampling methods can speed up the convergence, therefore smaller N could be used. In Figure 3-5, the computational costs against the number of uncertain input variables for all the MCS approaches are plotted as red lines, where the different settings (N values and orders of indices) are indicated by different markers. For

instance, assume that the modified estimator from (Jansen, 1999; Saltelli et al., 2010) could increase the efficiency by 50% (which is very optimistic), the cost of computing the first order indices (with $N = 500$) is indicated by the red line with asterisk markers.

The computational costs of PCE approaches are dependent on the detailed techniques for estimating the polynomial coefficients. For intrusive PCE, equation (3-30) can be used, where P is the order of truncation. The non-intrusive PCE can be further classified as full tensor approach, *Least Square Approximation* (LSA), and sparse grid approach. The cost of a full tensor approach grows exponentially as shown in (3-31), while LSA can provide the same cost as the intrusive approach (as adopted in (Cuneo, Traverso and Shahpar, 2017)). In (Eldred and Burkardt, 2009) and (Sudret, 2008), over-sampling is recommended, which increases the total number of model evaluations, as shown in equations (3-34) and (3-35). The cost of sparse grid approach can be estimated with equation (3-32), where K is the level of the grid. A technique named low-rank tensor approximation was proposed in (Konakli and Sudret, 2016), which was reported to be more efficient than the conventional LSA. The computational cost is not included in Table 3-1, as an iterative algorithm is used for sampling. The stopping criteria of this algorithm include a pre-defined tolerance for error and a number of maximum iterations. In Figure 3-5, the PCE related approaches are plotted as blue lines, with the markers indicating different settings. In equations (3-30), (3-31), and (3-32), P is chosen to be 2 and 4, and K is chosen to be 2 (as representative values).

The computational cost of Lamboni method for total effect indices is given by equation (3-36), where N is the sampling of a Latin Hyper Cube and Q is the order of quadrature. In the plot, N is assumed to be 250 and Q is assumed to be 2, indicated by the green line. The cost of the grid quadrature method used by (Kucherenko, Klymenko and Shah, 2017) is not compared here, because the quantities being calculated are not equivalent. However a higher cost can be expected as the cost of grid quadrature grows exponentially.

The cost of FAST is given by (3-37), where M is the order of interference (the value is set as 4) and ω_{max} is the maximum frequency in the Fourier Transformation, the set of frequencies are adopted from (McRae, Tilden and Seinfeld, 1982). In Figure 3-5, the

cost is indicated by the purple line. The cost of extended-FAST (which is not plotted here) will be higher due to re-samplings.

The Bayesian approach is not included in the table, as no explicit estimation of the computational cost has been provided in the original paper (Oakley and O'Hagan, 2004). However an example was given, in which a 40-dimensional problem was solved, using 101 model evaluations.

Table 3-1. Computational costs of various methods

Method		**Computational Cost**	
MCS for First Order Indices Only (Saltelli, 2002)		$N_{total} = N(n+1)$	**(3-26)**
MCS for First Order and Total Effect Indices (Saltelli, 2002)		$N_{total} = N(n+2)$	**(3-27)**
MCS for First Order, Second Order, and Total Effect Indices (Saltelli, 2002)		$N_{total} = N(n+2+\binom{n}{2})$	**(3-28)**
MCS using Random Balance Design (Mara and Rakoto Joseph, 2008)		$N_{total} = 2N$	**(3-29)**
Intrusive PCE for All the Indices (Xiu and Karniadakis, 2002)		$N_{total} = \frac{(n+P)!}{n!\,P!}$	**(3-30)**
Non-Intrusive PCE (using full tensor quadrature) for All the Indices (Eldred and Burkardt, 2009)		$N_{total} = (P+1)^n$	**(3-31)**
Non-Intrusive PCE (using sparse grid) for All the Indices (Xiu and Hesthaven, 2005)		$N_{total} \sim 2^K n^K / K!$	**(3-32)**
Non-Intrusive PCE (using least square approximation) for All the Indices	Adopted in (Cuneo, Traverso and Shahpar, 2017)	$N_{total} = \frac{(n+P)!}{n!\,P!}$	**(3-33)**
	Recommended in (Eldred and Burkardt, 2009)	$N_{total} = 2\frac{(n+P)!}{n!\,P!}$	**(3-34)**
	Reported in (Sudret, 2008)	$N_{total} = (n-1)\frac{(n+P)!}{n!\,P!}$	**(3-35)**
Lamboni Method (Lamboni, 2016)		$N_{total} = N(nQ+1)$	**(3-36)**
Classic FAST (McRae, Tilden and Seinfeld, 1982)		$N_{total} = 2M\omega_{max} + 1$	**(3-37)**

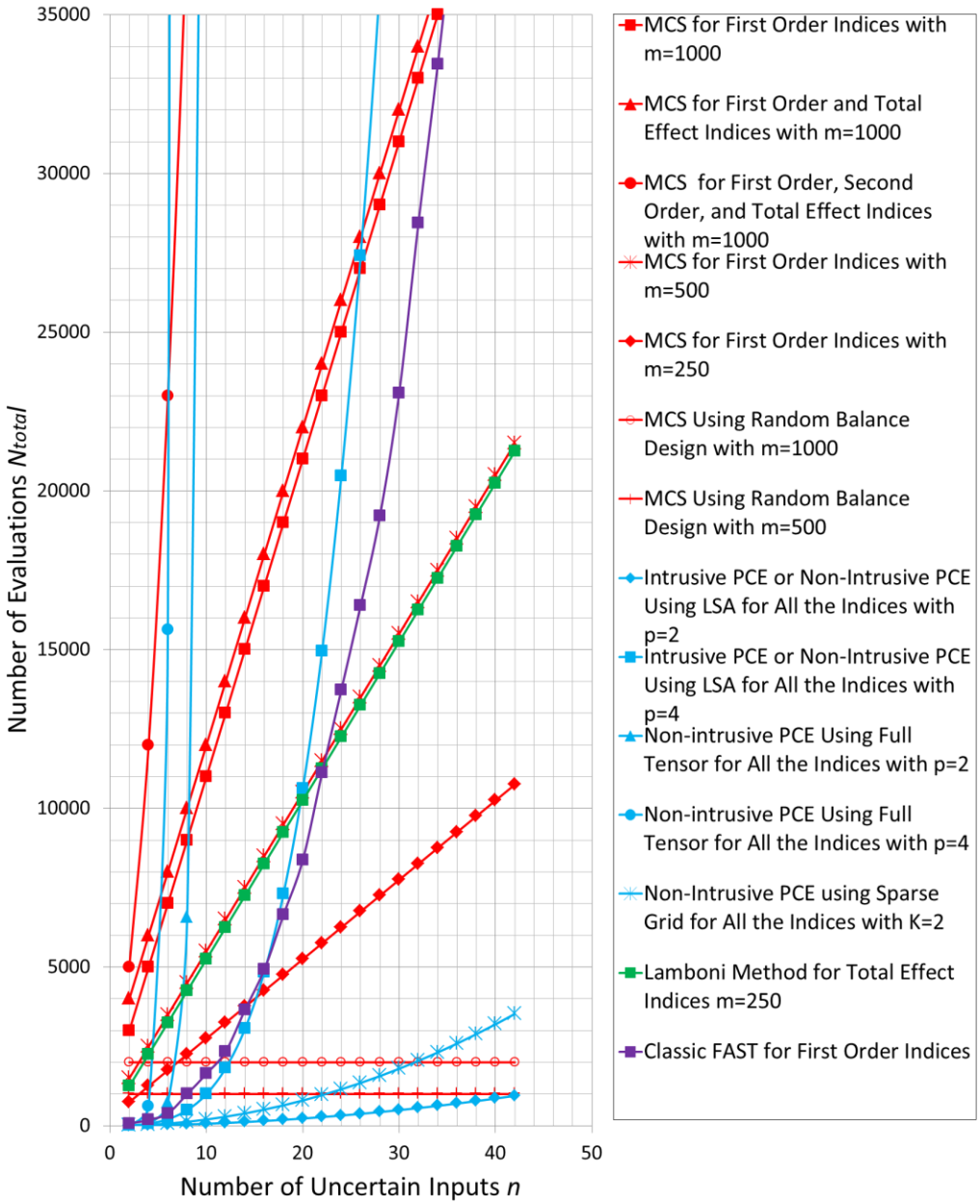

Figure 3-5. Number of evaluations for different SA techniques

It can be seen that, for most of cases (except the random balanced methods (Mara and Rakoto Joseph, 2008; Tarantola, Gatelli and Mara, 2006)), the computational cost is related with the number of uncertainty sources, therefore becomes very expensive for high dimensional problems. Thus improving efficiency (i.e. the calculation speed), is still an area requiring further research.

3.5 Summary and Conclusions

This chapter has reviewed the existing methods regarding the three specific problem areas of interest, namely: margin allocation and trade-off, uncertainty allocation, and sensitivity analysis. According to the literature review, several limitations can be identified for each of these problem areas:

Regarding margin allocation and trade-off, existing research has already addressed the explicit link between margin and uncertainty, based on the probabilistic formulation of the former. However there is a lack of systematic approach to perform trade-off studies, especially considering the interaction between different margins.

Regarding uncertainty allocation, similar problems can be identified among several different research communities. Although a few methods have been proposed, the implementation of these methods are complicated, which has restricted their application for interactive design exploration.

Regarding sensitivity analysis, the main challenge is related to the computational cost for large scale problems. With most of the current techniques, the required number of model evaluations will become prohibitive, as the dimension increases. Although some improved techniques are proposed, there are still different ways to address the problem, which have not been investigated.

In response to these issues, methods are developed as presented in the following Chapters.

4 Margin Allocation and Trade-offs

4.1 Introduction

This chapter presents a method for margin allocation and trade-off, within the context of UQ&M. The pros and cons of margins have already been discussed in the motivations (Section 1.2) and literature review (Section 3.2.1). According to the literature review, some of the problems have been addressed, especially regarding the probabilistic formulation of margins. The remaining challenges include:

- **Account for different types of uncertainties**: In the previous research (Cooke et al., 2015; Guenov, Nunez and Gondhalekar, 2011; Thunnissen, 2004, 2005; Thunnissen and Tsuyuki, 2004; Zang et al., 2015), margins were mainly introduced for uncertainties associated with computational models. In practice, margins can also be applied to account for uncertain requirements, due to potential design modifications and future developments.
- **Explore the interactions between different margins**: When designing a complex system, different margins can be competing and even conflicting with each other. For instance, given a constraint on the aircraft gross weight, increasing the margin on the weight of one subsystem will restrict the potential margins on the weights of others. In a more general case, the margins can be allocated across different disciplines. A method should be developed to explicitly explore this interaction, which has not been addressed in the previous research.
- **Enable the trade-off between margins and other quantities**: The previous research (Cooke et al., 2015; Guenov, Nunez and Gondhalekar, 2011; Thunnissen, 2004, 2005; Thunnissen and Tsuyuki, 2004; Zang et al., 2015) has addressed the trade-off between a margin and a probability of constraint satisfaction. In (Cooke et al., 2015), the trade-off between margins and performances are implicitly addressed by using a down selection process. In a broader picture, this should involve decision making upon margins, system performances, and probabilities of constraint satisfaction. A formalised approach is therefore required to organise such trade-off studies.

To address these problems, the proposed method adopts a set-based design approach (as explained in Section 4.2), along with techniques for uncertainty quantification and sensitivity analysis. The aforementioned interactions and trade-offs are classified as three categories: margin vs. margin, margin vs. probability of constraint satisfaction, and margin vs. performance.

The rest of the chapter is organized as follows. Section 4.2 presents specific concepts, tools, and enablers which are adopted in the proposed method. The latter is described in Section 4.3. The demonstration of the proposed approach, implemented in AirCADia is presented in Section 4.4. Finally, conclusions are drawn and future work is outlined in Section 4.5. It should be mentioned that, the development is a collaborative work of the entire research group (Guenov et al., 2017). Some initial ideas are proposed by the author's supervisors and colleagues, while the formulation and implementation are conducted as the author's work.

4.2 Background

The philosophy of the set-based design approach is to keep the design open "by the parallel development of multiple design solutions and delaying the critical decisions" (Riaz, Guenov and Molina-Cristóbal, 2017). Some early application and general introduction on set-base design can be found in (Singer, Doerry and Buckley, 2009; Sobek, Ward and Liker, 1999). More recent development of set-based approach for aircraft family design is given by (Riaz, 2015; Riaz, Guenov and Molina-Cristóbal, 2017), from which the following concepts and enablers are adopted in this work.

4.2.1 Design and Performance Space

A design space (DS) is defined as the hypercube confined by the specified (valid) intervals of all design variables. The design variables are assumed to be independent. If a correlation exists, there will be either a constraint, or a model (in which the dependent variable will not be considered as a design variable) to address the relationship.

For continues design variables, the specified (valid) interval of the i_{th} design variable is given as follows:

$$DV_i = [dv_i^{LB}, dv_i^{UB}], i = 1,2,3, \dots, u \quad \textbf{(4-1)}$$

Here dv_i^{LB} and dv_i^{UB} are the lower and upper bounds, respectively. Mathematically, the DS can be represented by equation (4-2), as the Cartesian product of all the design variable sets ($DV_i, i = 1,2,3, \dots, u$). Each element in the DS is a possible design solution, i.e., an ordered u-tuple $[dv_1, dv_2, \dots, dv_u]$.

$$DS = DV_1 \times DV_2 \times \dots \times DV_u \quad \textbf{(4-2)}$$

In practice, Design of Experiments (DoE) (Antony, 2003) can be applied to discretise and populate the design space. For this purpose, various sampling strategies can be used, for instance, full factorial, Monte Carlo, Latin hypercube, etc. In this research, full factorial approach is used for its simplicity. If the design variables are discrete in nature, the aforementioned process can still be applied. For instance, the interval of number of passengers is $[160, 180]$, and DoE can be conducted with 160, 165, 170, 175, and 180 passengers. However these cases are not demonstrated in this research.

For each sample, the computational workflow is executed to obtain the relevant performances. Similar to the concept of design space, a performance space (PS) can be defined as a hypercube confined by the intervals of all performance variables. The ranges however are not specified by the designer but rather defined by the maximum and minimum values of the corresponding variables, obtained from the DoE study:

$$PV_i = [y_i^{MIN}, y_i^{MAX}], i = 1,2,3, \dots, l \quad \textbf{(4-3)}$$

By using Cartesian product, the performance space is defined as:

$$PS = PV_1 \times PV_2 \times \dots \times PV_l \quad \textbf{(4-4)}$$

4.2.2 Parallel Coordinates Plot

The Parallel Coordinates Plot (PCP) is used to visualise high-dimensional spaces (Inselberg, 2009). One illustrative example is shown in Figure 4-1 (plotted in AirCADia), in which, every vertical axis represents one dimension of the design or performance space. The values of each variable are indicated as the vertices along the corresponding axes. One design solution could be represented as a polyline which connects the vertex on each axis, as indicated by the highlighted orange line.

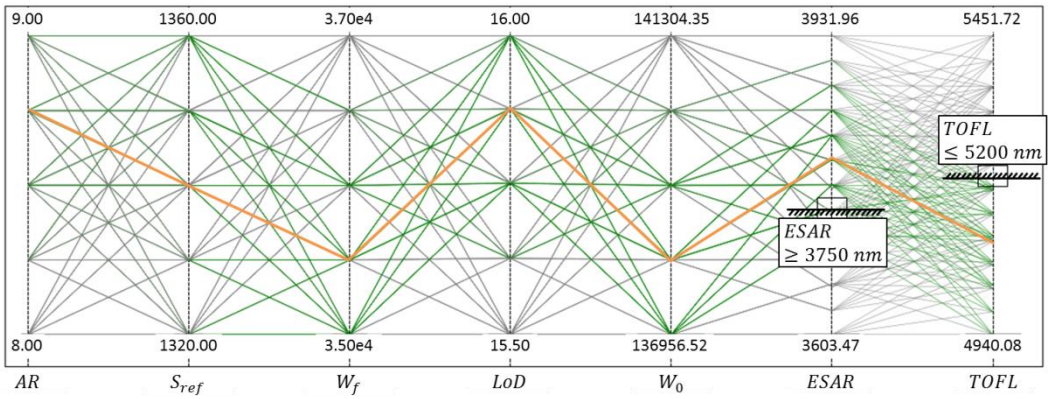

Figure 4-1. Example of parallel coordinates plot

In AirCADia, constraints can be specified dynamically for each variable. For instance in Figure 4-1, a lower bound and an upper bound are placed on the axes of $ESAR$ and $TOFL$, respectively. According to these constraints, the feasible solutions are marked as green polylines, while the infeasible ones indicated by grey. Further information of PCP for set-based design could be found in (Guenov et al., 2014a, 2014b; Nunez et al., 2009; Riaz, 2015; Riaz, Guenov and Molina-Cristóbal, 2017).

4.2.3 Iso-Contours Plot

The Iso-Contours Plot (ICP) divides a high-dimensional design space into multiple two-dimensional (2D) contours, on which the projections of performance constraints are plotted, based on numerical calculation (Guenov et al., 2014a; Riaz, 2015; Riaz, Guenov and Molina-Cristóbal, 2017). An illustrative example is shown in Figure 4-2, in which, the three-dimensional design space composed of AR, W_f, and S_{ref} (on the left) is projected as the 2D contours of AR and W_f (on the right). Within a 2D contour plot, the feasible regions are indicated as white areas, while the infeasible ones are shaded by grey. By changing the value of S_{ref}, the positions of constraints (on $ESAR$ and $TOFL$ in this case) will be changed accordingly, as indicated by the two different 2D contour plots on the right. In a more general case, the designer can arbitrarily select any two dimensions to plot multiple contours. It should be mentioned that, the method does not require any additional model evaluations, apart from the samples of the DoE study. The reader is referred to (Riaz, 2015; Riaz, Guenov and Molina-Cristóbal, 2017) for more information on this technique.

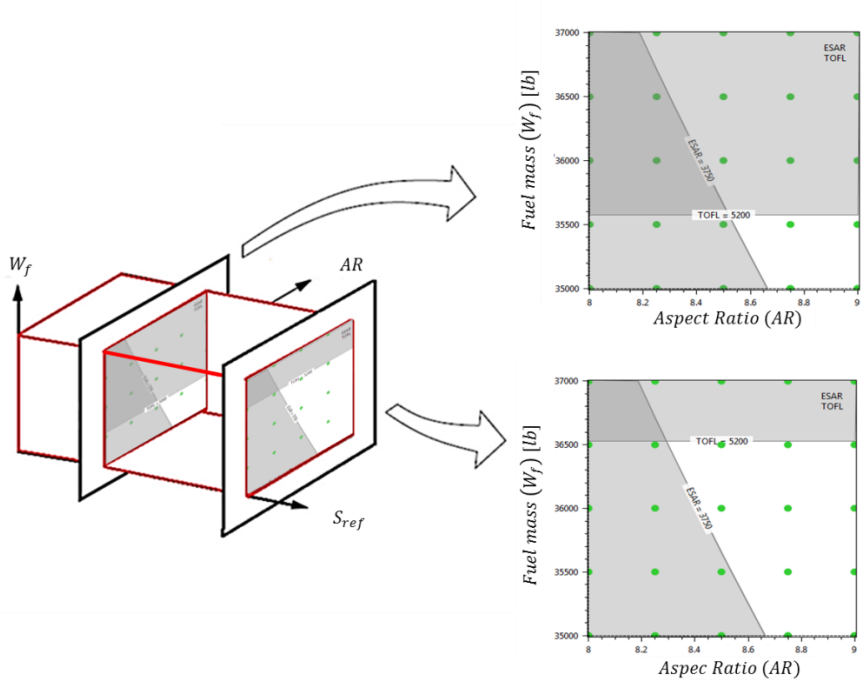

Figure 4-2. Design space exploration using 2D contour plots (adapted from (Riaz, 2015))

4.3 Methodology

The overview of the proposed method is illustrated by Figure 4-3. After design problem formulation, the process breaks into two streams. The left stream includes analyses conducted by an uncertainty expert, and the results are used to support the decision making in the right stream (by a designer). The final outcome of this method is a set of promising margin combinations.

This separation of the two streams is to make the method less intrusive to be incorporated in the traditional deterministic design process, which has already been widely adopted in industrial practice. For instance, the designer can focus on his/her domains of expertise (e.g. airframe or systems) without being heavily involved in the UQ&M techniques. On the other hand, the uncertainty expert is not required to fully understand the domain knowledge of aircraft design. It should be noted that, this distinction between the designer and uncertainty expert is not absolute, as an individual can have multiple capabilities and roles in the design study.

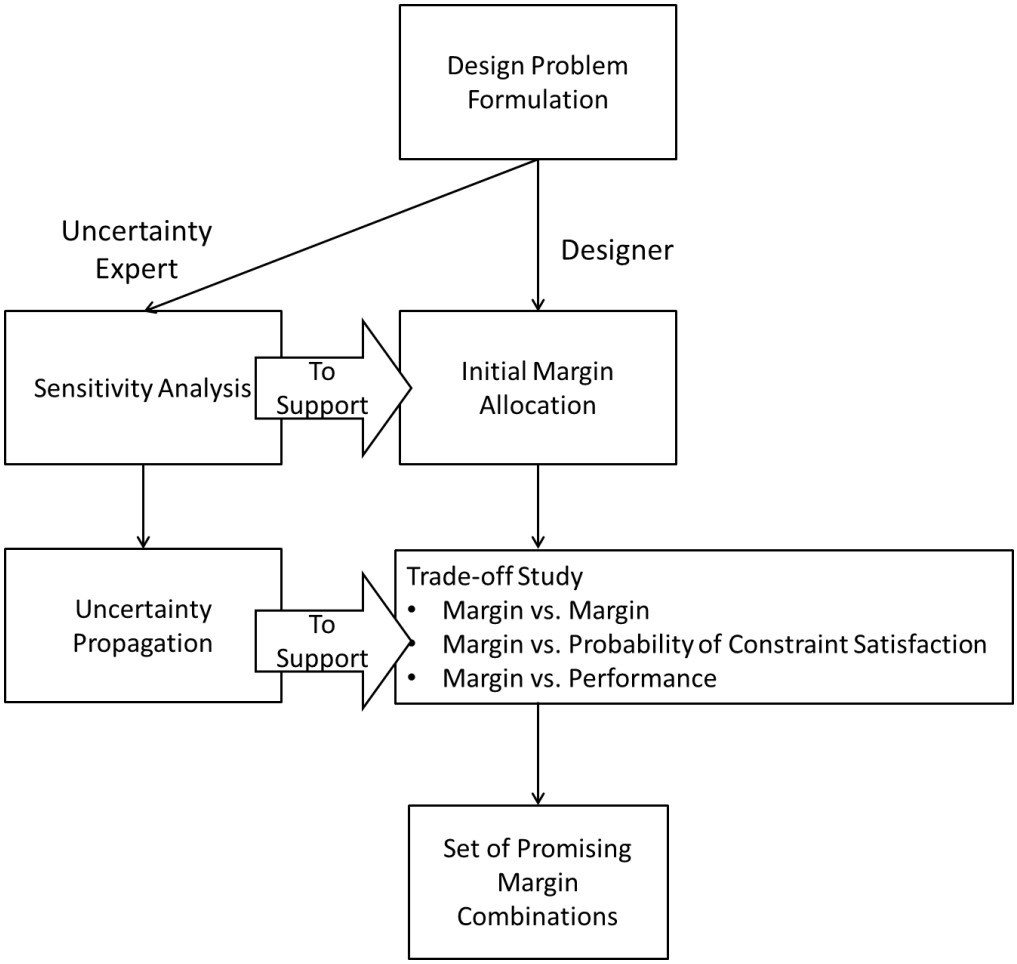

Figure 4-3. Overview of margin allocation and trade-off

In the following sections, a running example is used to explain details of the proposed method, using the SIMPCODE workflow as presented in Section 2.2. The deterministic setup is listed in Table 4-1.

Table 4-1. DoE setup using the illustrative example

	Variable	**Notation**	**Unit**	**Value (Range)**
Design Variable	Aspect Ratio	AR		[8,9]
	Fuel Weight	W_f	lb	$[35000, 37000]$
Parameter	Payload Weight	W_p	lb	28000
Performance	Equivalent Still Air Range	$ESAR$	$n.m$	$ESAR \geq 3500$
	Take-off Field Length	$TOFL$	ft	To minimise

4.3.1 Sensitivity Analysis

The first step is to identify all the potential margins in the design problem. In this research, margins can be allocated to account for two types of uncertainty:

- **Requirement Uncertainty:** is due to vagueness (Ayyub and Klir, 2006), design modifications, or potential needs for future product evolution. Considering margins on requirements offers the flexibility to handle such scenarios. For instance, in the running example, a margin (Mar_{W_p}) can be applied on required payload mass (W_p) to account for the potentially increase during the design process (e.g. catering to a changing market need). This margin can also enable the future development of a derivative aircraft with more passenger capability (van Heerden et al., 2016; Riaz, Guenov and Molina-Cristóbal, 2017).
- **Model Uncertainty:** is due to lack of knowledge as explained in Section 2.3. A margin can be applied as a knock-down factor on a model output to make the predictions more conservative. For example, margins (Mar_{LoD} and Mar_{W_e}) can be applied on the predicted lift-over-drag (LoD) and empty weight (W_e), respectively. Because the corresponding models in the running example are based on simplified empirical relationships, the results can be over- or under-estimated. By applying the margins, risk of rework can be mitigated if the true values turn out to be on the pessimistic side.

If the total number of potential margins is too large, sensitivity analysis (Section 1.2 and 3.4) can be applied to select a subset of the most critical margins. For instance, all the margins are assumed to follow uniform distributions within +/- 10% intervals around their nominal values. In this case, the mathematical meaning of the resulting sensitivity indices is not strictly the same as defined in Section 3.4, because the margins are not necessarily random variables and the distributions are only assumed. However, these results are used to down scale the problem, so that the following analysis can be computationally affordable.

4.3.2 Initial Margin Allocation

According to results of sensitivity analysis, margins are allocated on the most critical variables. Similar to the design variables, all the margins are assumed to be independent

and they interactions will be explored as trade-offs as shown in the following sections. The initial values of margins can be based on equation (3-8), if the probability distributions are already known. In a more likely scenario, the probability distributions are not yet available or even not attainable (for instance, a possible growth in the required payload). In such a case, the margins will be arbitrarily specified based on the designer's judgement. For example, margins are allocated on the predicted Lift-over-Drag Ratio ($Mar_{LoD} = 5\%$), Empty Weight ($Mar_{W_e} = 5\%$), and requirement of Payload Weight ($Mar_{W_p} = 10\%$).

The effect of margins on the computational workflow is illustrated in Figure 4-4, where the margins and additional calculations (compared with the original workflow in Figure 2-1) are indicated by yellow ovals and dashed-line blocks, respectively.

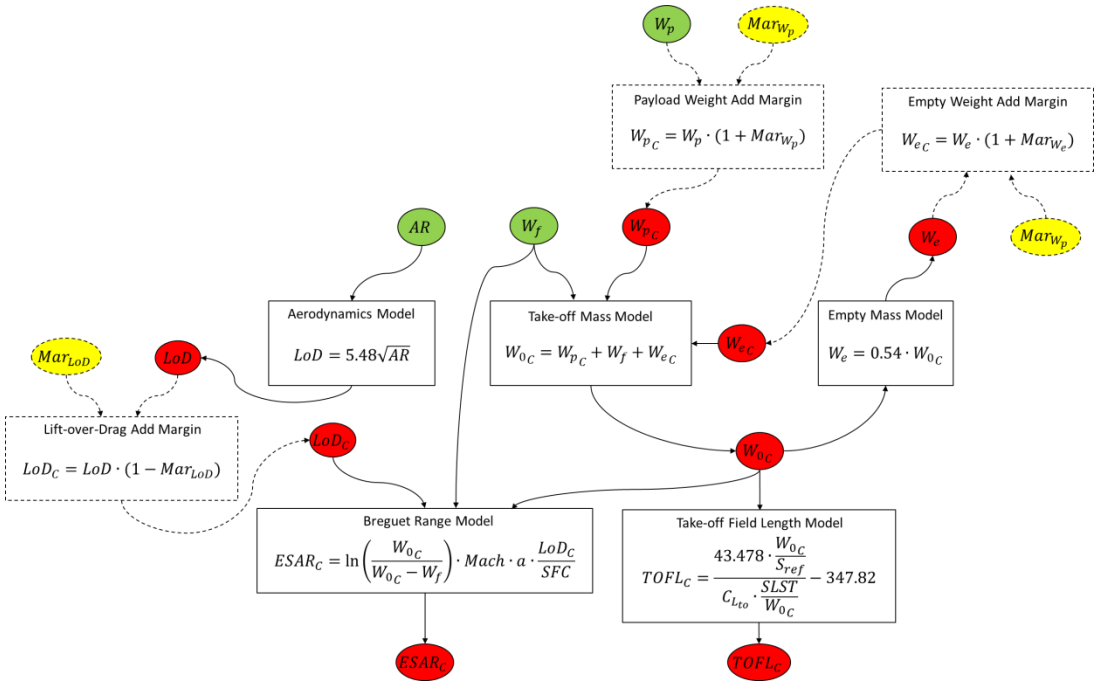

Figure 4-4. Workflow modified to include margins

Since the initial margins might not be feasible, a set-based approach is adopted to evaluate multiple solutions. Here the concept of Margin Space (MS) is introduced, as a hypercube similar to the design and performance spaces. For each margin of interest, an interval is defined, which should contain the initially specified values.

$$MAR_i = [Mar_i^{LB}, Mar_i^{UB}], \quad i = 1,2,3,\dots,w \tag{4-5}$$

The upper bounds Mar_i^{LB} and lower bounds Mar_i^{UB} are based on the maximum and minimum values of interest for each margin. The MS is generated by using the Cartesian product of all the margin intervals:

$$MS = MAR_1 \times MAR_2 \times \ldots \times MAR_w \quad \textbf{(4-6)}$$

Each element in the margin space is an ordered u-tuple $[Mar_1, Mar_2, \ldots, Mar_w]$, which will be referred to as a 'margin combination'.

A DoE study, as mentioned in Section 4.2.1 is used to populate both the design space and margin space. The step size is decided by the designer. To maintain the computational cost within an affordable level, a large step size can be used for initial assessment to explore the general trends of the design and margin spaces. Reduced step sizes can be applied later, on specific areas of interest. In this case, 5 samples are placed for each design variable and 6 samples are placed for each margin. This leads to 5400 design solutions in total, as shown in Figure 4-5.

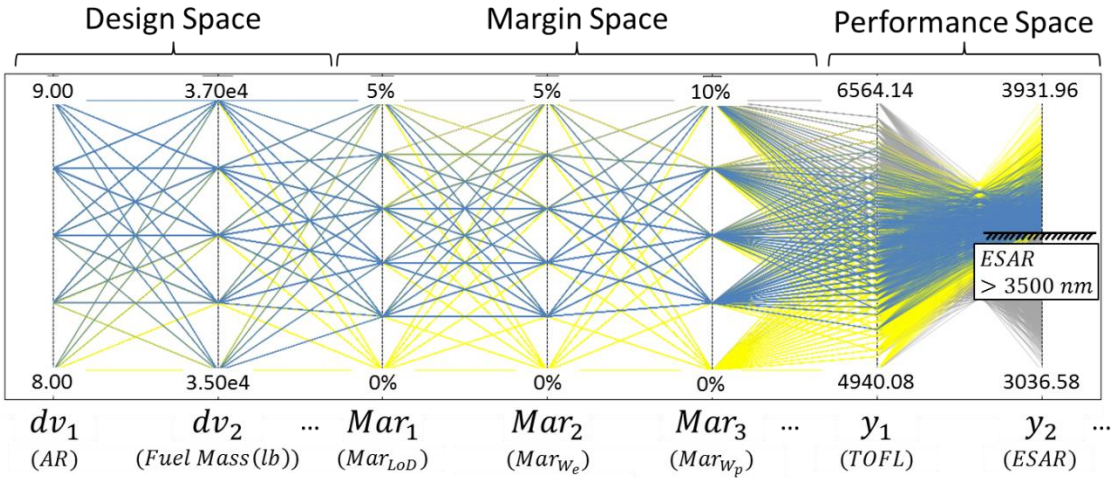

Figure 4-5. Result of initial margin allocation (to be updated please ignore the yellow lines in the plot)

Some of the margin combinations may not be feasible because of the resulting performance not meeting the constraints (as indicated by grey in the plot). On the other hand, the yellow polylines indicate solutions which are feasible, but with no margins on the uncertain variables.

The maximum value of a margin can be defined as a function of the design variable and other margins:

$$Mar_i \leq F_{Mar_i,C_r}(\boldsymbol{dv}, \boldsymbol{Mar}_{\sim i}), i = 1,2,3, \dots, w, r = 1,2,3, \dots, t \quad \textbf{(4-7)}$$

Here, F_{mar_i,C_r} is a function which returns the upper bound of the i_{th} margin, with regard to the r_{th} constraint. $\boldsymbol{dv}$ is the vector of values of design variables, and $\boldsymbol{Mar}_{\sim i}$ is the vector of all except the i_{th} margin, i.e. $\boldsymbol{Mar}_{\sim i} = (Mar_1,\ Mar_2,\ Mar_3,\ \dots, Mar_{i-1},\ Mar_{i+1},\ \dots,\ Mar_w)$.

The iso-contours plot, as explained in Section 4.2.3, can be used to visualise the feasible margin space for each design point (DP). This is demonstrated in Figure 4-6, where the design space is shown on the left, with two design variables dv_1-aspect ratio (AR) and dv_2-fuel mass (W_f). For selected design points (DP1 and DP2) in the design space, the corresponding feasible margin spaces are shown as the white regions in Figure 4-6 (a) and (b), respectively. It can be seen that, compared with DP1, DP2 leads to a reduced feasible area, which means that this design point has less room for margin allocation. Note that, for simplicity, only one constraint ($ESAR \geq 3\,500nm$) is applied in Figure 4-6. Introducing iso-contours of other constraints may further constrict the margin spaces. This will be illustrated in the evaluation section.

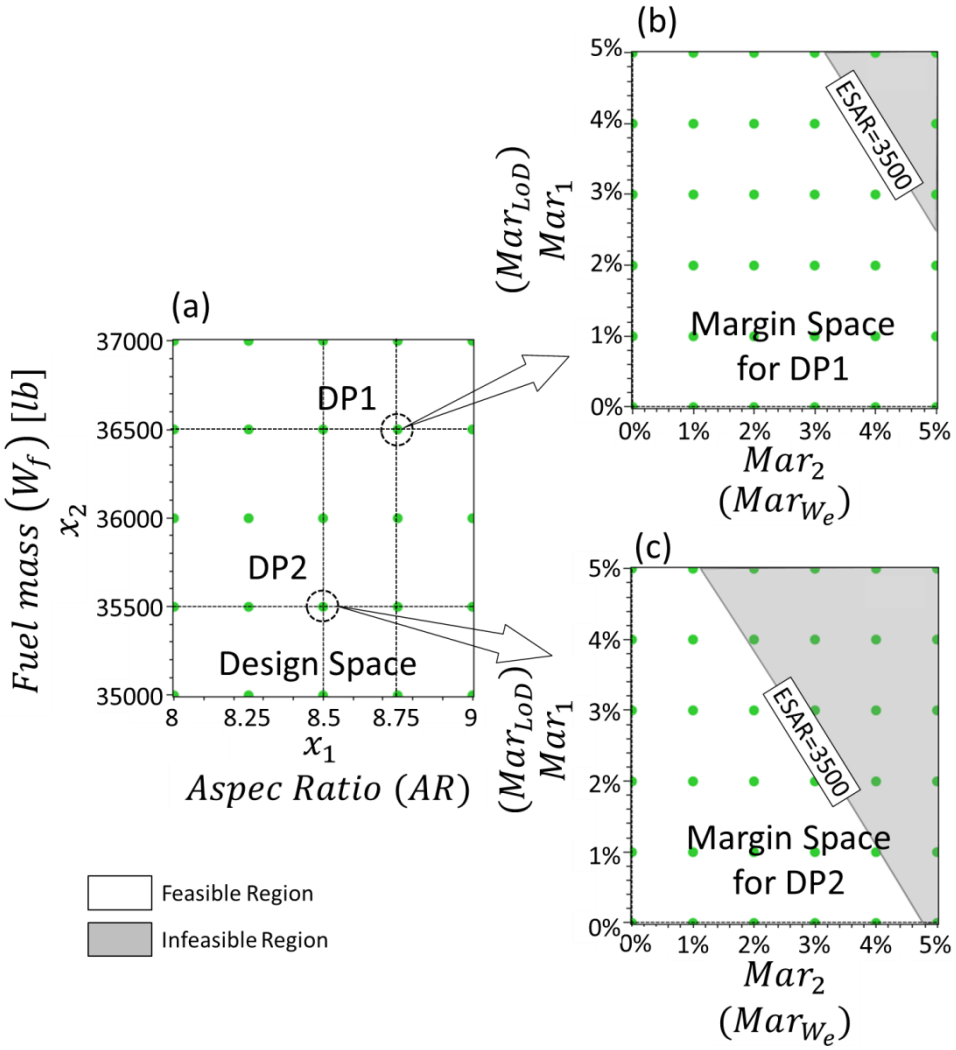

Figure 4-6. Effect of design point selection on the margin space

Similarly, if a margin combination is selected in the margin space, the feasible design space will also be changed accordingly. This is illustrated in Figure 4-7, again using the running example. Figure 4-7 (a) shows the design space with zero margins. Selection of a margin combination, represented by the dotted circle in Figure 4-7 (b), modifies the feasible design space, as shown in Figure 4-7 (c).

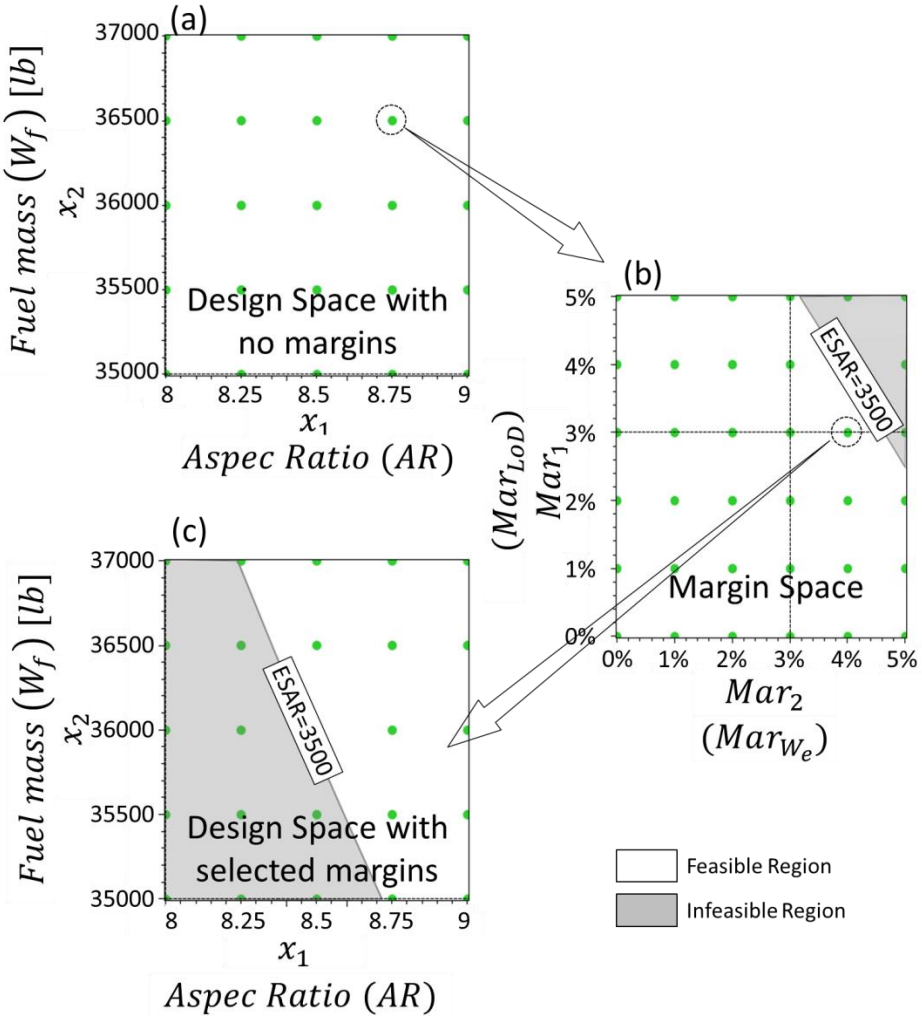

Figure 4-7. Effect of margin selection on the design space

The explicitly-defined and bi-directionally linked margin and design spaces constitute one of the central concepts of this research.

4.3.3 Uncertainty Propagation

Uncertainty analysis is performed for all the feasible design solutions. This includes forward propagation and probabilistic assessment of constraint satisfaction. The *Univariate Reduced Quadrature* (URQ) method (Padulo, Campobasso and Guenov, 2011), was employed to propagate the uncertainty for each design point (more details of this technique will be explained in Chapter 5).

To account for model uncertainty, the method from (Molina-Cristóbal et al., 2014) is adopted, as explained in Section 2.4. One issue to be addressed is that, for uncertainty

propagation, the margins corresponding to model uncertainty need to be removed first. Otherwise, the uncertainty will be double counted. However, simply removing the margins will lead to another problem, because the links between margins and probabilities of constraint satisfaction will be lost. As illustrated in Figure 4-8, for one design point with $AR = 8$, the deterministic prediction of lift-over-drag will be $LoD = 15.50$. By applying different Mar_{LoD} on the relevant model, the conservative estimation LoD_C will change accordingly. On the other hand, in uncertainty analysis, the same probability of constraint satisfaction will be obtained, if the margin is removed and the random factor rv_{LoD} is used. This gives a wrong impression that the margins are not related to the probability of constraint satisfaction at all.

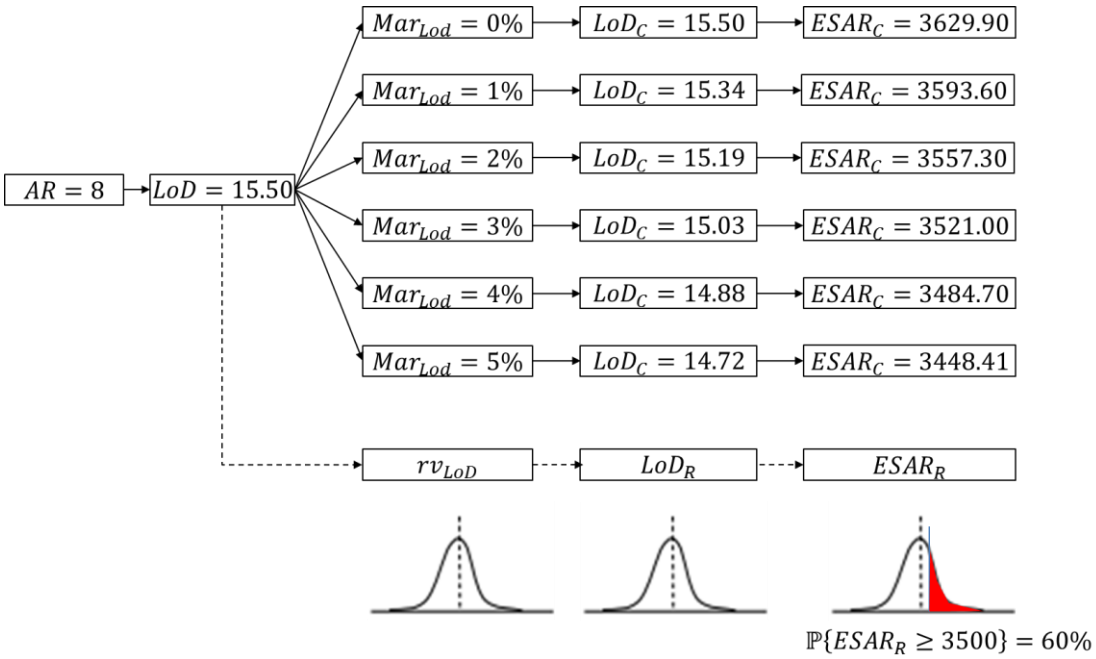

Figure 4-8. Link between margin and probability

In fact, for a fixed design point, its probability of constraint satisfaction cannot be improved by adding margins. The real effect of margins is: they tend to shift the projection of deterministic performance constraints towards more robust regions in the design space. This seems to be obvious if the margin is allocated directly on the constraints, but turns out to be implicit for intermediate variables of the workflow. As illustrated in Figure 4-9, by increasing the margins on LoD and W_e from 2% to 3%, the projection of constraint ($ESAR \geq 3500$) has been moved towards the top right corner. As a result, the design points which are feasible with small margins become infeasible

as the margins are increased (for instance the point A, as marked by the dashed circle). Although in this case the margins cannot improve the robustness of point A, they do help to filter out such points, which are likely to have lower probabilities of constraint satisfaction.

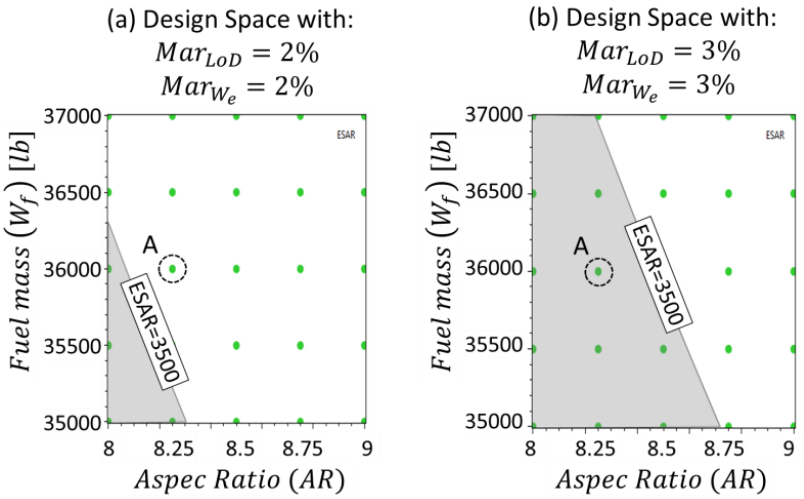

Figure 4-9. Effect of margins on the projection of deterministic constraint

Following this reasoning, one approach is adopted from (Cooke et al., 2015; Zang et al., 2015), in which the uncertainty propagation is implemented using two sets of models as shown in Figure 4-10. On the left stream, are the deterministic models, where margins are added; on the right stream are the randomized models, for which the outputs are distributions.

In this way, the links between the conservative estimations (due to margins) and the probability of constraint satisfaction (due to randomization) can be 'preserved'. The results are shown in Figure 4-11, in which three extra axes (μ_{TOFL_R}, μ_{ESAR_R}, and $\mathbb{P}_{ESAR_R}$) are added on the right, which are respectively the means of stochastic $TOFL$ and $ESAR$, and the probability of $ESAR \geq 3500$. The benefit of plotting both the margins and probability is that, the designer can put a lower bound on the probability, to identify which margin combinations are able to produce robust solutions. If this link is lost, such a down-selection will not be able to perform.

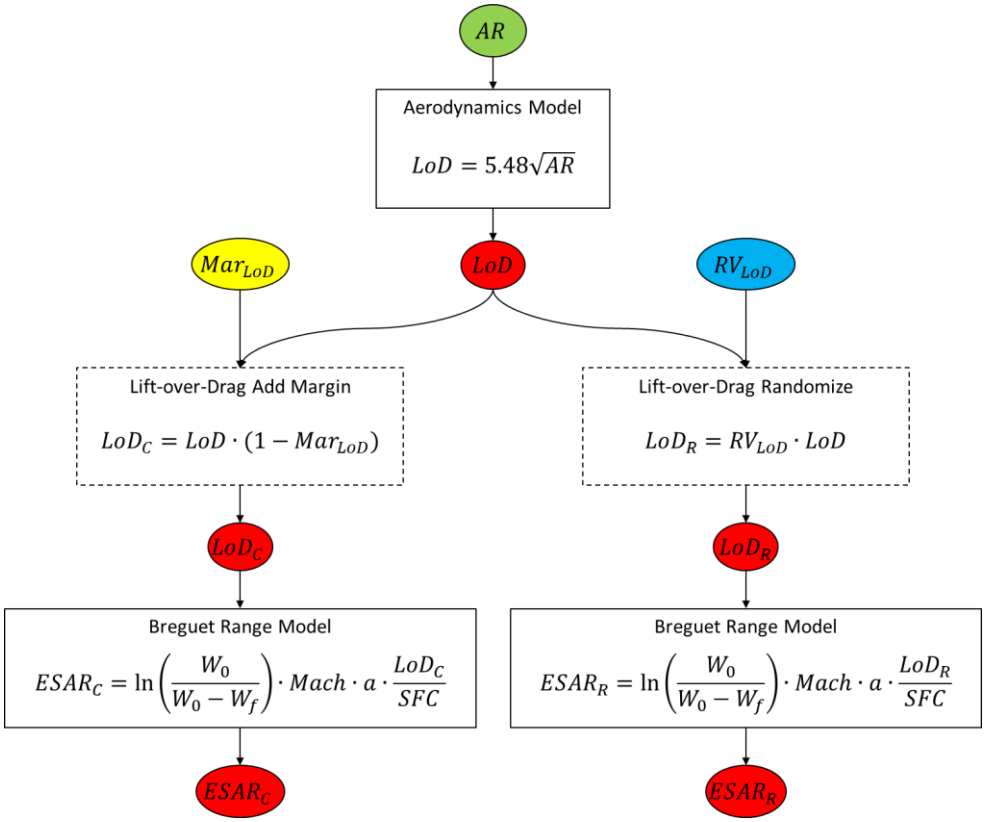

Figure 4-10. Two set of models for implementing the uncertainty analysis

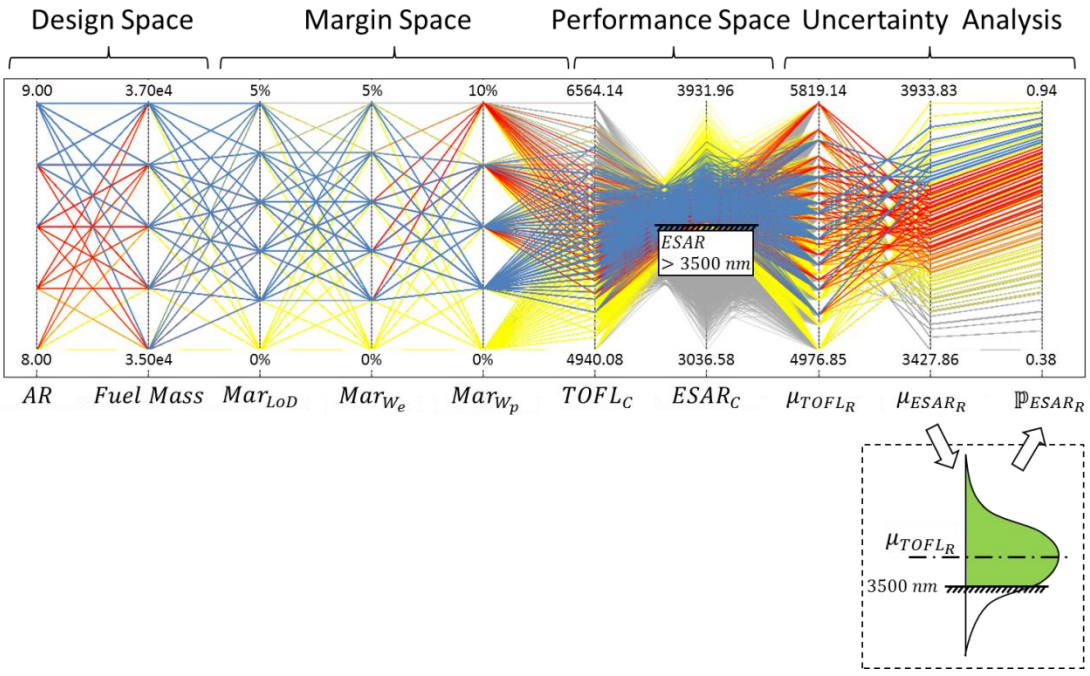

Figure 4-11. Combination of conservative estimations and result from uncertainty analysis

4.3.4 Trade-off Study

In the trade-off study, the designer explores the design/margin space and identifies one or more margin combinations that may provide reserves for uncertainty in models and requirements, while attempting to prevent degradation in design performance. The previous steps can be considered as a process to gather information, which is used for three types of trade-off studies presented in this section: between margins, between margins and performances, and between margins and probabilities of constraint satisfaction. It is important to note that in practice the designer can conduct these trade-offs in any order, subject to available information.

4.3.4.1 Margin vs. Margin

Due to the presence of performance constraints, margins on different variables and models can be conflicting with each other. Therefore trade-off studies are conducted between margins to find acceptable combinations which fulfil all the constraints.

This exploration is enabled by a matrix of two-dimensional Iso-contour plots of the margin space. This is demonstrated in Figure 4-12, using the running example. Given a selected design point, for instance DP1 in Figure 4-6, the projection of performance constraint ($ESAR \geq 3500nm$) is plotted in two-dimensional contours of Mar_{LoD} vs. Mar_{W_e}, Mar_{LoD} vs. Mar_{W_p}, and Mar_{W_e} vs. Mar_{W_p}. By selecting a margin combination point (MP) in one of the plots (for instance MP1: $Mar_{LoD} = 1\%, Mar_{W_e} = 3\%$), the feasible region in the remaining two plots will be updated.

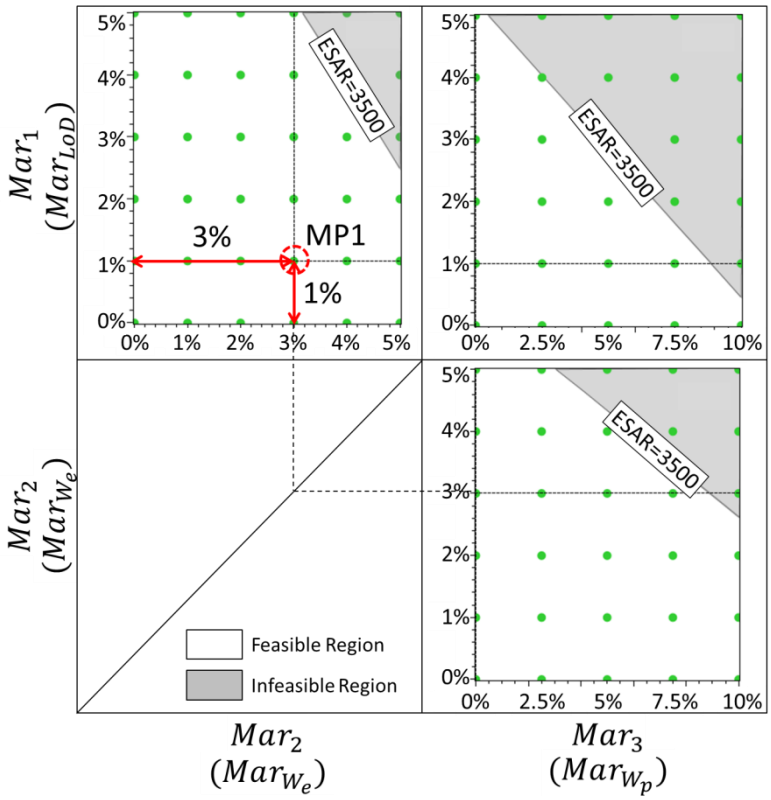

Figure 4-12. Selection of MP1 in margin space ($Mar_{W_e} = 3\%$ and $Mar_{LoD} = 1\%$).

If the margin combination point, MP2 ($Mar_{W_e} = 1\%, Mar_{LoD} = 4\%$), is subsequently selected, the iso-contour representing the aircraft range constraint shifts in the other contour plots of the MS, as shown in Figure 4-13. It can be seen from Figure 4-13 that the selection of MP2 results in a larger feasible margin space in plot Mar_{LoD} vs. Mar_{W_e}. However, the opposite trend is noticeable in the bottom-right plot (Mar_{W_e} vs. Mar_{W_p}).

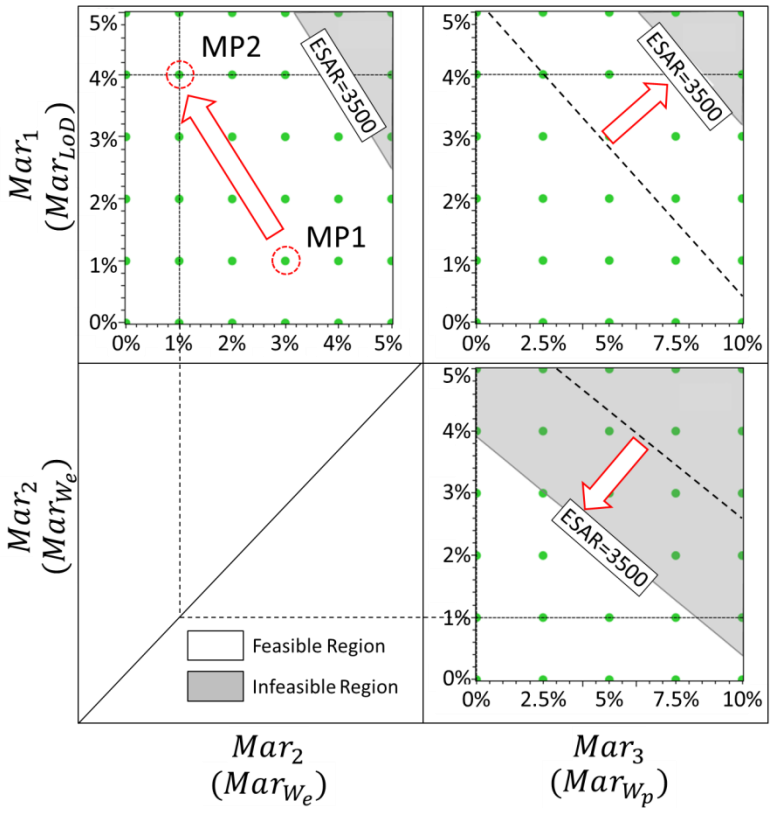

Figure 4-13. Selection of MP2 in margin space ($Mar_{W_e} = 1\%$ and $Mar_{LoD} = 4\%$).

By selecting different margin combinations in this manner, the designer can interactively visualize how the feasible margin space changes. Note that, the sequence of assigning combinations of margins is up to the designer and can be dependent on the problem under consideration.

After exploration of the relationships between margins (and design variables), parallel coordinates plot (PCP) is utilized to down select design solutions. The designer may filter out solutions with insufficient margins, by placing a lower bound on each margin axis. The remaining solutions constitute a deterministic design band (DB^D) in the parallel coordinates plot. The DB^D for the running example is shown in Figure 4-14. It is confined by the dashed blue lines. In addition to the design variables (AR and W_f) and performance parameters ($TOFL$ and $ESAR$), the plot in Figure 4-14 also includes the margin parameters (Mar_{LoD}, Mar_{W_e}, and Mar_{W_p}). Here, two criteria are used for filtering out undesirable points: 1) deterministic constraints ($ESAR \geq 3500nm$) and 2) minimum margin requirements (1%, 1% and 2.5%, for margins Mar_{LoD}, Mar_{W_e}, and

Mar_{W_p}, respectively). The grey solutions are discarded due to the violation of the deterministic constraint, whereas the yellow solutions are filtered out due to the required minimum margin values not being met.

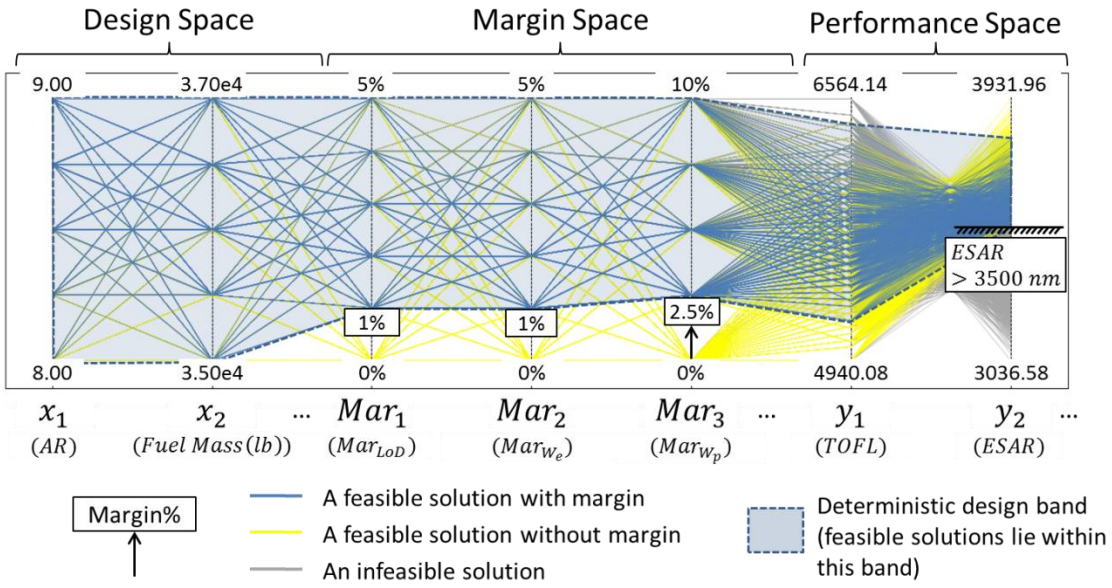

Figure 4-14. Deterministic design band (DB) filtering using parallel coordinates plots.

The seamless integration of parallel coordinates and Iso-contour plots is an effective means for interactive design space exploration and margins trade-off. For example, it may turn out that it is impossible to achieve the desired minimum values for all the margins under consideration, due to not being able to meet the constraints. In such a case, the parallel coordinates plot could be used to find an appropriate combination of minimum margin values by interactively varying the minimum values of the individual margins. Moreover, given that the parallel coordinates and the Iso-contours are interlinked, the designer can simultaneously investigate the positions of the points in the design and margin spaces and visualize the topology of the feasible regions (the white regions in Figure 4-7).

4.3.4.2 Margin vs. Probability of Constraint Satisfaction

The implementation of uncertainty propagation has been explained in Section 4.3.3. Based on those results, the designer can conduct trade-off studies between margins and probabilities of constraint satisfaction.

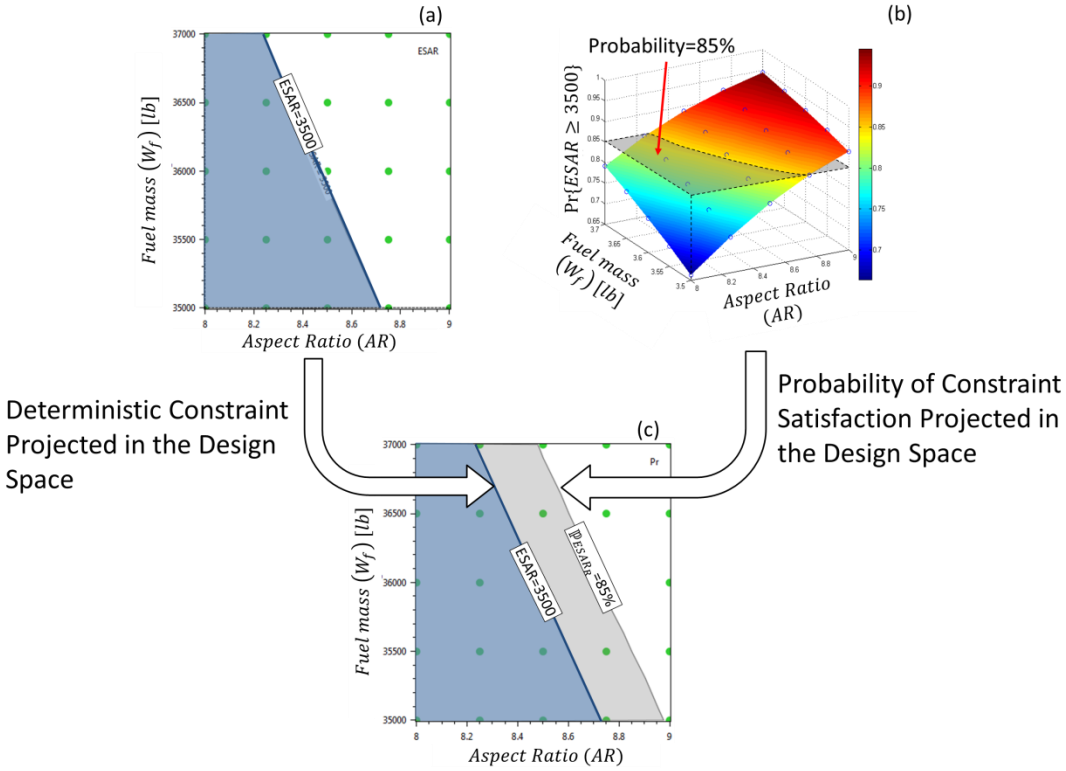

Figure 4-15. A visualization technique to combine a deterministic constraint and the probability of constraint satisfaction

A visualization technique has been developed by combining a deterministic constraint and the corresponding probability of constraint satisfaction, projected in one contour plot of the design space.

As shown in Figure 4-15 (a), the deterministic constraint ($ESAR \geq 3500$) projected in the design space (of AR and W_f), is affected by margins to account for model uncertainty (referring to Figure 4-9). In Figure 4-15 (b), for each design point, the probability of constraint satisfaction, $\mathbb{P}\{ESAR_R \geq 3500\}$, is calculated (by using the right stream in Figure 4-10) and plotted as a 3D surface. The horizontal transparent slide represents a designer specified threshold value for this probability, $\mathbb{P}\{ESAR_R \geq 3500\} \geq 85\%$. The intersection of this slide and the 3D surface is also projected on the

same two-dimensional design space. By combining these two contours as shown by Figure 4-15 (c), the designer is able to visually assess:

- Are the margins allocated on the right variable: The relationship between a margin and a probability of constraint satisfaction is straight forward, if there is only one dimension and the margin is directly allocated on the uncertain variable (as shown in Figure 3-2). But in high dimensional cases, this relationship will be complicated. For instance, the constraint is defined on the $ESAR$, but the margins are allocated on the LoD and W_e. If there are no dependencies between these variables, the margins will not be relevant to the probability we want to increase. From the plot, the two contour lines should be ideally 'parallel' (although the contour lines are not necessarily straight lines), and the feasible regions defined by these two lines should be on the same side. Large intersection angle of the two contour lines would indicate that the margins and uncertainty are not correlated. In the extreme case, the two contour lines are perpendicular to each other, thus altering the margin will not change the probability of constraint satisfaction. This implies that the margins in question should be reallocated elsewhere.
- Are the margins sufficient to account for the relevant uncertainty: this can be assessed by comparing the feasible (or infeasible) region defined by the two contour lines as illustrated in Figure 4-16. Here four possible margin combinations are considered, as listed in Table 4-2. If the feasible region defined by the deterministic constraint is fully contained within the region defined by probability contour line (Figure 4-16 (c)), then the margins are sufficient to accommodate the relevant uncertainty. Otherwise higher values should be applied on the margins (Figure 4-16 (a) and (b)). In another scenario, the designer may realise that too much margin is allocated, if the deterministic constraint defines a much smaller feasible region compared to that defined by the probability threshold (Figure 4-16 (d)).

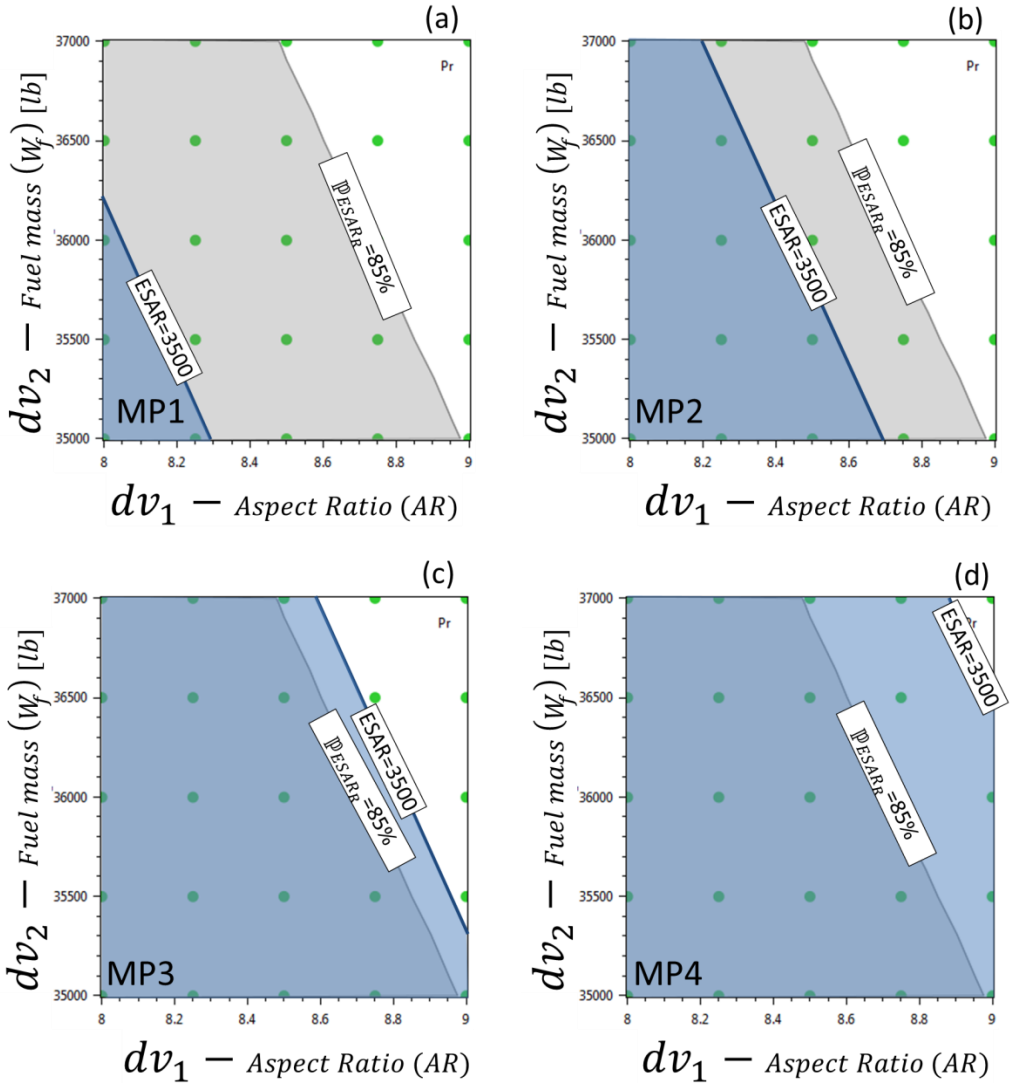

Figure 4-16. Probability of constraint satisfaction vs. margin combinations

Table 4-2. Trade-off between margins combinations

Combination	Mar_{LoD}	Mar_{W_e}
MP1	2%	1%
MP2	3%	2%
MP3	1%	5%
MP4	4%	4%

Similar to the previous step, filtering can be used after the exploration, to discard design solutions that do not meet the required probability of constraint satisfaction. The result of doing this is a robust design band (DB^R). In this running example, the threshold on the probability of satisfying the $ESAR$ constraint, $\mathbb{P}\{ESAR_R \geq 3500\}$, is set as 85%. Design solutions that fail to meet this threshold are shown by the red lines in Figure

4-17. These can be discarded as a further reduction of the polylines in the DB^D as shown in Figure 4-14. The resulting robust design band is shown in Figure 4-17 as the region enclosed by the dashed blue lines.

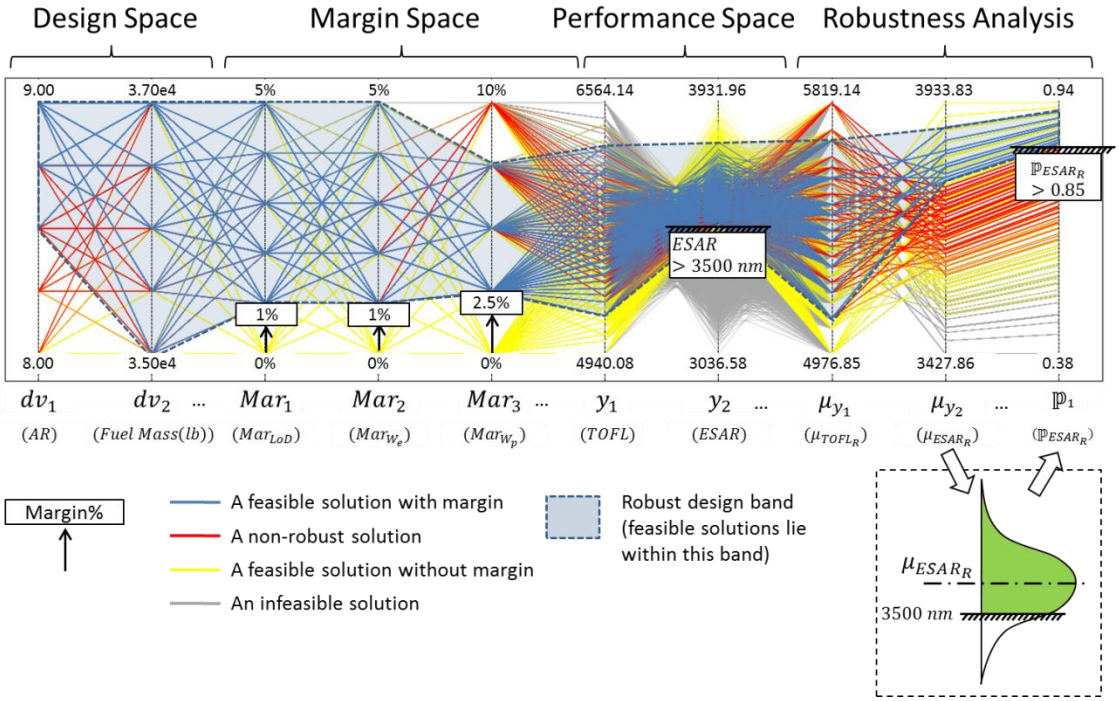

Figure 4-17. Robust design band (DB^R) filtering using parallel coordinates plots

4.3.4.3 Margins vs. Performance

In general, the allocation of margins adversely affects performance, as it involves placing 'reserves'. The purpose of trading-off margins with performance variables is therefore to identify combinations of margins that adequately account for uncertainty, but still allow acceptable performance. A technique using scatter plots is presented for interactively exploring the change in the location of the Pareto front when the margin values are varied.

This is demonstrated in Figure 4-18, where a two-dimensional performance space is plotted. The performances are $TOFL$ (smaller the better) and $ESAR$ (larger the better).

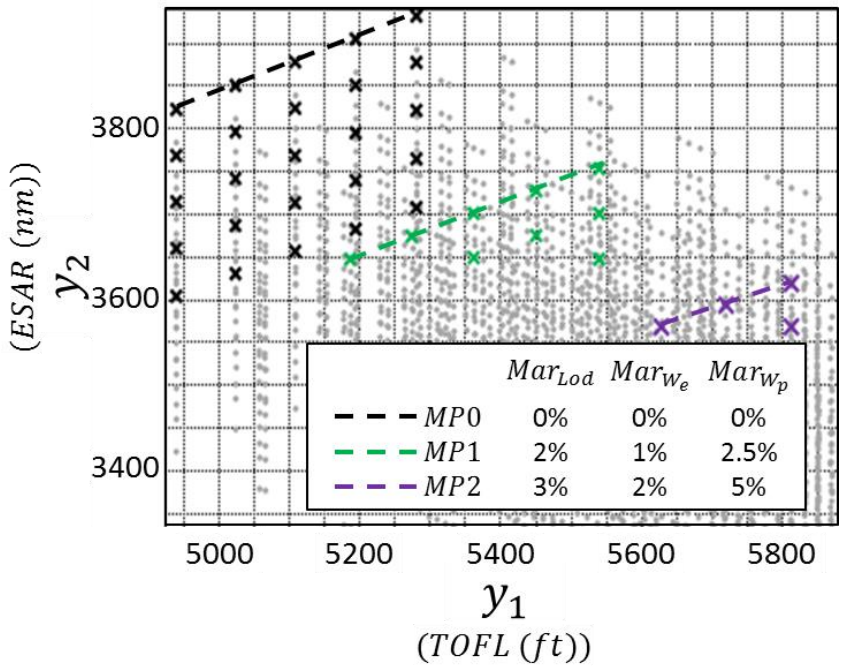

Figure 4-18. Exploration effect of different combinations of margins on the location of the Pareto front

In this plot, the Pareto front of the design solutions with no margins (MP0) is depicted by the black dashed line. The other two dashed lines (green and purple) represent the Pareto fronts corresponding to two different combinations of nonzero margin combinations MP1 and MP2, respectively. As can be seen from Figure 4-18, margin combination MP1 provides better performance compared to margin combination MP2. The designer can therefore use this technique to interactively vary the values of the margins and simultaneously observe the resulting movement of the Pareto front towards degradation or improvement of performance.

4.4 Evaluation

A qualitative evaluation and validation of the proposed method was undertaken as part of the TOICA project. In the evaluation, the proposed method was applied on an aircraft and subsystem sizing test-case, implemented with the in-house software AirCADia (as presented in Section 2.6). The feedback from the practicing experienced designers and airframe systems architects confirmed the potential usability of the method in an industrial context.

4.4.1 Test case description

A conceptual sizing study of a short-range passenger aircraft was considered. A simplified high level view of the computational workflow for this study is shown in Figure 4-19. The margins applied are shown in blue rectangles. For the sake of simplicity, only nine out of the 28 margins were shown in the figure. The full nomenclature of all the margins and design variables are shown in Table 4-3.

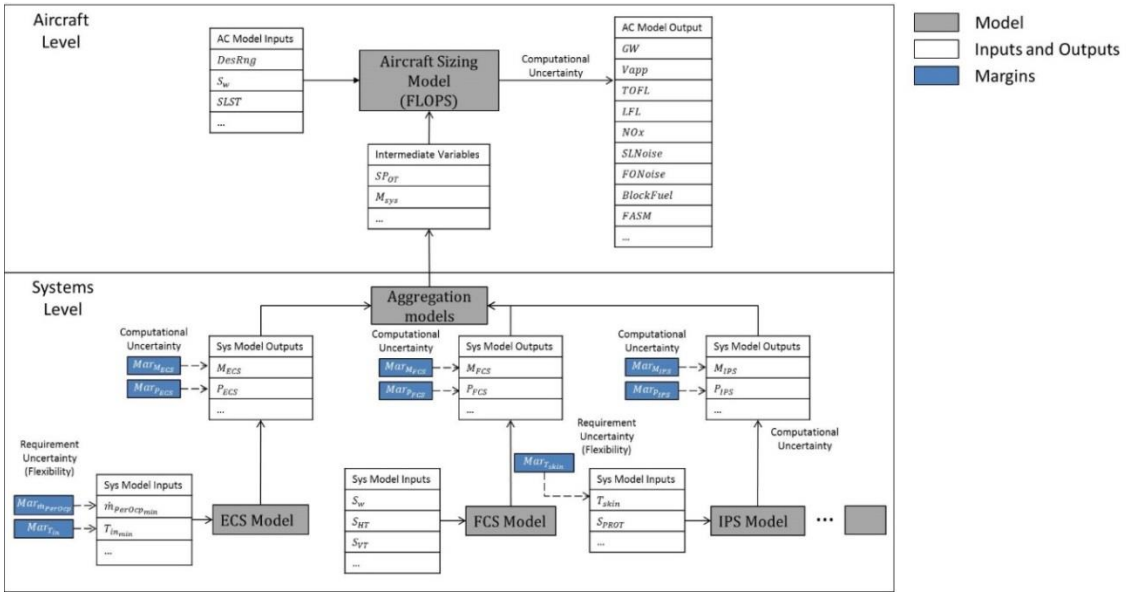

Figure 4-19. Schematic computational workflow of the design study

Table 4-3. Test case nomenclature

Category	Symbol	Name
Aircraft Level	$DesRng$	Design Range $[nm]$
	S_w	Wing Reference Area $[ft^2]$
	$SLST$	Engine Sea-Level Static Thrust $[lb]$
	GW	Gross Weight $[lb]$
	$Vapp$	Approach Velocity $[kt]$
	$TOFL$	Take-Off Field Length $[ft]$
	LFL	Landing Field Length $[ft]$
	NO_x	Nitrogen Oxide Emissions $[lb]$
	$SLNoise$	Sideline Noise $[EPNdB]$
	$FONoise$	Flyover Noise $[EPNdB]$
	$BlockFuel$	Block Fuel $[lb]$
	$FASM$	Fuel per Available Seat Mile $[lb/nm]$
	SP_{OT}	Total Shaft Power Off-Take $[kW]$
	M_{sys}	Total System Mass $[lb]$
	N_{Pax}	Number of Passengers
Environment Control System (ECS)	$\dot{m}_{PerOcp_{min}}$	Mass Flow Rate of Fresh Air per Occupant [kg/s]
	$T_{in_{min}}$	Temperature of Air Flow into the Cabin [°C]
	M_{ECS}	ECS Mass $[lb]$

	P_{ECS}	ECS Power Off-Take $[kW]$
	$Mar_{\dot{m}_{PerOcp}}$	Margin on Mass Flow Rate of Fresh Air per Occupant
	$Mar_{T_{in_{min}}}$	Margin on Temperature of Air Flow into the Cabin
	$Mar_{M_{ECS}}$	Margin on ECS Mass
	$Mar_{P_{ECS}}$	Margin on ECS Power Off-Take
Flight Control System (FCS)	S_{HT}	Horizontal Tail Reference Area $[ft^2]$
	S_{VT}	Vertical Tail Reference Area $[ft^2]$
	M_{FCS}	FCS Mass $[lb]$
	P_{FCS}	FCS Power Off-Take $[kW]$
	$Mar_{M_{FCS}}$	Margin on FCS Mass
	$Mar_{P_{FCS}}$	Margin on FCS Power Off-Take
Ice Protection System (IPS)	T_{skin}	Required Skin Temperature $[K]$
	S_{PROT_W}	Total Protected Area for Wing IPS $[m^2]$
	$S_{PROT_{NAC}}$	Total Protected Area for Cowling IPS $[m^2]$
	M_{IPS}	IPS Mass $[lb]$
	P_{IPS}	IPS Power Off-Take $[kW]$
	$Mar_{T_{skin}}$	Margin on Required Skin Temperature
	$Mar_{M_{IPS}}$	Margin on IPS Mass
	$Mar_{P_{IPS_W}}$	Margin on Wing IPS Power Off-Take
	$Mar_{P_{IPS_C}}$	Margin on Cowling IPS Power Off-Take
Fuel System (FS)	$Mar_{M_{FS}}$	Margin on FS Mass
	$Mar_{P_{FS}}$	Margin on FS Power Off-Take
Hydraulic Power System (HPS)	$Mar_{M_{HPS}}$	Margin on HPS Mass
Electrical Power System (EPS)	$Mar_{M_{EPS}}$	Margin on EPS Mass
Thrust Reverser System (ThrRev)	$Mar_{P_{ThrRev_{LND}}}$	Margin on Thrust Reverser Power Off-Take at Landing
Landing Gears (LG)	$Mar_{P_{LG}}$	Margin on Landing Gear Power Off-Take
Brakes (Brk)	$Mar_{P_{Brk}}$	Margin on Brake Power Off-Take
Avionics (Avion)	$Mar_{M_{Avion}}$	Margin on Avionics Mass
	$Mar_{P_{Avion}}$	Margin on Avionics Power Off-Take
Auxiliary Power Unit (APU)	$Mar_{M_{APU}}$	Margin on APU Mass
Instruments (Instr)	$Mar_{M_{Instr}}$	Margin on Instrument Mass
	$Mar_{P_{Instr}}$	Margin on Instrument Power Off-Take
Equipment (Equip)	$Mar_{M_{\mathrm{Equip}}}$	Margin on Equipment Mass
Passenger Service Units (PaS)	$Mar_{P_{PaS}}$	Margin on Passenger Service Units Power Off-Take
Inflight Entertainment System (IfEnt)	$Mar_{P_{IfEnt}}$	Margin on Inflight Entertainment System Power Off-Take
Lightings (Light)	$Mar_{P_{Light}}$	Margin on Lighting Power Off-Take
Galley (Gal)	$Mar_{P_{Gal}}$	Margin on Galley Power Off-Take
Lavatories (Lav)	$Mar_{P_{Lav}}$	Margin on Lavatory Power Off-Take

As can be seen, the workflow is divided into two hierarchical levels: aircraft (top) level and subsystems level. The top level contains the computational models for aircraft geometry and performance. Specifically, the NASA code, Flight Optimization System (FLOPS) (Mccullers, 1984), was used for overall aircraft sizing and analysis. Models at

the subsystems level were adopted from several research papers (Chakraborty et al., 2014; Lammering, 2014; Shibata et al., 2014; de Tenorio, 2010; Xia and Lawson, 2013), which compute subsystem attributes and performances. As shown in Figure 4-19, the outputs of the system models were aggregated and linked to the aircraft level, to compute the total systems weight and engine shaft power offtake.

All models were assembled into a computational workflow using the dynamic (automatic) workflow creation module (Guenov et al., 2010) in AirCADia. Figure 4-20 shows a screen capture of the computational workflow, where the purple rectangles represent the computational models, while the green and red ovals represent input and output variables, respectively. A fragment of the workflow is magnified for an intuitive image of how models and variables are connected. In total, the workflow comprises 171 models and 317 variables. The data employed are realistic, but do not represent any actual aircraft, since the purpose was solely to demonstrate the capabilities and benefits of the proposed method.

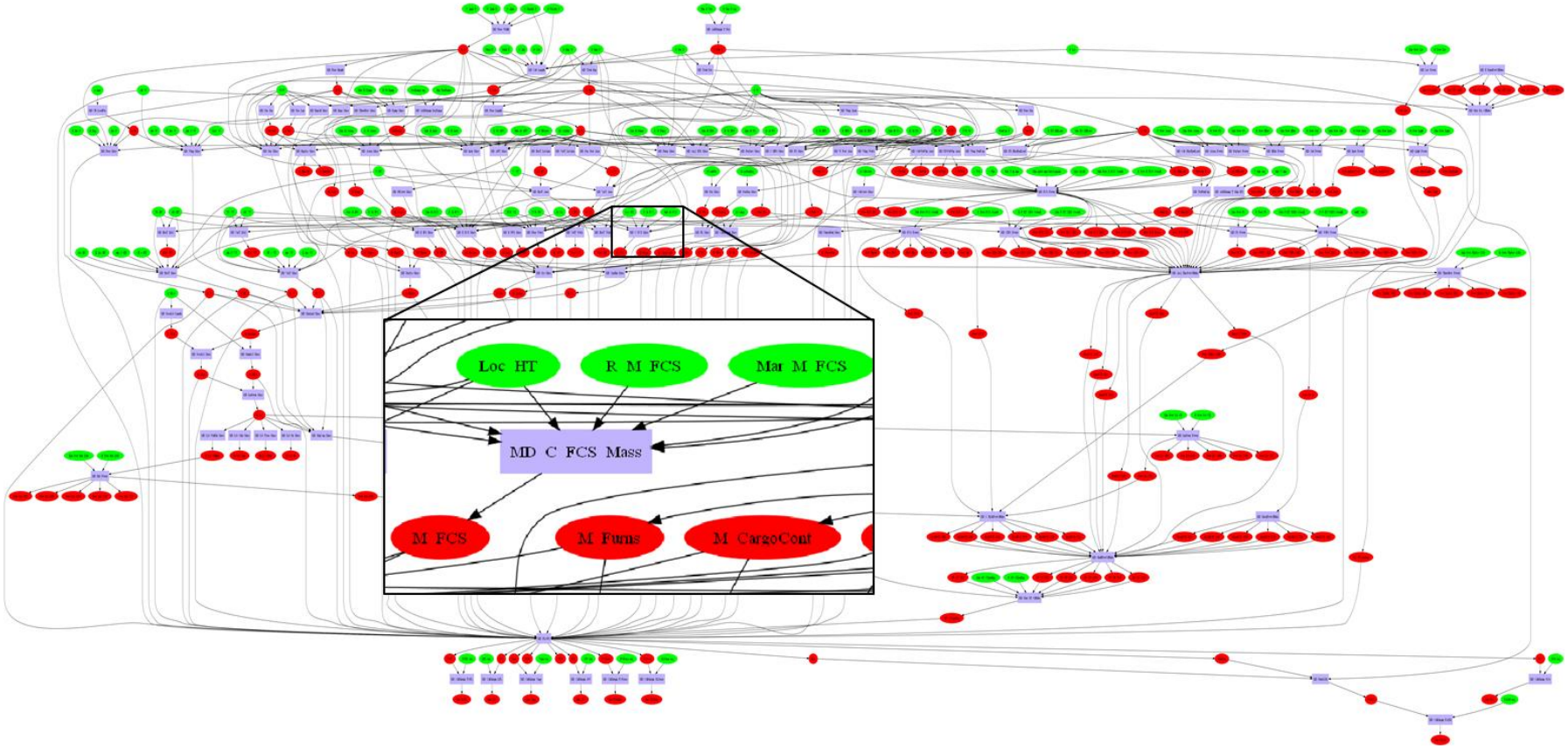

Figure 4-20. Computational workflow used for the evaluation case

4.4.2 Sensitivity Analysis and Initial Margin Allocation

After the set-up of the computational workflow, margins were allocated (on 28 variables with values of up to 30%) to address two types of uncertainty: uncertainty in the requirements and uncertainty in the computational models. For example, as shown in Figure 4-19, $Mar_{\dot{m}_{PerOcp}}$ was applied to account for changes of the requirement on the fresh air mass flow rate in the environmental control system (ECS), whereas $Mar_{M_{ECS}}$ was applied to account for uncertainty associated with the output of the ECS mass model.

DoE study with all the 28 margins will be computationally very expensive. In order to determine which margins are the most influential, a sensitivity analysis was performed, using the Fourier-Amplified Sensitivity Test (FAST) (Cukier, 1973; Cukier, Levine and Shuler, 1978) (also refer to Section 3.4). Uniform distributions were applied on all the margins, ranging between 0 and 30%. The result of the sensitivity analysis is shown in Figure 4-21.

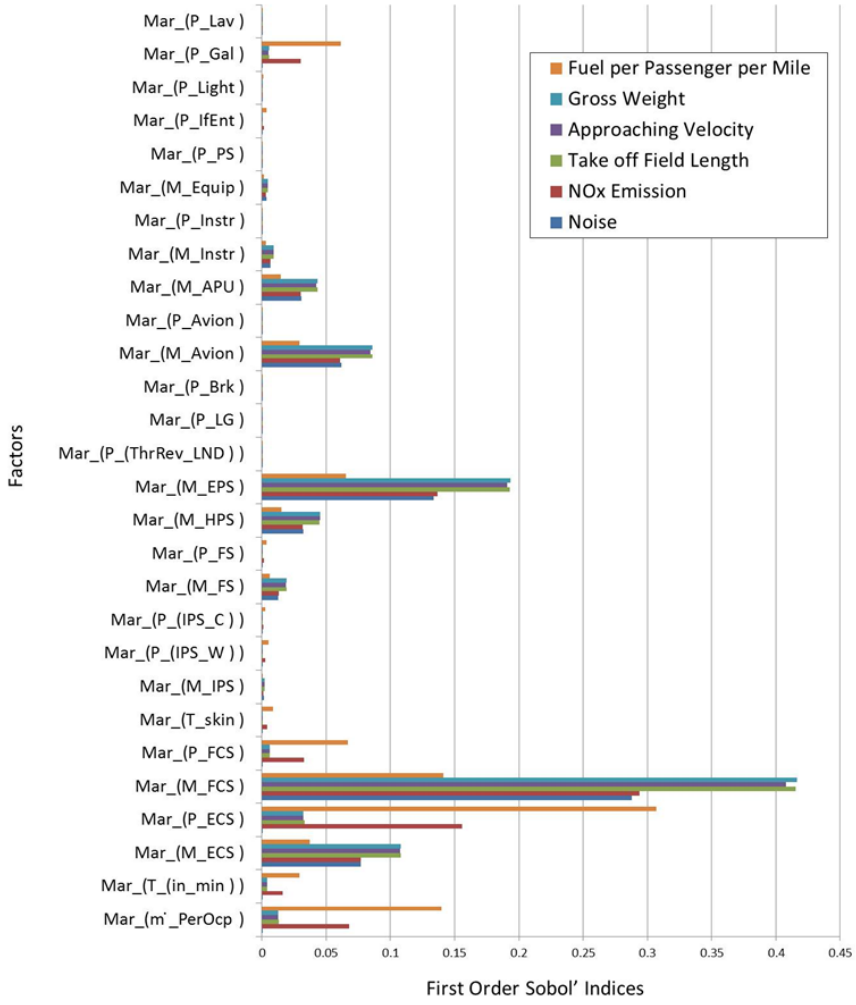

Figure 4-21. Sensitivity analysis result for all margins

As can be seen from Figure 4-21, the five most influential margins are $Mar_{\dot{m}_{PerOcp}}$, $Mar_{M_{ECS}}$, $Mar_{M_{FCS}}$, $Mar_{P_{ECS}}$, and $Mar_{M_{EPS}}$. For demonstration, three of these were of particular interest for this study, as summarized in Table 4-4.

Table 4-4. Three margins selected for trade-off

Parameter	**Margin**	**Need/Justification**
Min. fresh air mass flow rate per occupant	$Mar_{\dot{m}_{PerOcp}}$	Uncertainty in requirement
ECS mass	$Mar_{M_{ECS}}$	Computational model uncertainty
FCS mass	$Mar_{M_{FCS}}$	Computational model uncertainty

Two design variables, wing reference area (S_W) and sea-level static thrust ($SLST$) were considered. A full factorial DoE study was then performed in AirCADia to populate the design and margin spaces. The lower and upper bounds of the design variables and

margins are listed in Table 4-5, along with the corresponding numbers of levels for the DoE.

Table 4-5. Design of experiments study formulation

Category	Name	Parameter	Lower bound	Upper bound	Levels
Design Variables	Wing Area	$S_W\ (ft^2)$	1350	1390	5
	Sea-Level Static Thrust	$SLST\ (lbf)$	27000	29000	5
Margins	Fresh air mass flow rate	$Mar_{\dot{m}_{PerOcp}}$	0%	30%	4
	Environmental control system mass	$Mar_{M_{ECS}}$	0%	30%	4
	Flight control system mass	$Mar_{M_{FCS}}$	0%	30%	4

The performance constraints are listed in Table 4-6. For clarity of display, only three performance constraints are considered. The result of the DoE study is shown in Figure 4-22. The plot shows simultaneously the design space, the margin space, the performance space, and the performance constraint values. All feasible solutions are shown in blue, whereas the infeasible ones are shown in grey.

Table 4-6. Performance constraints

Constraint	Parameter	Limiting value
Take-off field length	$TOFL$	$\leq 5500\ ft$
Approach velocity	V_{app}	$\leq 136\ kt$
NOx emissions	NOx	$\leq 500\ lb$

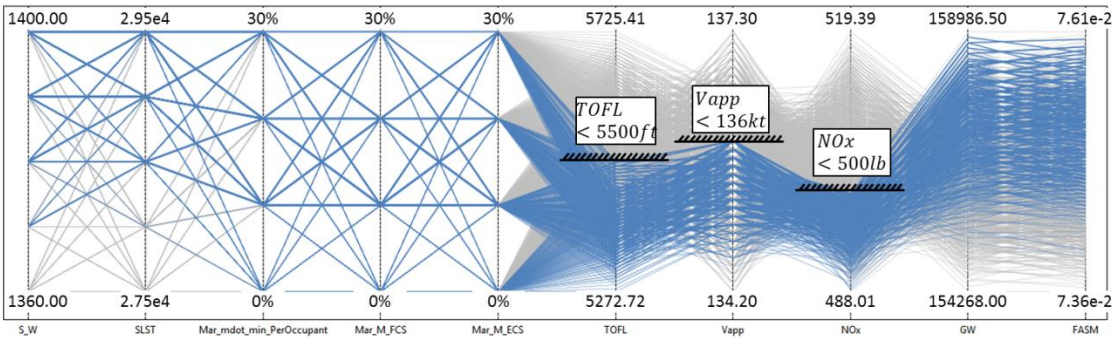

Figure 4-22. Results of the design of experiments study

4.4.3 Uncertainty Propagation

The sources of uncertainty considered, along with the assumed distributions, are listed in Table 4-7. These sources include both input and model uncertainties. Generally, such distributions can be obtained from historical data and/or expert elicitation. If this

information is not available, the designer may proceed to the margin vs. performance trade-off (as demonstrated in the next sub section).

Table 4-7. Sources of uncertainty

Parameter	Notation	Distribution
Fuselage skin heat transfer coefficient $[W/(m^2K)]$	c_{wall}	Uniform distribution: lower bound=1; upper bound=1.4
Heat load from the flight deck $[W]$	$\dot{Q}_{FD}$	Random factor $R_{Q_{FD}}$: uniform distribution: lower bound=0.8; upper bound=1.2
ECS mass $[lb]$	M_{ECS}	Random factor $R_{M_{ECS}}$: triangular distribution: left point=0.7; top point=1; right point=1.3
FCS mass $[lb]$	M_{FCS}	Random factor: $R_{M_{FCS}}$: triangular distribution: left point=0.7; top point=1; right point=1.3

Before uncertainty propagation, the margins on ECS and FCS mass had to be removed, because placing the distribution on a variable in addition to its margin would result in double accounting for uncertainty (as explained in Section 4.3.3). This was not required for the margin on the ECS fresh mass flow rate, since this margin was applied to accommodate flexible requirement, which would not be assigned any probability distribution.

The results of the uncertainty propagation are shown in Figure 4-23 as an extension of the deterministic results shown in Figure 4-22. The mean values of the output variables and the probabilities of constraint satisfaction have been added as the last eight parallel coordinates. For example, the arbitrarily selected poly line, highlighted in orange, represents a point in the DB^D with a particular combination of margins. The red error bars represent the standard deviation (±1σ) of the performance variables.

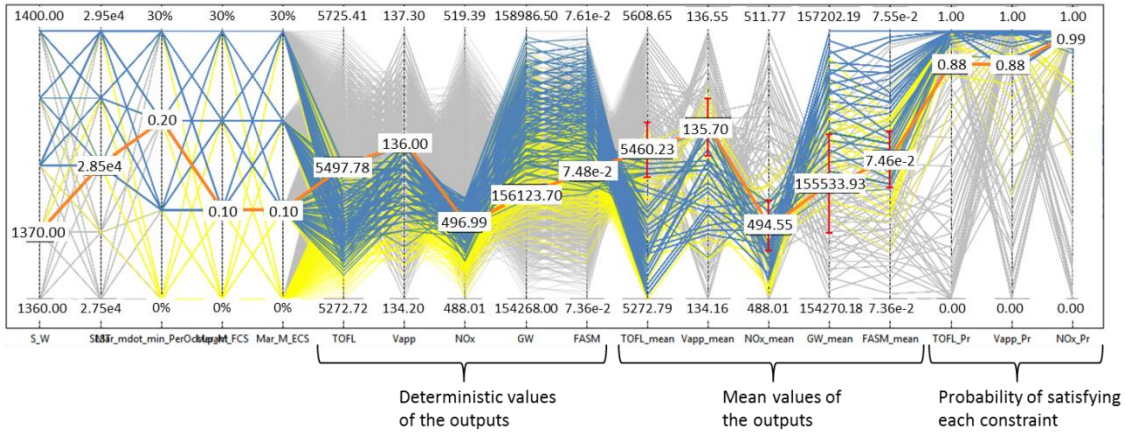

Figure 4-23. Uncertainty propagation result for a single design point

While not central to this particular research, it should be noted that the URQ method only estimates the means and standard deviations of the outputs. Therefore, to calculate the probabilities of constraint satisfaction, assumptions needed to be made regarding the output distributions. For this purpose, a Monte Carlo simulation was conducted with 10000 model evaluations. The resulting histogram (as shown in Figure 4-24) indicated that a Gaussian distribution could be assumed for the output variables. It should be mentioned that, in this test case, this assumption was based on visual comparison only, because the focus was not on statistics of the results, but on demonstration of the trade-off studies. In a practical design process, some more rigours approaches can be applied for normality test, for instance, the Kolmogorov-Smirnov analysis (Massey and Jr., 1951) and Kullback–Leibler divergence (Kullback and Leibler, 1951).

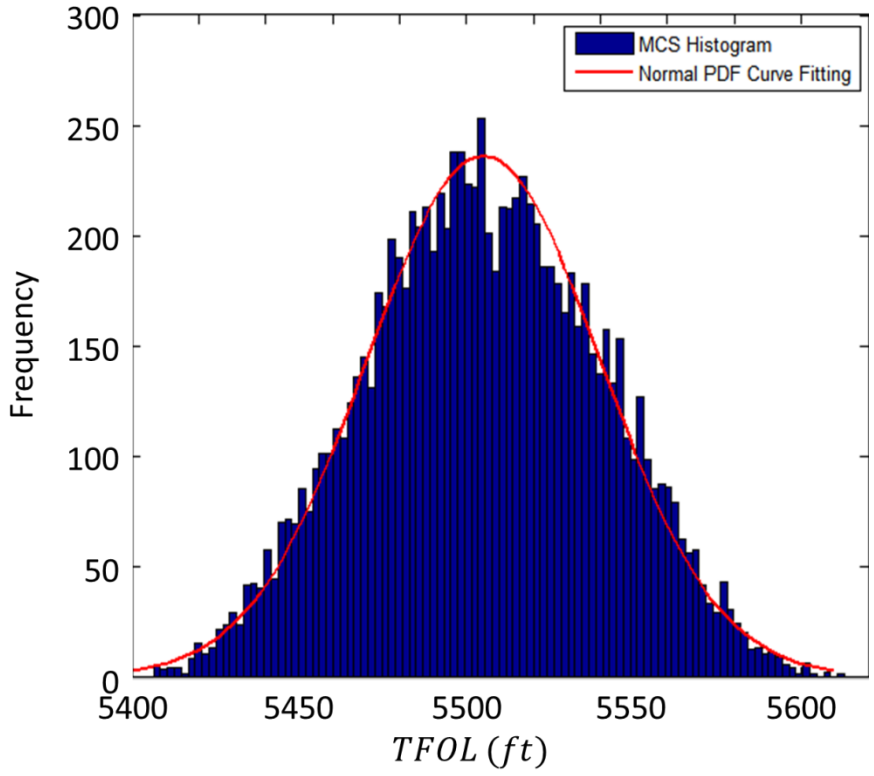

Figure 4-24. Histogram of MCS to assess the distribution of the outputs

4.4.4 Trade-off Study

4.4.4.1 Margin vs. Margin

Recall from Section 4.3.2 that the allowable 'room' for allocating a particular margin is dependent on the selected design point and the selected magnitudes of the other

margins. This can be seen again in Figure 4-25 and Figure 4-26, where the design space and corresponding margin spaces for design points DP1 ($SW = 1300\ ft^2$, $SLST = 28500\ lbf$) and DP2 ($SW = 1380\ ft^2$, $SLST = 28500\ lbf$) are shown, respectively. In both figures, the constraints on take-off field length ($TOFL$), approach velocity (V_{app}), and emission (NOx) are projected on the two-dimensional design and margin spaces. It can be observed that DP2 would provide more room for allocating margins compared with PD1, as indicated by the white feasible regions in the plots.

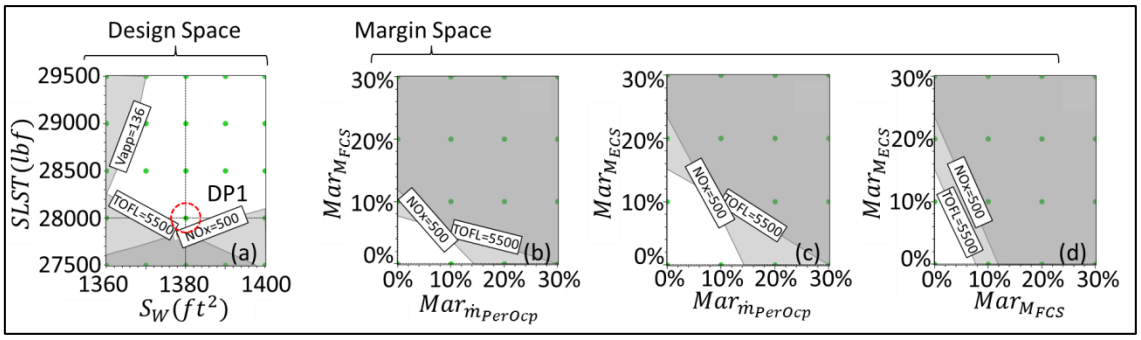

Figure 4-25. Margin exploration for DP1

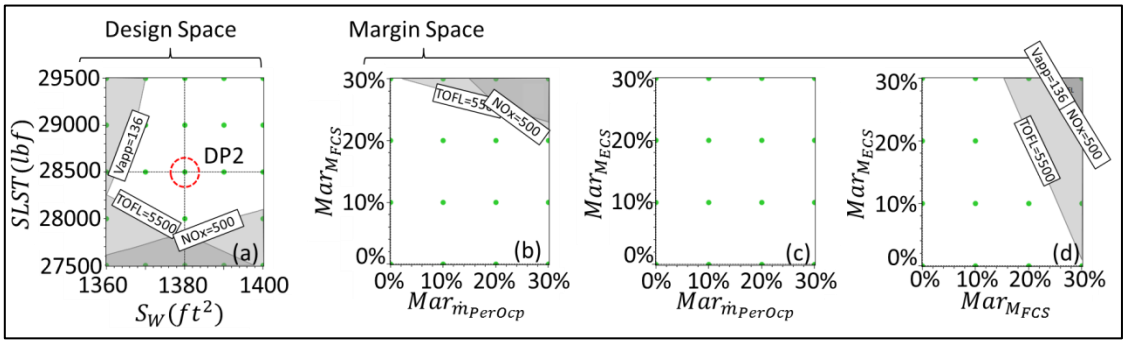

Figure 4-26. Margin exploration for DP2

After selecting a particular design point (for instance DP2), the designer can start trading-off the margins with each other (as described in Section 4.3.4.1). Assume that the designer decided to allocate a 20% margin on FCS mass ($Mar_{M_{FCS}} = 20\%$) and 10% margin on the ECS fresh air mass flow rate ($Mar_{\dot{m}_{PerOcp}} = 10\%$), the resulting allowable room for the margin on the ECS mass ($Mar_{M_{ECS}}$) may not be sufficiently large, which is indicated by the green arrows in Figure 4-27 (a).

Subsequently, selecting a different point in the margin space, i.e. 10% margin on FCS mass and 20% margin on the ECS fresh air mass flow rate will modify the feasible (allowable) margin space, as shown in Figure 4-27 (b).

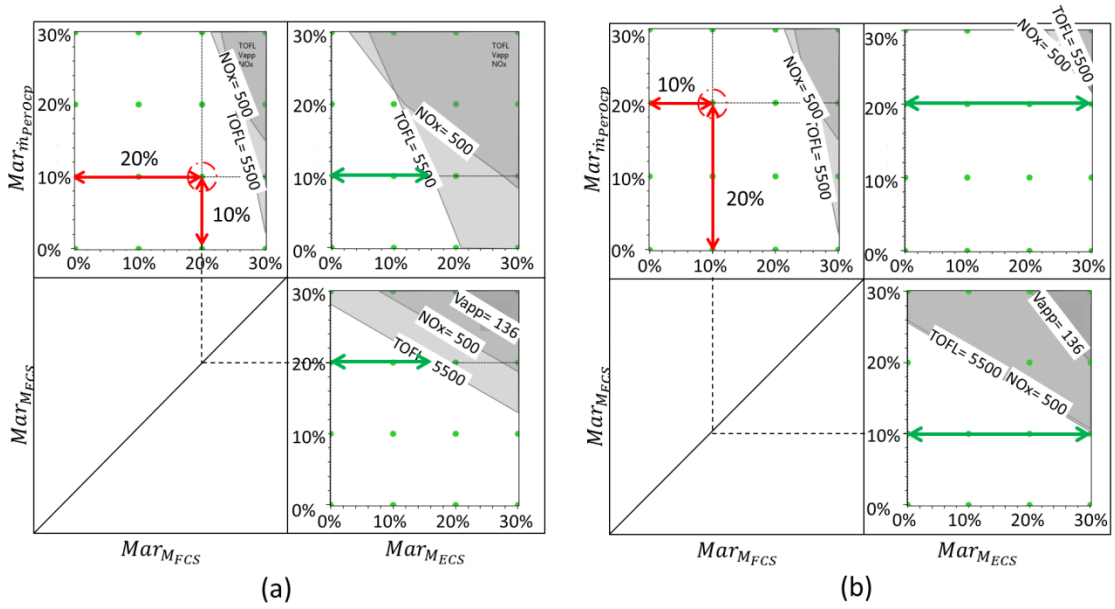

Figure 4-27. Margin space assignment for DP2 in the design space

Finally, suppose that the designer wishes to have a value of at least 10% for all the margins. This will reduce the feasible design space to that shown in Figure 4-28 and Figure 4-29. This leads to a reduction in the number of design points with different margin combinations from 1600 to 124. The parallel coordinates plot in Figure 4-29 shows the deterministic design band which is represented by the region enclosed by the dashed-blue lines. All the design points in this band could therefore maintain at least 10% margins on ECS mass, FCS mass, and ECS mass flow rate, and still remain feasible.

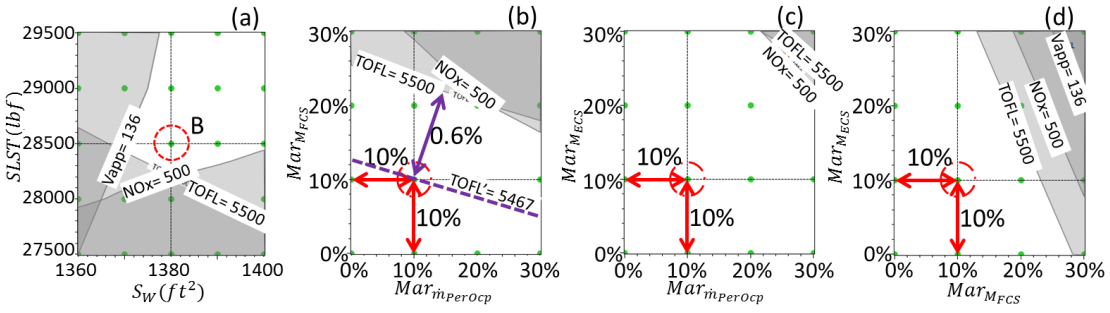

Figure 4-28. An example of a feasible combination of margins

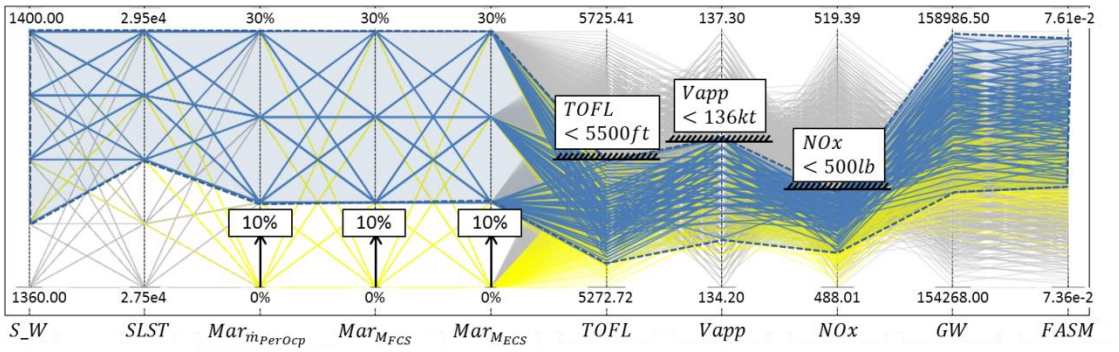

Figure 4-29. Deterministic design band

It can be noted from Figure 4-28 (b) that, after selecting the minimum value of 10% on all the chosen margins, the allowable room for margin on the $TOFL$ requirement becomes only 0.6%. This value is calculated by comparing the current constraint ($TOFL \leq 5500\ ft$) and the take-off field length calculated at the selected point in the design/margin space (in this case, $TOFL = 5467\ ft$). It should be noted that the purple double headed arrow in Figure 4-28 (b) is not a margin, but the distance between the current constraint and the best achievable one through trade-off (indicated by the purple dash line).

This can be further illustrated in Figure 4-30, where the blue surface is the model response ($TOFL$) over the margin space of $Mar_{\dot{m}_{PerOcp}}$ and $Mar_{M_{FCS}}$. The two grey planes are respectively the current constraint at $5500\ ft$ and the best achievable constraint which goes through the selected design/margin point ($5467\ ft$). The projections of these two constraints lead to the contour plots shown on the left. The actual margin on $TOFL$ can be defined as the distance between these two grey planes (as indicated by the vertical red arrow), therefore it is orthogonal to the margin space of $Mar_{\dot{m}_{PerOcp}}$ and $Mar_{M_{FCS}}$, and cannot be projected to the counter plots on the left.

In the current test-case, these two constraints can be considered as straight lines and parallel to each other. This is because the total aircraft weight is promotional to $Mar_{\dot{m}_{PerOcp}}$ and $Mar_{M_{FCS}}$, and within this small increment of weight, the response of $TOFL$ is close to linear. Therefore the purple arrow is plotted to be perpendicular to the two constraints. It should be stated that this arrow is only to visually indicate the shift of the constraint, while rigorous definition of this arrow (as a vector) is beyond the scope

of this research. In a highly non-linear case, this distance might be represented as a set of vectors, as illustrated in Figure 4-31 (using an assumed model).

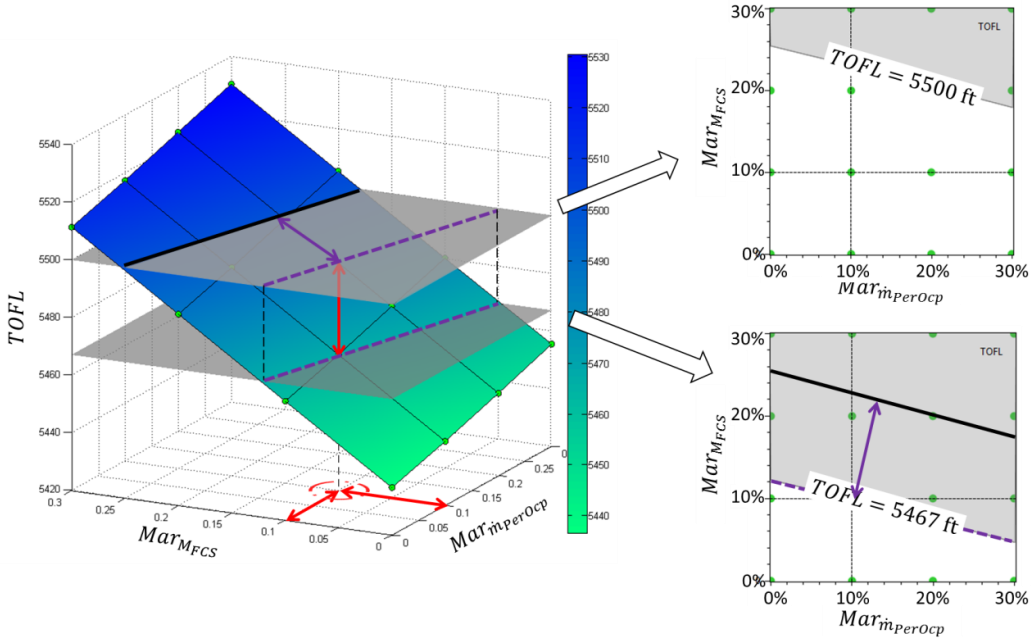

Figure 4-30. 3-Dimensional plot of the model response and projections of constraints

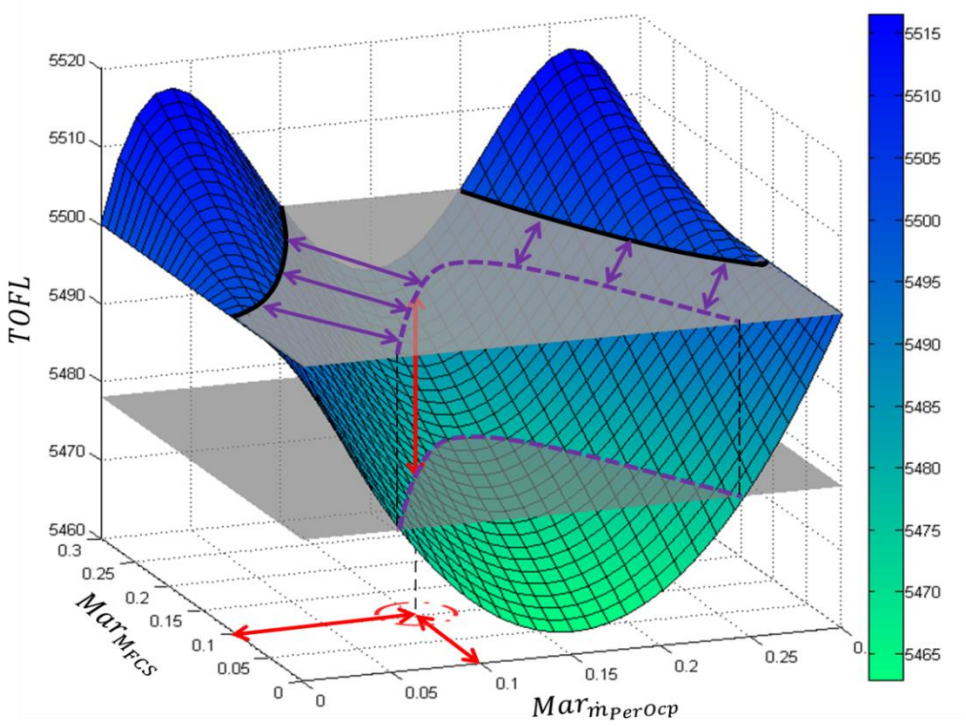

Figure 4-31. Model response and constraints of an assumed non-linear model

4.4.4.2 Margin vs. Probability of Constraint Satisfaction

Next, the user could interactively adjust the margins and probability thresholds to perform trade-off studies, as shown in Figure 4-32 for margin combinations listed in Table 4-8. Assume that the required probabilities of constraint satisfaction for $TOFL$, $Vapp$, and NOx were all selected to be 90%. This defines the robust (white) and non-robust (grey shaded) regions in the design space, as indicated in Figure 4-32 (a).

Following the approach described in Section 4.3.4.2, the contours formed in the design space, resulting from the allocation of deterministic margins, can now be combined (overlapped) with the probabilistic contours. This is illustrated in Figure 4-32 (a)-(f) where the blue lines represent the deterministic constraints and the blue-shaded regions represent the infeasible regions for particular margin combinations (Table 4-8). As can be seen in the figure, more stringent combinations constrict the design space further. A reasonable selection of margins would be MP3 which would 'just' cover the probabilistic constraints (i.e. rendering the design space slightly smaller than what it would have been with only the probabilistic constraints present). An undesirable case will be the margin combination MP6, where the NOx constraint renders the design space infeasible.

Table 4-8. A selection of possible margin combinations for trade-off with probability of constraint satisfaction.

Combination	$Mar_{M_{ECS}}$	$Mar_{M_{FCS}}$	$Mar_{\dot{m}_{PerOcp}}$
MP1	10%	10%	10%
MP2	10%	20%	10%
MP3	10%	20%	20%
MP4	30%	30%	10%
MP5	30%	30%	20%
MP6	30%	30%	30%

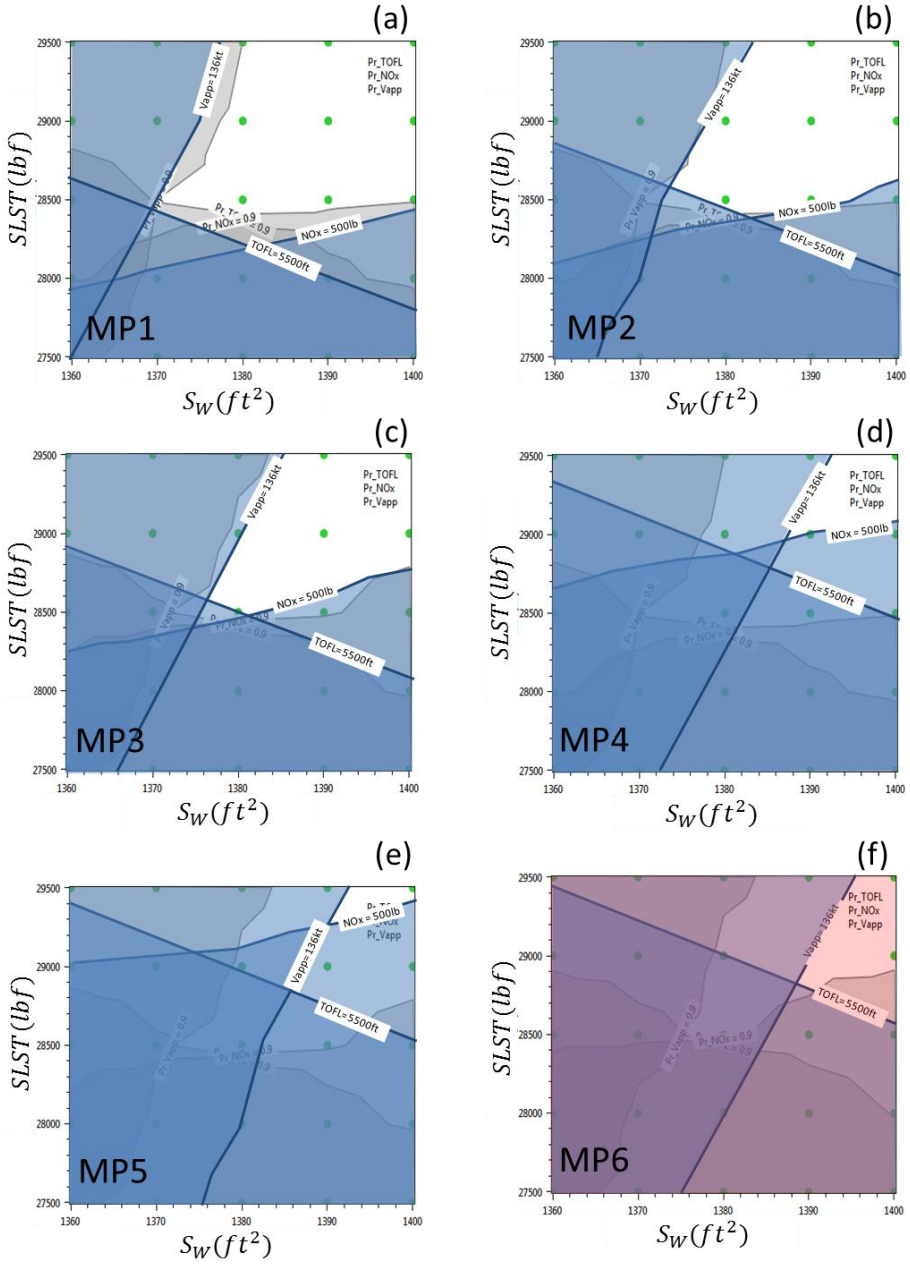

Figure 4-32. Different combinations of margins with 90% probability constraint satisfaction for all three constraints ($TOFL$, $Vapp$, and NOx)

To determine the robust design band, 90% thresholds were placed on the probabilistic constraints. The results are shown in Figure 4-33. In this case, only three points were discarded (as indicated by red), which means that the minimum values placed on margins in Section 4.4.4.1 were sufficient to account for uncertainty. Depending on the assessment of the available feasible space, the decision maker may attempt to reduce (some of) the margins to ensure improved performance.

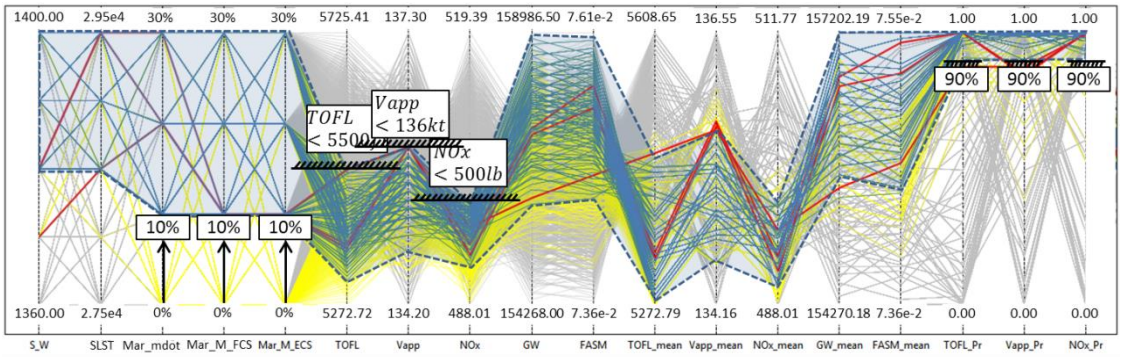

Figure 4-33. Robust design band filtering using parallel coordinates plot

4.4.4.3 Margins vs. Performance

As described in Section 4.3.4.3, assigning different combinations of margins lead to the 'shifting' of the Pareto front in the performance space. A Pareto front investigation can be applied to trade-off the margins and performances as illustrated in Figure 4-34, for instance, using the margin combinations listed in Table 4-9. The performance variables considered are the nitrogen oxide emissions (NOx) and fuel per available seat mile ($FASM$). Combination MP1 (represented by the red dashed line) provides the best performance compared to the other margin combinations.

Table 4-9. A selection of possible margin combinations for performance trade-off

Combination	$Mar_{\dot{m}_{PerOcp}}$	$Mar_{M_{FCS}}$	$Mar_{M_{ECS}}$
MP1	10%	10%	10%
MP2	20%	10%	20%
MP3	20%	20%	10%
MP4	30%	10%	30%

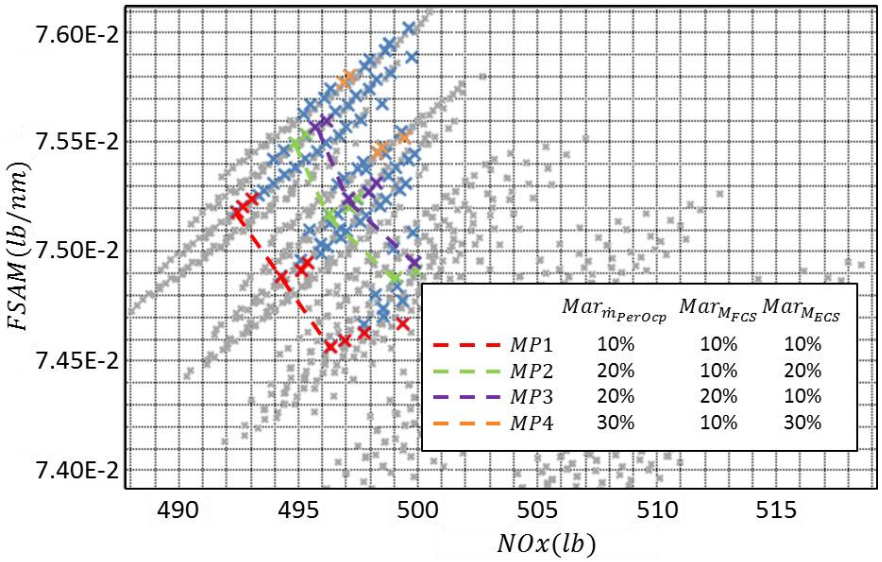

Figure 4-34. Exploring the effect of different combinations of margins on the location of the Pareto front

A parallel coordinates plot can be used to further down-select the design points. For instance, if the designer identifies that the Pareto front corresponding to margin combination MP3 (represented by the purple dashed line in Figure 4-34) is the minimum accepted performance, then he/she can use the parallel coordinates plot to filter out the margin combinations which are more conservative than MP3. As shown in Figure 4-35, this leads to the reduction of the number of design points from 121 down to 30. In AirCADia, the designer can 'drag' the filter bars for the different margin values in the parallel coordinates plot and immediately view the effect on the performance parameters. Furthermore, the designer may also apply filtering on performance parameters, e.g., fuel per available seat mile ($FASM$).

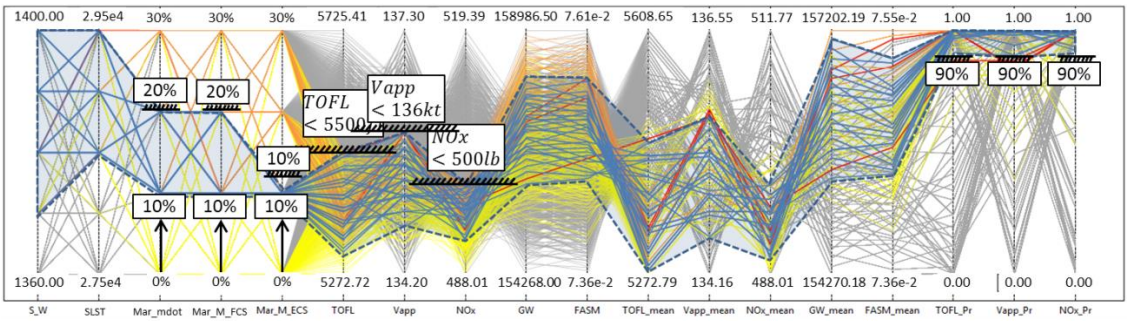

Figure 4-35. Effects of filtering margins on performance

4.4.5 What-if scenarios

The trade-off techniques presented in the previous section enable the designer to answer certain what-if questions.

1. If the weight growth for a certain system exceeds the assigned margin, then the designer can reduce margins on other systems, to re-allocate them on the overweight system. For example, if at a stage of the design process, it is found that the selected margin on the FCS mass, $Mar_{M_{FCS}}$, is not sufficient, then the designer can attempt to reduce the margin on the ECS mass, $Mar_{M_{ECS}}$, and assign extra margin on the mass of the FCS. Such a case is illustrated in Figure 4-36, where the initial margin combination is shown by the red arrows, while a possible revised selection is indicated by the blue arrows.

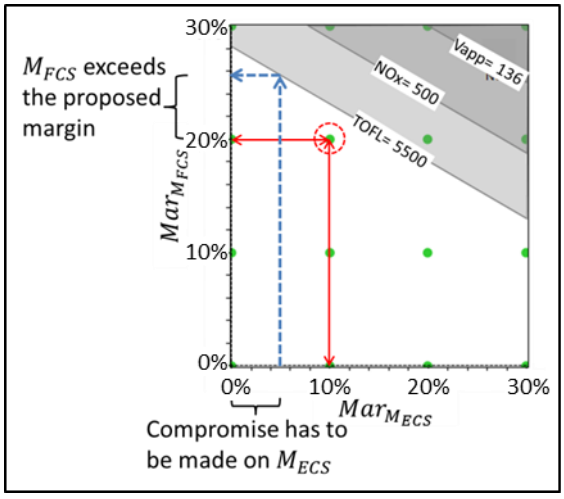

Figure 4-36. Reallocation of margins from one system to another

2. If, for whatever reason, a requirement is changed during the design process (for example, a reduction in required take-off field length, $TOFL$), the designer can proceed with one of the following two options:

   A. Trade between different margins (as described above). This is shown in Figure 4-37 (a), the selected point in the margin space allows a scope for only 0.6% reduction on the constraint of $TOFL$. (Note that the meaning of the purple arrows is the same as explained earlier with regard to Figure 4-28). By reducing the margins $Mar_{M_{FCS}}$ and $Mar_{M_{\dot{m}_{PerOcp}}}$, this scope could be increased to 1%, as

shown in Figure 4-37 (b). However, if the required reduction on the $TOFL$ is 5%, no margin combination will be feasible, as shown in Figure 4-37 (c).

B. An investigation in the design space shows that, even when changing to another design point, the margin of 5% can still not be achieved by any design and margin combination. Based on the knowledge gained during the design and margin space exploration, a possible course of action is to expand the design space. Not surprisingly, in this case study, the harder $TOFL$ constraint can be met by increasing the wing area and/or sea-level static thrust, as shown by the direction of the purple arrow in Figure 4-38 (c).

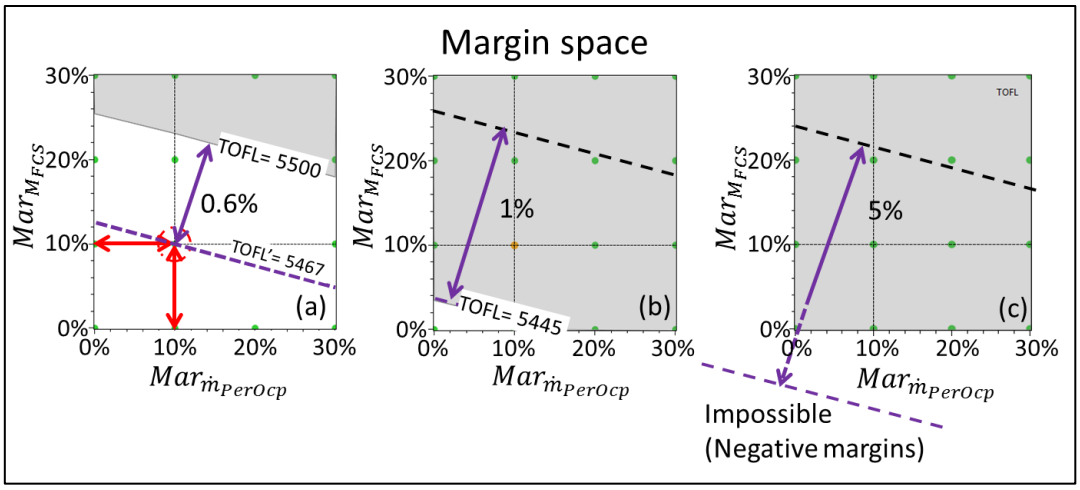

Figure 4-37. Accommodating change of take-off field length $TOFL$ requirement

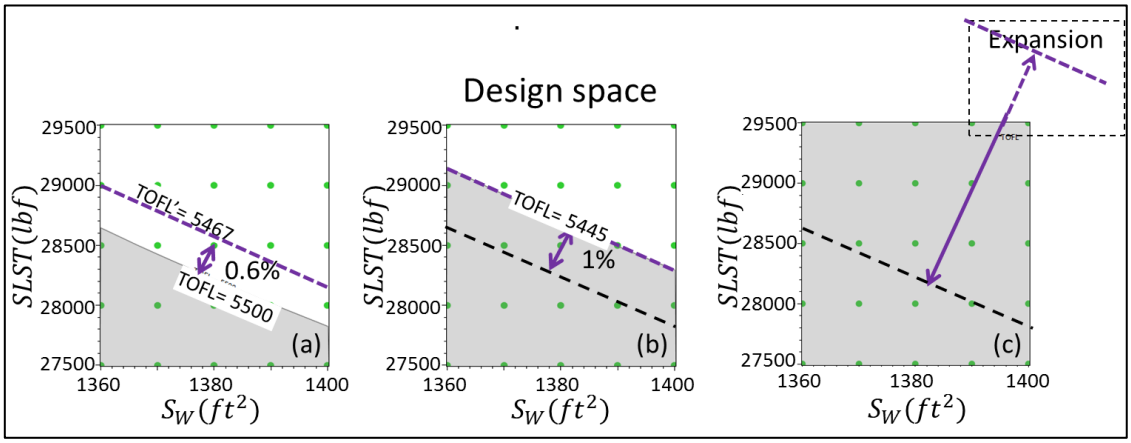

Figure 4-38. Required expansion of the design space to accommodate the desired margin

4.5 Summary and Conclusions

Presented in this chapter is a novel method for margin allocation and trade-off study. The proposed method integrates deterministic margin allocations with uncertainty analysis (sensitivity analysis and propagation), following a set-based design paradigm for different trade-off studies. The method was explained with a running example and demonstrated with a realistic test case. Compared with the existing approaches (Birman, 2013; Cooke et al., 2015; Thunnissen, 2004, 2005; Thunnissen and Tsuyuki, 2004; Yuan et al., 2016; Zang et al., 2015), the novelty includes:

- Incorporation of different types of margins to handle uncertainties not only due to computational models, but also from flexible requirements. This improves the margin allocation at early design stages, when the design requirements are not fully defined. Also it provides means to account for possible future developments.
- The interactions between margins and design variables are explored in an explicit manner. This helps the designer to understand the trends in the design/margin space, rather than only down selecting the design solutions. With more knowledge of the design/margin space, better decisions can be made to handle the competing or conflicting relationships between margins and design variables. Additionally, the designer will be able to predict the direction to expand the design space for further exploration, if there is no feasible solution.
- The trade-off studies are organised in a systematic way, and are enabled by several novel techniques. These developments will improve the decision making for margin management under multiple criteria.

Future work will focus on three aspects:

- Currently the margins are allocated from bottom-up and propagated to the top-level performances. The next step is to account for margin decomposition from the top-down in the design hierarchy. For instance, if the margin on the $TOFL$ is defined *a priori*, the challenge will be to properly apportion this margin on the relevant design variables.
- In the current implementation, the computational workflow is fixed throughout the design study. A future research is to investigate the evolution of margins

during the design process, where the model fidelity is gradually increased. As the maturity of design solution grows, the proportion of the margins due to computational uncertainty (lack of knowledge) should ideally diminish, while the proportion accounting for design flexibility will be maintained.

- Currently, full factorial design of experiment is applied for its simplicity. If the designer wants to treat a large number of margins simultaneously, Latin Hypercube and other sampling strategies can be adopted to reduce the computational cost of the proposed method.

5 Inverse Uncertainty Propagation

5.1 Introduction

In this chapter, a method is proposed for inverse uncertainty propagation, as part of the broader problem of uncertainty allocation. Recall that the latter is defined in such a context that controlling the amount of uncertainty is also part of the decision making. Within this context, inverse uncertainty propagation is to estimate the uncertainty from different sources, given the known or pre-defined amount of uncertainty associated with the system outputs.

According to the literature review, several numerical approaches have been proposed to address this problem (Section 3.3.4). However, the implementation of the existing methods is complicated, and requires different setups for every specific case study. On the other hand, we would like to enable the interactive exploration at early design stages, by asking "what-if" types of questions. For instance, the designer can specify on-the-fly different amount of uncertainty associated with different output variables, and immediately receive a feedback regarding if the target is attainable, and if so, how much input uncertainty can be tolerated. This requires dynamic reformulation of the problem, which has been largely restricted by the existing methods.

To overcome this issue, the proposed method integrates a fast forward propagation method with the workflow reversal capability of AirCADia. These enablers are briefly explained in Section 5.2, while the proposed method is presented in Section 5.3. In Section 5.4, several numerical test-cases are used for validation. Finally, summary and conclusions are presented in Section 5.5.

5.2 Background

5.2.1 Univariate Reduced Quadrature Method

The Univariate Reduced Quadrature (URQ) method was proposed by (Padulo, Campobasso and Guenov, 2011) for efficient uncertainty propagation. It takes the first four statistical moments (mean μ_{x_i}, standard deviation σ_{x_i}, skewness γ_{x_i}, and kurtosis Γ_{x_i}) of the stochastic inputs, and produces the means and variances of the outputs:

$$E(y) \approx E_{URQ}(y) = W_0 f(\boldsymbol{\mu_x}) + \sum_{i=1}^{n} W_i \left[\frac{f(\boldsymbol{x_i^+})}{h_i^+} - \frac{f(\boldsymbol{x_i^-})}{h_i^-} \right] \quad \textbf{(5-1)}$$

$$V(y) \approx V_{URQ}(y) \quad \textbf{(5-2)}$$
$$= \sum_{i=1}^{n} \left\{ W_i^+ \left[\frac{f(\boldsymbol{x_i^+}) - f(\boldsymbol{\mu_x})}{h_i^+} \right]^2 + W_i^- \left[\frac{f(\boldsymbol{x_i^-}) - f(\boldsymbol{\mu_x})}{h_i^-} \right]^2 + W_i^{\pm} \frac{[f(\boldsymbol{x_i^+}) - f(\boldsymbol{\mu_x})][f(\boldsymbol{x_i^-}) - f(\boldsymbol{\mu_x})]}{h_i^+ h_i^-} \right\}$$

The coefficients in equation (5-1) and (5-2) are listed in Table 5-1. To produce the results, this method requires $2n + 1$ model evaluations in total.

Table 5-1. Coefficients and vectors used in URQ (Padulo, Campobasso and Guenov, 2011)

$h_i^+ = \frac{\gamma_{x_i}}{2} + \sqrt{\Gamma_{x_i} - \frac{3\gamma_{x_i}^2}{4}}$	$W_0 = 1 + \sum_{i=1}^{n} \frac{1}{h_i^+ h_i^-}$
$h_i^- = \frac{\gamma_{x_i}}{2} - \sqrt{\Gamma_{x_i} - \frac{3\gamma_{x_i}^2}{4}}$	$W_i = \frac{1}{h_i^+ - h_i^-}$
$\boldsymbol{\mu_x} = [\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}]$	$W_i^+ = \frac{(h_i^+)^2 - h_i^+ h_i^- - 1}{(h_i^+ - h_i^-)^2}$
$\boldsymbol{x_i^+} = [\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i} + h_i^+ \sigma_{x_i}, \dots, \mu_{x_n}]$	$W_i^- = \frac{(h_i^-)^2 - h_i^+ h_i^- - 1}{(h_i^+ - h_i^-)^2}$
$\boldsymbol{x_i^-} = [\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i} + h_i^- \sigma_{x_i}, \dots, \mu_{x_n}]$	$W_i^{\pm} = \frac{2}{(h_i^+ - h_i^-)^2}$

5.2.2 AirCADia Workflow Reversal

Reversal is one of the unique capabilities of AirCADia, which enables the designer to swap the input and output variables of a model/workflow. An illustrative example is shown in Figure 5-1. On the left side is the original workflow, consisted of two inputs (x_1, x_2) and two outputs (y_1, y_2). On the right side is a reversed workflow, constructed by swapping x_1 with y_1 (arbitrarily selected). In this new workflow, y_1 becomes an input, for which the designer can specify a value, and x_1 will be calculated accordingly. This capability allows the designer to ask "what-if" questions and conduct flexible computational studies, for instance to inversely size a system based on a pre-defined output performance.

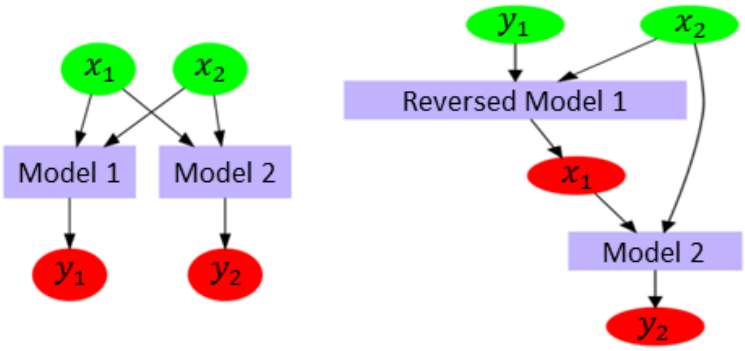

Figure 5-1 Example of workflow reversal: Original workflow (left) and Reversed workflow (right)

The reversed model/workflow is solved numerically as shown in Figure 5-2 (adapted from (Balachandran, 2007)). In this process, the value of x_2 is directly passed to the original Model 1 and a temporary value of x_1 (noted as x_1') is updated by a solver to minimise the difference between the designer specified y_1 (as a target) and the actual value y_1' (as a result of x_2 and x_1'). Once the difference falls below the tolerance (currently set as 10^{-6}), the x_1' will be returned as the output of the reversed model. In a general case where multiple input and output variables are involved in the reversal, the norm of all the differences will be used.

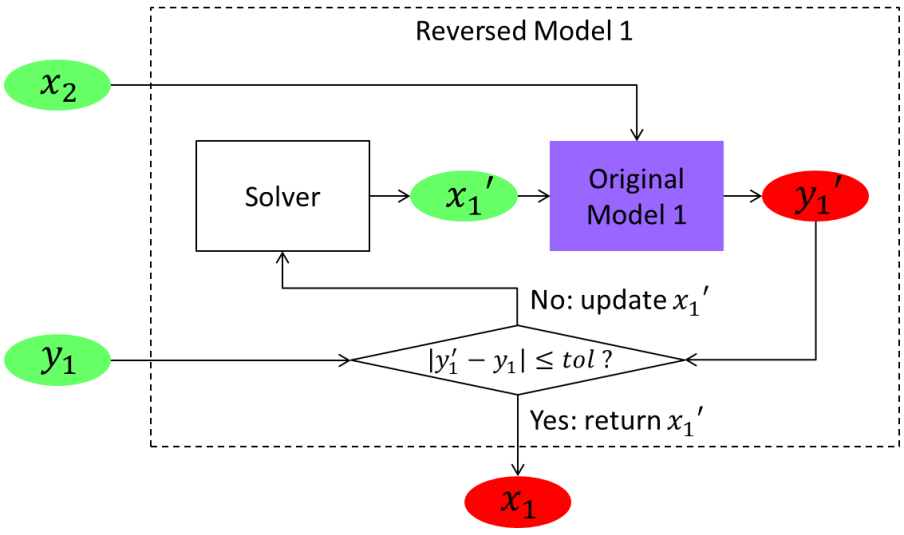

Figure 5-2 Process of solving reversed model/workflow (adapted from (Balachandran, 2007))

In AirCADia, the solving of a reversed workflow will be firstly attempted with Gaussian-Newton (GN) method (Keane and Nair, 2005). If no solution can be found

within the pre-specified limit of iterations, Particle Swarm Optimization (PSO) (Kennedy and Eberhart, 1995) will be used. In addition, algorithms have been developed in (Datta, 2014) to perform the calculation in the *Local Reversal* mode or *Global Reversal* mode. The former refers to updating only the corresponding models in the optimization loop (as in the case of Figure 5-2), while the latter evaluates the entire workflow in each iteration. It should be noted that the local/global reversal is different from the local/global solver, and is intended for reducing the computational cost.

Regarding an arbitrarily defined workflow reversal, the existence and uniqueness of a solution can be partially assessed, using the Incidence Matrix (IM) method (Guenov et al., 2006) and algorithms proposed in (Datta, 2014). Two necessary conditions for a determined solution are listed below:

- There is a dependency between the input and output variables to be reversed.
- The total number of input relationships in the IM of a reversed workflow is the same as the original workflow (if there are more input relationships, the reversal is over-determined, and vice versa under-determined)

As the models are treated as black-boxes, the aforementioned conditions are not sufficient for a determined solution. In addition, feasibility of a reversal is also dependent on the specified values of the reversed input variables. Currently, the solver will stop at a local solution which is nearest from the initial values (for example the initial value of x_1 in Figure 5-2). The designer should check if this solution is meaningful in the engineering context.

5.3 Methodology

5.3.1 Problem formulation

Considering the design workflow $f(\boldsymbol{x})$, the uncertainty of input variables are represented as a joined probability distribution defined as $\mathbb{f}_{\boldsymbol{x}}(x_1, x_2, x_3, \dots x_n)$. The task of forward propagation is to find the probability distribution of the outputs, represented as $\mathbb{f}_{\boldsymbol{y}}(y_1, y_2, y_3, \dots y_l)$. In the inverse propagation, the target is to solve for the unknown $\mathbb{f}_{\boldsymbol{x}}(x_1, x_2, x_3, \dots x_n)$, which leads to the pre-defined $\mathbb{f}_{\boldsymbol{y}}(y_1, y_2, y_3, \dots y_l)$.

To simplify the problem, statistical moments are used to represent the uncertainty, instead of the original probability density functions.

$$\boldsymbol{\mathcal{M}_x} = \begin{bmatrix} \langle x_1^1 \rangle & \langle x_1^2 \rangle & \langle x_1^3 \rangle & \cdots & \langle x_1^\alpha \rangle \\ \langle x_2^1 \rangle & \langle x_2^2 \rangle & \langle x_2^3 \rangle & \cdots & \langle x_2^\alpha \rangle \\ \langle x_3^1 \rangle & \langle x_3^2 \rangle & \langle x_3^3 \rangle & \cdots & \langle x_3^\alpha \rangle \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ \langle x_n^1 \rangle & \langle x_n^2 \rangle & \langle x_n^3 \rangle & \cdots & \langle x_n^\alpha \rangle \end{bmatrix} \tag{5-3}$$

$$\boldsymbol{\mathcal{M}_y} = \begin{bmatrix} \langle y_1^1 \rangle & \langle y_1^2 \rangle & \langle y_1^3 \rangle & \cdots & \langle y_1^\beta \rangle \\ \langle y_2^1 \rangle & \langle y_2^2 \rangle & \langle y_2^3 \rangle & \cdots & \langle y_2^\beta \rangle \\ \langle y_3^1 \rangle & \langle y_3^2 \rangle & \langle y_3^3 \rangle & \cdots & \langle y_3^\beta \rangle \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ \langle y_l^1 \rangle & \langle y_l^2 \rangle & \langle y_l^3 \rangle & \cdots & \langle y_l^\beta \rangle \end{bmatrix} \tag{5-4}$$

Here, $\langle x_i^j \rangle$ (or $\langle y_i^j \rangle$) is the j_{th} statistical moment of the i_{th} input (or output) variable.

A matrix of target output statistical moments $\boldsymbol{\mathcal{M}_y^*}$ is defined by equation (5-5). The values inside this matrix are supposed to be known or specified *a priori* by the designer based on engineering requirements.

$$\boldsymbol{\mathcal{M}_y^*} = \begin{bmatrix} \langle y_1^1 \rangle^* & \langle y_1^2 \rangle^* & \langle y_1^3 \rangle^* & \cdots & \langle y_1^\beta \rangle^* \\ \langle y_2^1 \rangle^* & \langle y_2^2 \rangle^* & \langle y_2^3 \rangle^* & \cdots & \langle y_2^\beta \rangle^* \\ \langle y_3^1 \rangle^* & \langle y_3^2 \rangle^* & \langle y_3^3 \rangle^* & \cdots & \langle y_3^\beta \rangle^* \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ \langle y_l^1 \rangle^* & \langle y_l^2 \rangle^* & \langle y_l^3 \rangle^* & \cdots & \langle y_l^\beta \rangle^* \end{bmatrix} \tag{5-5}$$

An optimization is then set up to minimise the squared sum of the differences between the actual and target output statistical moments:

$$F_{obj}(\boldsymbol{\mathcal{M}_x}) = \sqrt{\sum_{i=1}^{l} \sum_{j=1}^{\beta} \left(\langle y_i^j \rangle - \langle y_i^j \rangle^* \right)^2}, \tag{5-6}$$

where $\langle y_i^j \rangle$ is obtained from forward uncertainty propagation (as a function of $\boldsymbol{\mathcal{M}_x}$).

5.3.2 Proposed process

In the implementation, assuming that a forward propagation has already been conducted and the designer would like to reduce the uncertainty of some output variables. To solve the aforementioned optimization problem with workflow reversal, the proposed process consists of three steps:

1. **Sensitivity analysis (optional):** The first step is to conduct a sensitivity analysis, so that the most influential sources of uncertainty can be selected for inverse propagation. This step is optional, as the designer can specify arbitrary input variables in the study. However, the most influential sources are preferred.
2. **Construction of an *Outer Workflow***: An outer workflow is constructed by integrating the original workflow with forward uncertainty propagation. The new inputs and outputs of the outer workflow now become the statistical moments of $\boldsymbol{x}$ and $\boldsymbol{y}$ as defined by:

$$\boldsymbol{\mathcal{M}_y} = F_{Outer}(\boldsymbol{\mathcal{M}_x}) \tag{5-7}$$

In the proposed method for inverse propagation, the outer workflow is constructed with URQ. Therefore the new inputs of the outer workflow include the first four statistical moments (means μ_{x_i}, standard deviations σ_{x_i}, skewness's γ_{x_i}, and kurtosis's Γ_{x_i}) of the original input variables, and the first two statistical moments (means μ_{y_j} and standard deviations σ_{y_j}) of the original output variables:

$$\boldsymbol{\mathcal{M}_x^{URQ}} = \begin{bmatrix} \mu_{x_1} & \sigma_{x_1} & \gamma_{x_1} & \Gamma_{x_1} \\ \mu_{x_2} & \sigma_{x_2} & \gamma_{x_2} & \Gamma_{x_2} \\ \mu_{x_3} & \sigma_{x_3} & \gamma_{x_3} & \Gamma_{x_3} \\ \cdots & \cdots & \cdots & \cdots \\ \mu_{x_n} & \sigma_{x_n} & \gamma_{x_n} & \Gamma_{x_n} \end{bmatrix} \tag{5-8}$$

$$\boldsymbol{\mathcal{M}_y^{URQ}} = \begin{bmatrix} \mu_{y_1} & \sigma_{y_1} \\ \mu_{y_2} & \sigma_{y_2} \\ \mu_{y_3} & \sigma_{y_3} \\ \cdots & \cdots \\ \mu_{y_l} & \sigma_{y_l} \end{bmatrix} \tag{5-9}$$

Without loss of generality, this is illustrated in Figure 5-3. In this case, the original workflow has two inputs and two outputs, while the outer workflow has eight inputs and four outputs.

3. **Reversal of the outer workflow**: The third step utilises AirCADia's workflow reversal capability to swap the inputs' standard deviations with those of the outputs. The reversed workflow is illustrated in Figure 5-4, where $\sigma_{y_1}, \sigma_{y_2}$ have now become the inputs and $\sigma_{x_1}, \sigma_{x_2}$ have been swapped to be the outputs. The designer can set values for $\sigma_{y_1}, \sigma_{y_2}$, and calculate $\sigma_{x_1}, \sigma_{x_2}$ as the results of the reversed workflow

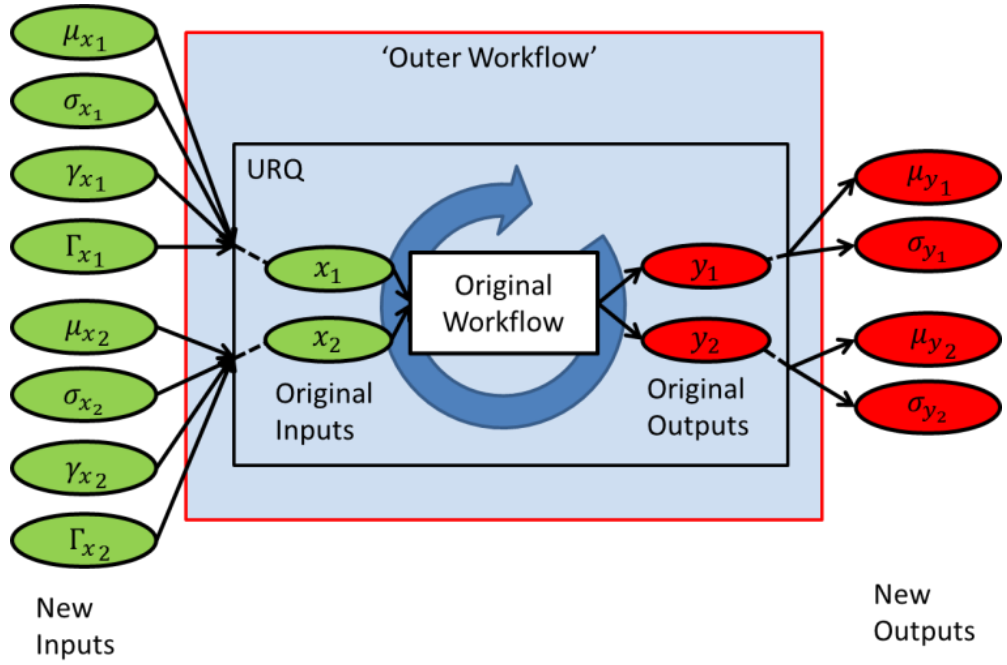

Figure 5-3 Example of an Outer Workflow

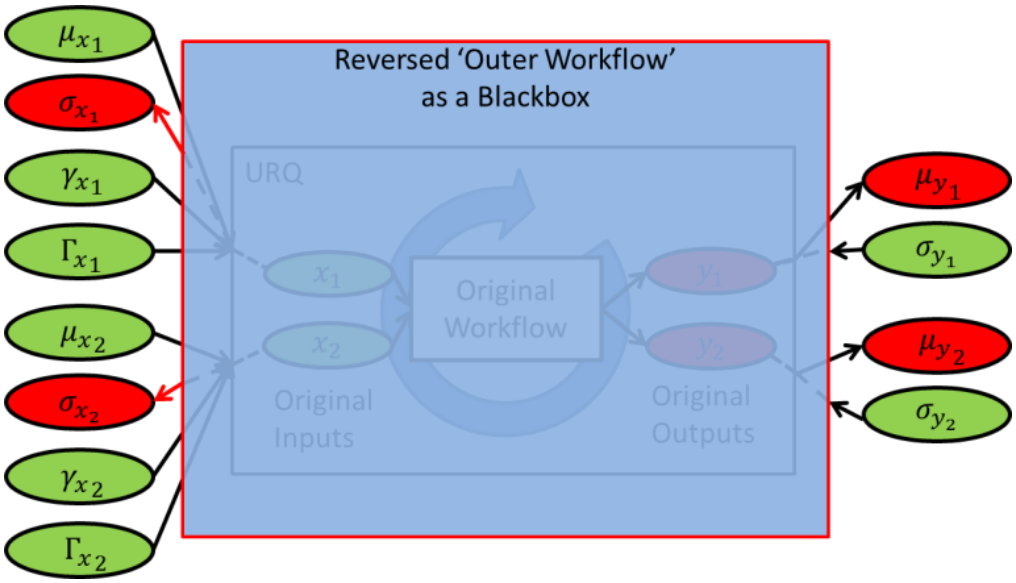

Figure 5-4 Reversal of an Outer Workflow

5.4 Validation

The validation is conducted via a series of numerical experiments. The proposed method is tested with progressively more complex test-cases, which includes two analytical examples and one practical design case study.

The analytical examples are proposed based on linear and non-linear algebraic equations. Given the target output variances, the theoretical input variances could be calculated and compared with the numerical results produced by the proposed method.

The practical design case study is based on an aircraft sizing problem, with considerable scale and complexity. In this case, theoretical values are not available, therefore the results of inverse propagation (input variances) are sent again to a forward propagation, to validate if the target output variances could be achieved.

5.4.1 Analytical Test-cases

The analytical test-cases are defined by:

$$\begin{cases} y_1 = x_1 + 2x_2 \\ y_2 = 3x_1 - 4x_2 \end{cases} \quad \textbf{(5-10)}$$

$$\begin{cases} y_3 = x_3^2 + e^{x_4} \\ y_4 = sin(x_3) + x_4 \end{cases} \quad \textbf{(5-11)}$$

In both cases the x's will be considered as the inputs and y's are taken as the outputs. The distributions of input variables are shown in Table 5-2.

Table 5-2. Distributions of the input variables

Variables	Distributions
x_1	Gaussian
x_2	Gaussian
x_3	Uniform
x_4	Uniform

For the first case, based on the linear combination of independent normal distributions, it could be analytically obtained that,

$$\sigma_{y_1} = \sqrt{\sigma_{x_1}^2 + 4\sigma_{x_2}^2} \quad \textbf{(5-12)}$$

$$\sigma_{y_2} = \sqrt{9\sigma_{x_1}^2 + 16\sigma_{x_2}^2} \quad \textbf{(5-13)}$$

Therefore, if the target output standard deviations are given, the theoretical input standard deviation (non-negative) for input variables could be calculated as:

$$\sigma_{x_1} = \sqrt{\frac{\sigma_{y_2}^2 - 4\sigma_{y_1}^2}{5}} \quad \textbf{(5-14)}$$

$$\sigma_{x_2} = \sqrt{\frac{9\sigma_{y_1}^2 - \sigma_{y_2}^2}{20}} \quad \textbf{(5-15)}$$

Here due to the square roots, σ_{y_1} and σ_{y_2} should follow:

$$0 < 2\sigma_{y_1} \leq \sigma_{y_2} \leq 3\sigma_{y_1} \quad \textbf{(5-16)}$$

For validation, a series of target σ_{y_1} and σ_{y_2} are given as shown in Table 5-3. The theoretical values of σ_{x_1} and σ_{x_2} (obtained from equation (5-14) and (5-15) are compared with the results from the proposed inverse propagation method. It should be mentioned that although in this case μ_{x_1} and μ_{x_2} do not influence the standard deviations, we need to set their values to execute the workflow. In this test, both of the values are set as 1.

Table 5-3. Results of the first analytical test case.

Output Standard Deviations (Target)		**Input Standard Deviations (Theoretical and from the Inverse Propagation)**			
σ_{y_1}	σ_{y_2}	σ_{x_1}: Theoretical	σ_{x_1}: Inverse Propagation	σ_{x_2}: Theoretical	σ_{x_2}: Inverse Propagation
0.05	0.12	0.0296648	0.0296666	0.0201246	0.0201254
0.7	1.7	0.4312772	0.4312772	0.2756810	0.2756810
1	2.5	0.6708204	0.6708204	0.3708099	0.3708099
2	5	1.3416408	1.3416407	0.7416198	0.7416199
3	8	2.3664319	2.3664319	0.9219544	0.9219544

From Table 5-3, it can be seen that the proposed method is able to provide considerable accuracy for such linear models.

It should be mentioned that the coefficients in equation (5-10) are selected arbitrarily. Given a general linear system defined as:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_l \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{33} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{l1} & a_{l2} & a_{l3} & \dots & a_{ln} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix} \quad \textbf{(5-17)}$$

If the input distributions are all normal, the standard deviations of outputs can be calculated as:

$$\begin{bmatrix} {\sigma_{y_1}}^2 \\ {\sigma_{y_2}}^2 \\ {\sigma_{y_3}}^2 \\ \dots \\ {\sigma_{y_l}}^2 \end{bmatrix} = \begin{bmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 & \dots & a_{1n}^2 \\ a_{21}^2 & a_{22}^2 & a_{33}^2 & \dots & a_{2n}^2 \\ a_{31}^2 & a_{32}^2 & a_{33}^2 & \dots & a_{3n}^2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{l1}^2 & a_{l2}^2 & a_{l3}^2 & \dots & a_{ln}^2 \end{bmatrix} \begin{bmatrix} {\sigma_{x_1}}^2 \\ {\sigma_{x_2}}^2 \\ {\sigma_{x_3}}^2 \\ \dots \\ {\sigma_{x_n}}^2 \end{bmatrix} \quad \textbf{(5-18)}$$

To ensure the existence of solution, the following condition must be fulfilled (Williams, 1978):

$$rank(\boldsymbol{\mathcal{A}}) = rank(\boldsymbol{\mathcal{A}}|\boldsymbol{b}), \quad \textbf{(5-19)}$$

where $\boldsymbol{\mathcal{A}}$ is the matrix of coefficients, and $\boldsymbol{\mathcal{A}}|\boldsymbol{b}$ is the augmented matrix:

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 & \dots & a_{1n}^2 \\ a_{21}^2 & a_{22}^2 & a_{33}^2 & \dots & a_{2n}^2 \\ a_{31}^2 & a_{32}^2 & a_{33}^2 & \dots & a_{3n}^2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{l1}^2 & a_{l2}^2 & a_{l3}^2 & \dots & a_{ln}^2 \end{bmatrix}, \quad \textbf{(5-20)}$$

$$\boldsymbol{b} = \left[{\sigma_{y_1}}^2, {\sigma_{y_2}}^2, {\sigma_{y_3}}^2, \dots, {\sigma_{y_l}}^2 \right]^T \quad \textbf{(5-21)}$$

If $rank(\boldsymbol{\mathcal{A}}) = n$, the solution will be unique, otherwise ($\boldsymbol{\mathcal{A}}$ is singular), there will be multiple solutions (Williams, 1978). In the latter case, the solver will converge to the solution which is closest to the initial values of ${\sigma_{x_i}}^2$'s (as discussed in Section 5.2.2).

It should also be noted that, as the solutions are all standard deviations, only non-negative values are of interest. If $\boldsymbol{\mathcal{A}}$ is square matrix ($n = l$), a sufficient condition for positive solution is given by (Kaykobad, 1985):

$$a_{ii}^2 > 0, i = 1,2,3, \dots, n \quad \textbf{(5-22)}$$

$$\sigma_{y_i}^2 > \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}^2 \frac{\sigma_{y_j}^2}{a_{jj}^2} \quad \textbf{(5-23)}$$

Further discussion of the problem is beyond the scope of the current research. In practice, the models maybe provided as black-boxes and the check of solution should be conducted as a post-process.

For the second test-case, the relationship between input and output standard deviations is given by:

$$\sigma_{y_3}^2 = \frac{\left(\mu_{x_3} + \sqrt{3}\sigma_{x_3}\right)^5 - \left(\mu_{x_3} - \sqrt{3}\sigma_{x_3}\right)^5}{10\sqrt{3}\sigma_{x_3}} - \frac{\left[\left(\mu_{x_3} + \sqrt{3}\sigma_{x_3}\right)^3 - \left(\mu_{x_3} - \sqrt{3}\sigma_{x_3}\right)^3\right]^2}{108\sigma_{x_3}^2} + \frac{e^{2\mu_{x_4}}\left(e^{2\sqrt{3}\sigma_{x_4}} - \frac{1}{e^{2\sqrt{3}\sigma_{x_4}}}\right)}{4\sqrt{3}\sigma_{x_4}} - \frac{e^{2\mu_{x_4}}\left(e^{\sqrt{3}\sigma_{x_4}} - \frac{1}{e^{\sqrt{3}\sigma_{x_4}}}\right)^2}{12\sigma_{x_4}^2} \quad \textbf{(5-24)}$$

$$\sigma_{y_4}^2 = \frac{\sin\left(2\mu_{x_3} - 2\sqrt{3}\sigma_{x_3}\right) - \sin\left(2\mu_{x_3} + 2\sqrt{3}\sigma_{x_3}\right)}{8\sqrt{3}\sigma_{x_3}} + \frac{1}{2} - \frac{\left[\cos\left(\mu_{x_3} - \sqrt{3}\sigma_{x_3}\right) - \cos\left(\mu_{x_3} + \sqrt{3}\sigma_{x_3}\right)\right]^2}{12\sigma_{x_3}^2} + \sigma_{x_4}^2 \quad \textbf{(5-25)}$$

In this case, it is difficult to represent the input standard deviations (σ_{x_3} and σ_{x_4}) as functions of the output ones (σ_{y_3} and σ_{y_3}). The validation will be conducted by firstly setting σ_{x_3} and σ_{x_4} to calculate σ_{y_3} and σ_{y_4} from equation (5-24) and (5-25). Then, these values of σ_{y_3} and σ_{y_4} will be considered as targets for inverse propagation. Ideally the same σ_{x_3} and σ_{x_4} should be achieved. In this case, μ_{x_3} and μ_{x_4} will also have an impact on the values of σ_{y_3} and σ_{y_4}. The settings are shown in Table 5-4.

Table 5-4. Results of the second analytical test case.

Mean		**Output Standard Deviations (Target)**		**Input Standard Deviations (Theoretical and from the Inverse Propagation)**			
μ_{x_3}	μ_{x_4}	σ_{y_3}	σ_{y_4}	σ_{x_3}: Theoretical	σ_{x_3}: Inverse Propagation	σ_{x_4}: Theoretical	σ_{x_4}: Inverse Propagation
1	0.7	1.3365	0.4803	0.5	0.5006	0.4	0.3996
1.2	1	1.1466	0.2515	0.4	0.4002	0.2	0.1998

1.4	1.2	2.7179	0.5461	0.7	0.7057	0.5	0.4963
1.6	0.8	1.7575	0.3184	0.5	0.5002	0.3	0.2994
2	1.5	3.6073	0.5756	0.65	0.6536	0.5	0.4979
1.7	3	17.1104	0.7188	0.6	0.5618	0.7	0.7032
1.3	1.6	3.3117	0.5587	0.7	0.7653	0.5	0.4793
1.7	2	6.3964	0.7051	0.4	Not Converged	0.7	Not Converged

As could be seen from Table 5-4, while most of the settings produce satisfactory results, the last three groups show relatively low accuracy. For the last one, the solver did not converge to the theoretical value. By running a sensitivity analysis, it is noted that for these three settings, the output uncertainty is very sensitive to σ_{x_4} due to the relatively large value of μ_{x_4}. This makes it more difficult for the solver to find a solution. Future work will involve determining the limitations of the solver and proposing adequate solvers for highly non-linear problems.

5.4.2 Realistic Design Case Study

To demonstrate a practical application, we applied the method to an aircraft sizing code, USMAC (Ultra Simplified Model of Aircraft), which was provided by an industrial partner, in the context of the European project, *Value Improvement through a Virtual Aeronautical Collaborative Enterprise*[2] (VIVACE). In this code, aircraft performances are calculated based on the top level requirements and geometry specification defined by the designer.

The following scenario is used to demonstrate the inverse uncertainty propagation. Among the performance outputs, attention has been focused on the range R and Maximum Take-Off Weight ($MTOW$), where the requirements are given as:

$$R \geq 9500\ km \quad \textbf{(5-26)}$$

$$MTOW \leq 104460\ kg$$

Five sources of uncertainty are considered as shown in Table 5-5. The first two are model uncertainties, regarding the estimation of drag coefficient C_D and Specific Fuel Consumption (SFC). The method to account for model discrepancy is explained in

[2] Project webpage: https://cordis.europa.eu/result/rcn/47814_en.html

Section 2.4, where the original outputs are randomised with factors (rv_{C_D} and rv_{SFC}). Here Gaussian and triangular distributions are chosen for these random factors. Temperature at cruise (T_{crz}) is considered as the external source of uncertainty, which is modelled as a Gaussian distribution. Uncertainties associated with the two design variables: Bypass Ratio (BPR) and Sea Level Static Thrust ($FNslst$) are caused by lack of definition at early stage of the design. For these two uncertainties, uniform distributions are used which represent the ranges of reasonable values. For all these distributions, the parameters are arbitrarily assigned. It should be emphasised that the purpose of this study is to validate the proposed method, rather than conduct an investigation into the merits of the particular aircraft design. In a real design scenario, the distributions could be based on historical data or expert elicitations.

Table 5-5. Sources of Uncertainty

Source	**Type**	**Distribution**	**Standard Deviation**
Drag Model	Model Uncertainty	Normal $\mu_{rv_{C_D}} = 1$ $\sigma_{rv_{C_D}} = 0.05$	0.05
SFC Model	Model Uncertainty	Triangular $LB_{rv_{SFC}} = 0.97$ $UB_{rv_{SFC}} = 1.03$ $MV_{rv_{SFC}} = 1$	0.0122
Temperature at Cruise (K)	Input Uncertainty	Normal $\mu_{T_{crz}} = 218$ $\sigma_{T_{crz}} = 10.9$	10.9
Bypass Ratio	Input Uncertainty	Uniform $LB_{BPR} = 7.5$ $UB_{BPR} = 8.5$	0.2886
Sea level static thrust (N)	Input Uncertainty	Uniform $LB_{FNslst} = 123500$ $UB_{FNslst} = 136500$	3752.7767

A Monte Carlo simulation (MCS) was conducted a priori to propagate uncertainty from the five sources. The results are shown in Figure 5-5 and Figure 5-6, where the mean values μ_R and μ_{MTOW} are indicated by red lines and the constraints defined in equation (5-26) are showed by the black ones. Based on the ratio of sampling points meeting the requirements, over the total sampling points of the MCS, the inferred probabilities of meeting these constraints are 84.36% and 83.89%, respectively.

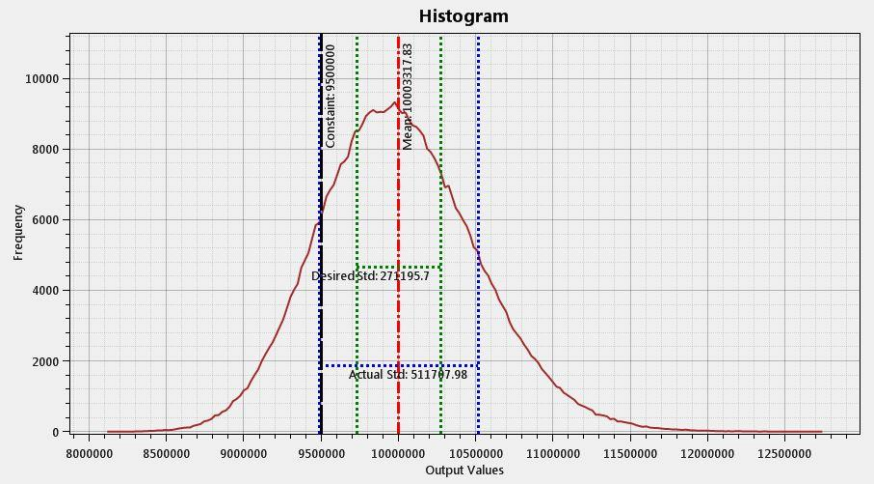

Figure 5-5. Histogram of the Range (m)

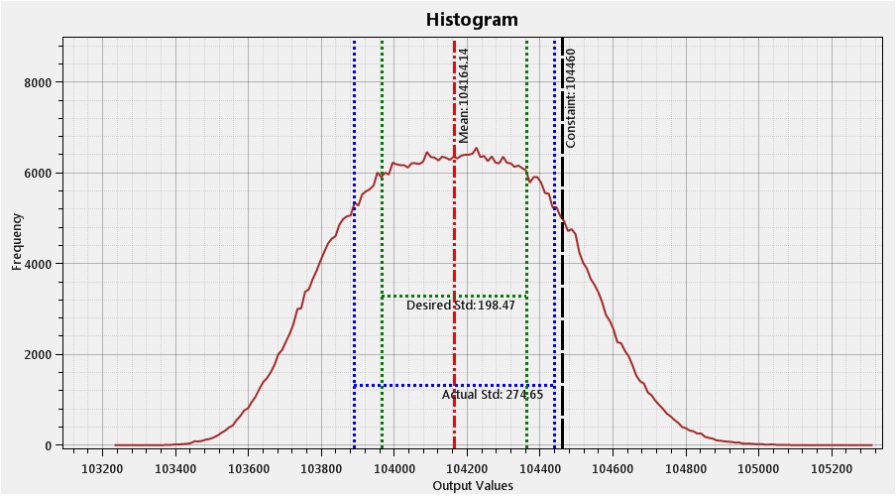

Figure 5-6. Histogram of the maximum take-off weight (kg)

The inverse uncertainty propagation is motivated, for example, if the designer wished to increase both probabilities of constraint satisfaction to 90%. Therefore the distributions of these two output variables should be 'narrowed' and the problem becomes how much uncertainty is allowed from the five sources listed in Table 5-5.

The first step is to construct the outer workflow. In this case, 20 input variables and 4 output variables will be automatically generated in the software, which are illustrated as green and red ovals in Figure 5-7.

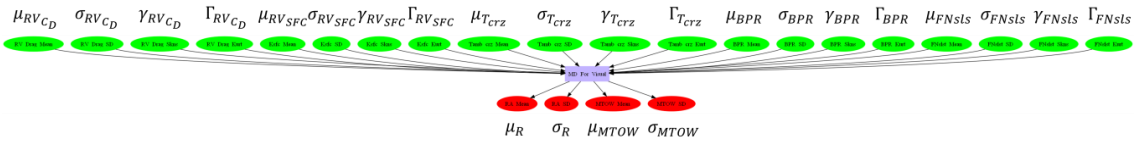

Figure 5-7. Outer workflow for USMAC in AirCADia

The second step is to set the targets of output standard deviations. The method (Padulo and Guenov, 2011) discussed in Section 2.5.2 is used to quickly approximate these standard deviations based on probabilities, in which:

$$\mathbb{P}\left\{\left|y_i - \mu_{y_i}\right| \geq k_{P_{y_i}}\sigma_{y_i}\right\} \leq 1 - \mathbb{p}_{y_i} \tag{5-27}$$

Here y_i is a stochastic output which could be R or $MTOW$ in this case. $\mathbb{p}_{y_i}$ is a function of $k_{P_{y_i}}$ and the detailed formulation is according to the shape of the distribution of y_i.

According to Table 2-2, for unimodality distributions (in this case, the distribution of R):

$$\mathbb{p}_R = \frac{9k_{P_R}^2 + 5}{9k_{P_R}^2 + 9} \tag{5-28}$$

Therefore, for $\mathbb{p}_R = 90\%$

$$k_{P_R} = 1.8559 \tag{5-29}$$

$$\sigma_R \leq \frac{\mu_R - 9500}{k_{P_R}} = 271.1957 = \sigma_R^* \tag{5-30}$$

For unimodality and symmetry distributions (In this case, the distribution of $MTOW$):

$$\mathbb{p}_{MTOW} = 1 - \frac{2}{9k_{P_{MTOW}}^2} \tag{5-31}$$

Therefore, for $\mathbb{p}_{MTOW} = 90\%$, $k_{P_{MTOW}} = 1.4907$

$$\sigma_{MTOW} \leq \frac{104460 - \mu_{MTOW}}{k_{P_{MTOW}}} = 198.4718 = \sigma_{MTOW}^* \tag{5-32}$$

In Figure 5-5 and Figure 5-6, the original standard deviations and the target ones are indicated by the blue and green lines respectively, and the values are summarised in Table 5-6.

Table 5-6. Original and target standard deviations of the outputs

Standard Deviations	Original Value	Target Value
σ_R (km)	511.7080	271.1957
σ_{MTOW} (kg)	274.6514	198.4718

Before reversing the workflow, two of the five input standard deviations are chosen, based on the results of sensitivity analysis. In this particular case Fourier Amplitude Sensitivity Test (FAST) (Cukier, 1973) is utilised and the results are plotted in Figure 5-8.

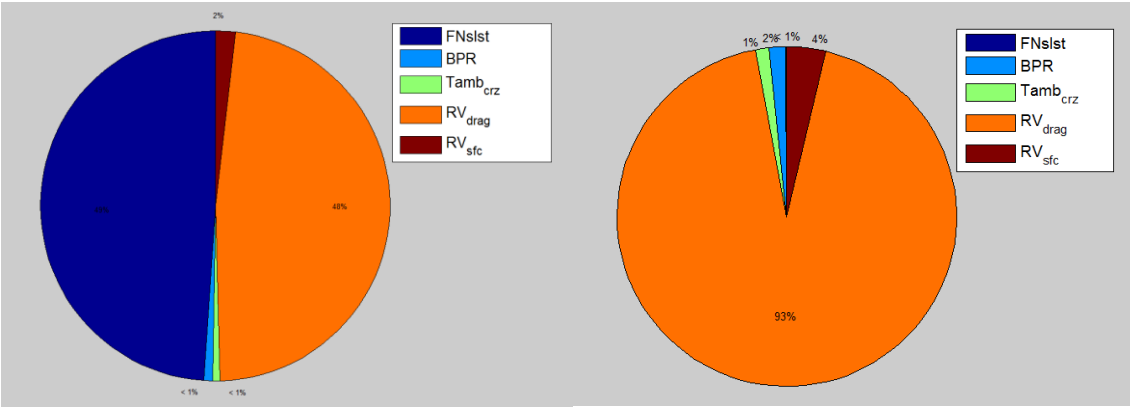

Figure 5-8. First order sensitivity indices of different inputs for $MTOW$ (left) and R (right)

From the figure, it can be seen that RV_{drag} and $FNslst$ are the main contributors to the variances of selected outputs. The outer workflow is then reversed by switching σ_{MTOW} and σ_R with σ_{FNslst} and $\sigma_{RV_{C_D}}$. The reversed workflow is shown by Figure 5-9, where the switched input/output variables are indicated by red dashed-line blocks.

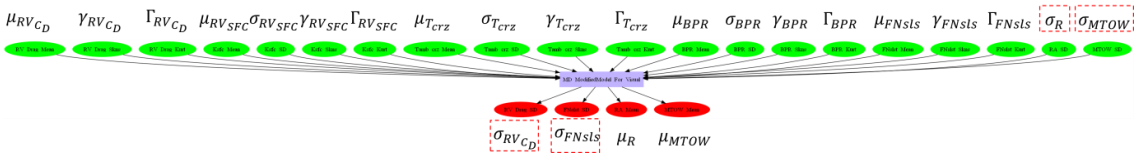

Figure 5-9. Reverse workflow in AirCADia

The target standard deviations in Table 5-6 are now given to the reversed workflow, which calculates $\sigma^*_{RV_{C_D}}$ and σ^*_{FNslst} as 0.0236 and 2924.5705, respectively. Using these values, the updated distributions of RV_{C_D} and $FNslst$ are shown in Figure 5-10.

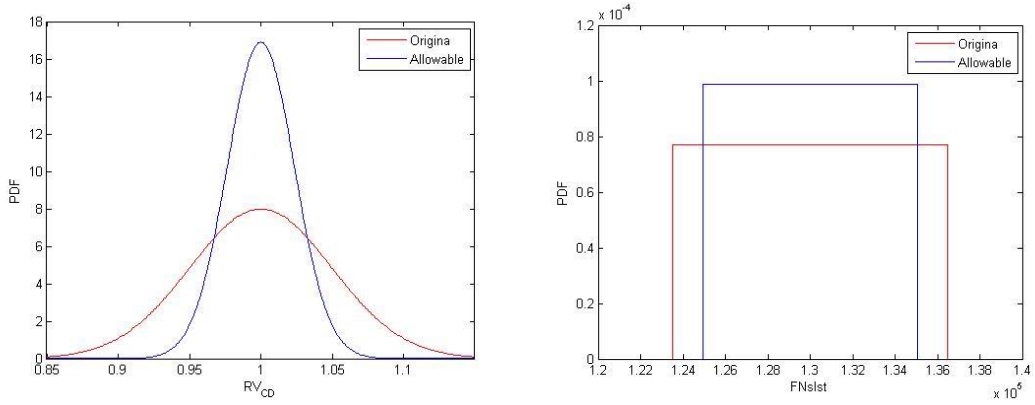

Figure 5-10. Original and allowable distributions of RV_{C_D} and $FNslst$

For validation, the new values of $\sigma_{FNslst}{}'$ and $\sigma_{RV_{C_D}}{}'$ are put back into Monte-Carlo forward propagation (1000000 runs) to check if the desired output uncertainty can be achieved. The results are shown in Table 5-7, which demonstrated the accuracy of the proposed method.

Table 5-7. Validation of the results

Output Variables	**Target *std***	**Achieved *std***	**Error**
$MTOW\ (kg)$	198.4718	198.6092	0.0692%
$R\ (km)$	271.1957	271.4027	0.0763%

5.5 Summary and Conclusions

Presented in this section is a novel method for inversed propagation of uncertainty. It incorporates enablers for the reversal of computational workflows and for the efficient propagation of uncertainty. This contributes towards the identification and allocation of desired variance to design inputs. The methodology is illustrated with representative analytical and numerical examples. The results have demonstrated the accuracy of the method.

Mathematically, the objective function based on actual and target statistical moments (defined in equation (5-6)) can be considered as equivalent to the ones presented in (Congedo et al., 2012; Zabaras and Ganapathysubramanian, 2008) (defined in equations

(3-13) and (3-16), respectively). However compared with the latter, several advantages can be achieved with the proposed method:

- The first is the simplicity of implementation. With the proposed method, the inverse propagation can be constructed on-the-fly, for various computational workflows, and various setup of the problem. The designer does not need to be heavily involved with the numerical setup of the problem and the explorations can be conducted interactively.
- The workflow reversal capability has integrated several techniques to improve the convergence of the inverse propagation. For instance, the strategies to select between a global and local reversal, the strategies to select from several solvers, and the ability to assess if a reversal is feasible *a priori*.
- The computational cost of URQ is significantly lower than the numerical treatment in the existing methods, which will leads to a reduction in the overall computational cost.

Future work includes four aspects:

- The first is regarding the robustness of the workflow reversal capability. In the second analytical test-case as presented in Section 5.4.1, some of the setups caused a failure of the numerical solver. Further research is needed to investigate the solver's numerical limitations.
- The second is to perform a more rigorous study to compare the proposed method with existing techniques, regarding the numerical accuracy, computational cost, and convergence speed.
- The third is to implement this approach with other forward uncertainty propagation techniques, apart from the URQ. Currently, the method considers only the standard deviation for inverse propagation. With the generalised Taguchi method (Seo and Kwak, 2002), for instance, it is feasible to inversely propagate the skewness and kurtosis as well.
- Last but not least is to investigate further the scenario of many-to-many reversals (regarding the broader picture of uncertainty allocation). Specifically, the following two cases can be considered:

- **Underdetermined Case**: There are multiple combinations of the input standard deviations, which can lead to the same output ones. In this case, additional constraints should be added, for instance, by specifying a proportion between the reductions of different input standard deviations. Alternatively, a method has been proposed by Molina-Cristóbal to explore a set of solutions and conduct trade-off studies (Molina-Cristóbal et al., 2018).
- **Overdetermined Case**: There are no feasible combinations of the input standard deviations, which can exactly lead to the target output ones. If the problem is overdetermined, the designer would like to be as close as possible to the target values. The values of $\left|\langle y_i^j \rangle - \langle y_i^j \rangle^*\right|$'s are to be minimised but there is no tolerance to be specified (or a gradually increased tolerance following a try-and-error process). The solver will stop according to the pre-defined limits on computing resources (e.g. number of interactions or optimization time).

6 Efficient Method for Variance Based Sensitivity Analysis

6.1 Introduction

This chapter presents a method for variance based sensitivity analysis. According to the literature review, most of the current techniques are computational expensive, especially for large scale design problems and high fidelity simulations. Although some improved techniques are proposed, there are still different ways to address the problem, which have not been investigated.

In this research, we propose a general approach to approximate the sensitivity indices based on the formulation from Saltelli (Saltelli et al., 2008). As been explained in Section 3.4.3, this formulation is not suitable for the traditional MSC approach, due to the nested loops in the formulation. However by using approximation techniques, such nested loops become computationally affordable. In particular, we propose one implementation with the URQ method as presented in Section 5.2.1. Compared with traditional quadrature techniques, in which the computational cost grows exponentially with the dimensions, the URQ method requires a much lower number of sampling points. Thus it is especially suitable for analysis of large scale problems.

The remaining part of this chapter is structured as follows. Section 6.2 reviews the formulations of variance based sensitivity analysis. In Section 6.3, the general approach for approximation is presented, followed by the detailed implementation based on the URQ method. The proposed method is evaluated in Section 6.4, using a number of test-cases and is compared to the traditional (benchmark) MCS approach. Finally, conclusions and future work are presented in Section 6.5.

6.2 Background

The Variance Base Sensitivity Indices can be classified as first order indices S_i, higher order indices S_{ij}, and total effect indices S_i^T. The deduction of the first order indices has been explained in Section 3.4.2, based on Eq. (3-19) to (3-21). In this section, the final formulation is repeated for the reader's convenience. Similar to the Section 3.4.2, $y = f(\boldsymbol{x})$ is considered to be the only output of the function. In the general case of multiple outputs, the following process shall be repeated for each variable in the vector

$\boldsymbol{y}$, and correspondingly, different sets of Sobol' indices (S_i, S_{ij}, and S_i^T) will be obtained for different outputs. As mentioned above, this situation will not be demonstrated in this thesis, therefore no additional notations will be introduced to represent Sobol' indices for multi-output functions.

6.2.1 First-order Indices

The first-order index accounts for the portion of variance caused by uncertainty from only one of the input variables. For instance, regarding the i_{th} input variable x_i, its corresponding first-order Sobol' index can be defined as:

$$S_i = \frac{V(y) - E_{X_i}\left(V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)\right)}{V(y)} \quad \textbf{(6-1)}$$

By further expansion and derivation, equation (6-1) could be reformulated to the following forms:

$$S_i = \frac{E_{X_i}(E^2_{x_{\sim i}}(y|_{x_i=X_i})) - E^2(y)}{V(y)} \quad \textbf{(6-2)}$$

$$S_i = \frac{V_{X_i}(E_{x_{\sim i}}(y|_{x_i=X_i}))}{V(y)} \quad \textbf{(6-3)}$$

$$S_i = \frac{E\left(f_+(\boldsymbol{x_{+i}})\right) - E^2(y)}{V(y)} \quad \textbf{(6-4)}$$

The reader is referred to (Saltelli et al., 2004, 2008) for more details on the derivation of equation (6-3), and to (Ishigami and Homma, 1990; Saltelli, 2002) for the derivation of equations (6-2) and (6-4).

It should be noted that in equation (6-4), the problem is converted into a single loop expectation of the new function $f_+(\boldsymbol{x_{+i}})$, which is defined as:

$$f_+(\boldsymbol{x_{+i}}) = f(x_1, x_2, \dots, x_i, \dots x_n) \cdot f(x_1', x_2', \dots x_{i-1}', x_i, x_{i+1}', \dots x_n'), \quad \textbf{(6-5)}$$

where $\boldsymbol{x_{+i}}$ is the new input vector, which consists of $2n-1$ variables. In this vector, x_r and $x_r{}'$ are considered as independent variables (for each $r = 1,2, \dots, n; r \neq i$) but with the same PDF. Also note that there is no $x_i{}'$ in vector, $\boldsymbol{x_{+i}}$:

$$\boldsymbol{x_{+i}} = [x_1, x_2, \dots, x_i, \dots x_n, x_1', x_2', \dots x_{i-1}', x_{i+1}', \dots x_n'] \quad \textbf{(6-6)}$$

6.2.2 Second-order Indices

The interaction effect refers to the portion of variance caused by particular combinations of the input variables (Saltelli et al., 2004, 2008). This effect is captured by the high order index. For instance, the second order index S_{ij} refers to the interaction effect between the i_{th} and the j_{th} input variables. Note that this interaction effect refers to a portion in the output variance, while x_i and x_j are still independent as inputs. In this research, only the second order indices are considered, but the same principle could also be applied to calculate other higher order indices.

Similar to the first order indices, S_{ij} could be calculated by solving the expectation of conditional variance with regard to two input variables. It can be proven that this formulation also includes the first order effects (Saltelli et al., 2004, 2008), therefore the first order indices need to be subtracted. The final equation becomes:

$$S_{ij} = \frac{V(y) - E_{X_{i,j}}\left(V_{x_{\sim i,j}}\left(y|_{x_i=X_i, x_j=X_j}\right)\right)}{V(y)} - S_i - S_j \quad \textbf{(6-7)}$$

Some alternatives formulations (Saltelli et al., 2004, 2008) include,

$$S_{ij} = \frac{E_{X_{i,j}}\left(E^2_{x_{\sim i,j}}\left(y|_{x_i=X_i, x_j=X_j}\right)\right) - E^2(y)}{V(y)} - S_i - S_j \quad \textbf{(6-8)}$$

$$S_{ij} = \frac{V_{X_{i,j}}(E_{x_{\sim i,j}}(y|_{x_i=X_i, x_j=X_j}))}{V(y)} - S_i - S_j \quad \textbf{(6-9)}$$

$$S_{ij} = \frac{E\left(f_{++}(\boldsymbol{x_{++ij}})\right) - E^2(y)}{V(y)} - S_i - S_j \quad \textbf{(6-10)}$$

Using similar reasoning as applied to equations (6-4) - (6-6), $f_{++}(x_{++ij})$ is defined by multiplying the original function $f(x)$ with itself, taking two different sets of independent inputs, but this time sharing the same x_i and x_j in both sets.

$$f_{++}(\boldsymbol{x_{++ij}}) = f(x_1, x_2, \dots, x_i, \dots, x_j \dots x_n) \cdot f(x_1', x_2', \dots, x_{i-1}', x_i, x_{i+1}', \dots, x_{j-1}', x_j, x_{j+1}', \dots, x_n') \quad \textbf{(6-11)}$$

Here x_{++ij} is the corresponding input vector, consists of $2n-2$ variables. In this vector, x_r and $x_r{}'$ are considered as independent variables (for each $r = 1,2,\dots,n;\ r \neq i,j$), but with the same PDF. However there is no $x_i{}'$ and $x_j{}'$ in this vector.

$$\boldsymbol{x_{++ij}} = [x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n, x_1{}', x_2{}', \dots, x_{i-1}{}', x_{i+1}{}', \dots, x_{j-1}{}', x_{j+1}{}', \dots, x_n{}'] \quad \textbf{(6-12)}$$

6.2.3 Total Effect Indices

The total effect index accounts for the variable's first order effect and all its interactions with other variables (Saltelli et al., 2004, 2008). That is,

$$S_i^T = S_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} S_{ij} + \sum_{\substack{j,r=1 \\ r\neq j\neq i}}^{n} S_{ijr} + \cdots \quad \textbf{(6-13)}$$

Apart from by adding all the indices as shown in equation (6-13), which may become impractical when the number of inputs is high, this index is more widely calculated by using a nested structure as well:

$$S_i^T = \frac{E_{\boldsymbol{X}_{\sim i}}\left(V_{x_i}\left(y|_{\boldsymbol{x}_{\sim i}=\boldsymbol{X}_{\sim i}}\right)\right)}{V(y)}, \quad \textbf{(6-14)}$$

where all variables except x_i are firstly fixed for the calculation of the conditional variance, and then are varied in the expectation loop. The reader is referred to (Saltelli et al., 2004, 2008) for more rigorous mathematical derivation. By expansion and further deduction, equation (6-14) could be transferred as the following alternatives:

$$S_i^T = \frac{V(y) + E^2(y) - E_{\boldsymbol{X}_{\sim i}}(E_{x_i}^2(y|_{\boldsymbol{x}_{\sim i}=\boldsymbol{X}_{\sim i}}))}{V(y)} \quad \textbf{(6-15)}$$

$$S_i^T = \frac{V(y) - V_{\boldsymbol{X}_{\sim i}}(E_{x_i}(y|_{\boldsymbol{x}_{\sim i}=\boldsymbol{X}_{\sim i}}))}{V(y)} \quad \textbf{(6-16)}$$

$$S_i^T = \frac{V(y) + E^2(y) - E(f_{-}(\boldsymbol{x}_{-i}))}{V(y)} \quad \textbf{(6-17)}$$

Again, the same reasoning as applied to equations (6-4) - (6-6), $f_{-}(\boldsymbol{x}_{-i})$ is defined by multiplying the original function $f(\boldsymbol{x})$ with itself. This time all the inputs are the same except x_i and $x_i{}'$.

$$f_{-}(\boldsymbol{x}_{-i}) = f(x_1, x_2, \dots, x_i, \dots x_n) \cdot f(x_1, x_2, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_n) \quad \textbf{(6-18)}$$

Therefore $\boldsymbol{x}_{-i}$ consists of only $n+1$ variables. In this vector, x_i and x_i' are considered as independent variables, but with the same PDF.

$$x_{-i} = [x_1, x_2, \dots, x_i, \dots, x_n, x_i'] \quad \textbf{(6-19)}$$

6.3 Proposed Method

6.3.1 General Approach

The formulations reviewed in Section 6.2 are summarized in Table 6-1, where the equations are categorised into four options. In the traditional MCS approach (Ishigami and Homma, 1990; Sobol', 1993), only Option 4 is adopted, because the other three options (1, 2, and 3) are computationally too expensive for MSC, due to the nested integrals in the formulations.

Table 6-1. Equations used in four options to implement the proposed approach

Index	**Option 1: Nested Expectation of Variance**	**Option 2: Nested Expectations**	**Option 3: Nested Variance of Expectation**	**Option 4: Single Loop Expectation**
First Order: S_i	Eq. (6-1)	Eq. (6-2)	Eq. (6-3)	Eq. (6-4)
Second Order: S_{ij}	Eq. (6-7)	Eq. (6-8)	Eq. (6-9)	Eq. (6-10)
Total Effect: S_i^T	Eq. (6-14)	Eq. (6-15)	Eq. (6-16)	Eq. (6-17)

The rationale of the proposed approach is that, since the nature of these integrals is to solve nested expectations/variances, the calculation can be transformed into a standard uncertainty propagation process. As there are plenty of more efficient uncertainty propagation techniques compared with MCS, Option 1, 2, and 3, may become computationally affordable. The critical part is to construct such nested loops of propagation, which is illustrated in Figure 6-1, by taking S_i in Option 1 as an example.

The uncertainty propagation is firstly applied to calculate the variance of the original model, as indicated by the number (1) in Figure 6-1. In this process, the i_{th} variable is temporarily fixed, thus the calculated variance is the conditional one as defined by equation (3-19).

Then this propagation process itself is taken as a 'new model', with x_i as its input, and the conditional variances as the outputs. Another propagation loop is conducted on top of this 'new model', regarding to the uncertainty of x_i, as indicated by the number (2) in Figure 6-1. This calculation provides the mean (expectation) of the conditional variances, which could be used to calculate the indices as defined by Eq. (6-1). For Option 2 and 3, similar approaches could be applied by modifying the sequence of calculating the means and variances.

Option 4, as discussed previously, does not require any nested loops. The approach is to apply single-loop uncertainty propagation for the mean of $f_+(x_{+i})$ defined by Eq. (6-5).

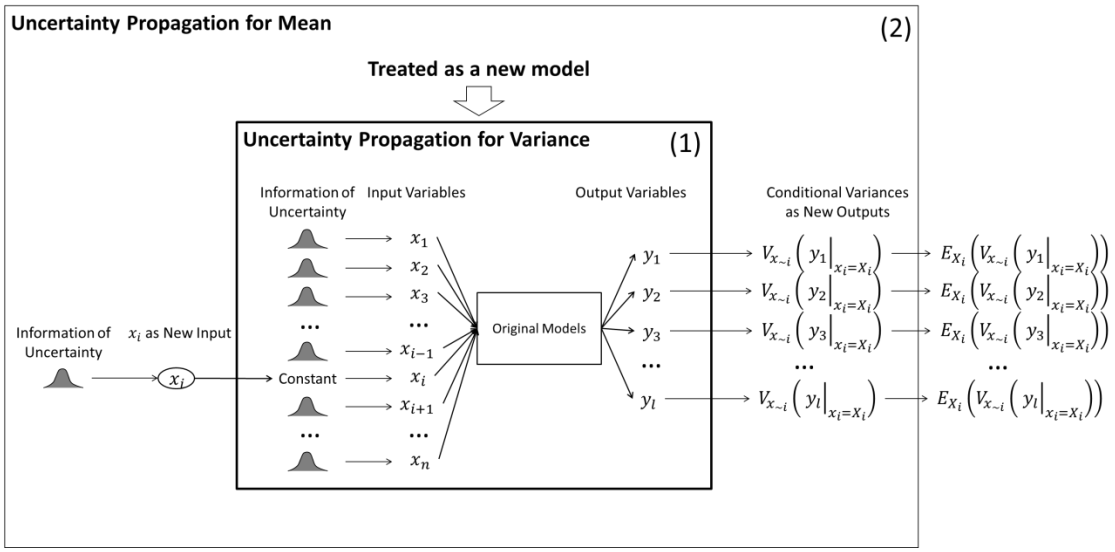

Figure 6-1. The general process, Illustrated with First Order Index Using Option 1

6.3.2 Formulation with URQ

Following the general approach, the specific formulation with URQ is deduced in this section, regarding the first order, second order, and total effect indices. To avoid repetition, only the derivation of Option 1 will be explained in detail. However all four options are implemented and compared in Section 6.4. The reader is referred to Appendix A.1 for the equations of Option 2, 3, and 4.

6.3.2.1 First-order Indices

Given equation (6-1), as the computation of $V(y)$ via URQ has already been specified in equation (5-2), the following section will focus on the calculation of $E_{X_i}\left(V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)\right)$.

To calculate the inner loop $V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)$, firstly x_i is temporarily fixed as a constant X_i. Compared to the original function $y=f(x)$ which has n input variables, a new function $f_1^{(\sim i)}$ can be defined, with $n-1$ input variables.

$$f_1^{(\sim i)}(x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = y|_{x_i=X_i} = f(x_1, x_2, \dots, x_{i-1}, X_i, x_{i+1}, \dots, x_n) \quad \textbf{(6-20)}$$

Equation (5-2) from URQ could be used to calculate the variance of this new function $f_1^{(\sim i)}$.

$$V_{x_{\sim i}}\left(y|_{x_i=X_i}\right) = V\left(f_1^{(\sim i)}\right) \quad \textbf{(6-21)}$$

$$= \sum_{\substack{r=1 \\ r\neq i}}^{n} \left\{ W_r^+ \left[\frac{f_1^{(\sim i)}(x_r^{+(\sim i)}) - f_1^{(\sim i)}\left(\mu_x^{(\sim i)}\right)}{h_r^+}\right]^2 + W_r^- \left[\frac{f_1^{(\sim i)}(x_r^{-(\sim i)}) - f_1^{(\sim i)}\left(\mu_x^{(\sim i)}\right)}{h_r^-}\right]^2 + W_r^{\pm} \frac{\left[f_1^{(\sim i)}(x_r^{+(\sim i)}) - f_1^{(\sim i)}\left(\mu_x^{(\sim i)}\right)\right]\left[f_1^{(\sim i)}(x_r^{-(\sim i)}) - f_1^{(\sim i)}\left(\mu_x^{(\sim i)}\right)\right]}{h_r^+ h_r^-} \right\},$$

where W_r^+, W_r^-, $W_r^{\pm}$, h_r^+, and h_r^- are defined by the original URQ method (as shown in Table 5-1), while $\mu_x^{(\sim i)}$, $x_r^{+(\sim i)}$, and $x_r^{-(\sim i)}$ are defined by removing the i_{th} variable from the original input vectors, as shown in the following equations

$$\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i)} = [\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_{i-1}}, \mu_{x_{i+1}}, \dots, \mu_{x_n}] \tag{6-22}$$

$$\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i)^+} = [\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_r} + h_r^+ \sigma_{x_r}, \dots, \mu_{x_{i-1}}, \mu_{x_{i+1}}, \dots, \mu_{x_n}], \qquad r \neq i \tag{6-23}$$

$$\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i)^-} = [\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_r} + h_r^- \sigma_{x_r}, \dots, \mu_{x_{i-1}}, \mu_{x_{i+1}}, \dots, \mu_{x_n}], \qquad r \neq i \tag{6-24}$$

Recalling the definition of $f_1^{(\sim i)}$, since the value of X_i is still not specified, equation (6-21) now becomes a function of X_i, which could be defined as,

$$g_1(X_i) = V_{x_{\sim i}}\left(y|_{x_i = X_i} \right) \tag{6-25}$$

Then, the expected value of $g_1(X_i)$ can be calculated employing Equation (5-1) from URQ method,

$$\begin{aligned} E_{X_i}\left(V_{x_{\sim i}}\left(y|_{x_i = X_i} \right)\right) &= E_{X_i}\left(g_1(X_i)\right) \\ &= W_0^{(i)} g_1\left(\mu_{x_i}\right) + W_i \left[\frac{g_1(\mu_{x_i} + h_i^+ \sigma_{x_i})}{h_i^+} - \frac{g_1(\mu_{x_i} + h_i^- \sigma_{x_i})}{h_i^-}\right], \end{aligned} \tag{6-26}$$

where W_i, h_i^+, and h_i^- are defined in Table 5-1 and $W_0^{(i)}$ is defined as:

$$W_0^{(i)} = 1 + \frac{1}{h_i^+ h_i^-} \tag{6-27}$$

Substituting Equation (6-26) and (5-2) to Equation (6-1), the first order Sobol' index is obtained as:

$$S_i \approx \frac{V_{URQ}(y) - W_0^{(i)} g_1\left(\mu_{x_i}\right) + W_i \left[\frac{g_1(\mu_{x_i} + h_i^+ \sigma_{x_i})}{h_i^+} - \frac{g_1(\mu_{x_i} + h_i^- \sigma_{x_i})}{h_i^-}\right]}{V_{URQ}(y)} \tag{6-28}$$

In this equation, $V_{URQ}(y)$ is the total variance calculated by equation (5-2).

6.3.2.2 Second Order Indices

Given equation (6-7), the objective is to calculate $E_{X_{i,j}}\left(V_{x_{\sim i,j}}\left(y|_{x_i = X_i, x_j = X_j} \right)\right)$. The process is similar to that presented in Section 6.3.2.1. Since two variables are now

involved in the nested loop, function $f_2^{(\sim i,j)}$ could be defined by fixing the i_{th} and j_{th} input variables.

$$f_2^{(\sim i,j)}(x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_{j-1}, x_{j+1}, \dots, x_n) = y|_{x_i=X_i, x_j=X_j} = f(x_1, x_2, \dots, x_{i-1}, X_i, x_{i+1}, \dots, x_{j-1}, X_j, x_{j+1}, \dots, x_n) \quad \textbf{(6-29)}$$

By calling equation (5-2), the inner loop variance $V_{x_{\sim i,j}}\left(y|_{x_i=X_i, x_j=X_j}\right)$ is calculated by,

$$V_{x_{\sim i,j}}\left(y|_{x_i=X_i, x_j=X_j}\right) = V\left(f_2^{(\sim i,j)}\right) \quad \textbf{(6-30)}$$

$$= \sum_{\substack{r=1 \\ r\neq i,j}}^{n} \left\{ W_r^+ \left[\frac{f_2^{(\sim i,j)}(x_r^{(\sim i,j)+}) - f_2^{(\sim i,j)}\left(\mu_x^{(\sim i,j)}\right)}{h_r^+} \right]^2 + W_r^- \left[\frac{f_2^{(\sim i,j)}(x_r^{(\sim i,j)-}) - f_2^{(\sim i,j)}\left(\mu_x^{(\sim i,j)}\right)}{h_r^-} \right]^2 + W_r^{\pm} \frac{\left[f_2^{(\sim i,j)}(x_r^{(\sim i,j)+}) - f_2^{(\sim i,j)}\left(\mu_x^{(\sim i,j)}\right) \right] \left[f_2^{(\sim i,j)}(x_r^{(\sim i,j)-}) - f_2^{(\sim i,j)}\left(\mu_x^{(\sim i,j)}\right) \right]}{h_r^+ h_r^-} \right\},$$

where $\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i,j)}$, $\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i,j)+}$, and $\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i,j)-}$ are defined by removing the i_{th} and j_{th} variable from the original input vector, as shown in the following equations:

$$\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i,j)} = [\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_{i-1}}, \mu_{x_{i+1}}, \dots, \mu_{x_{j-1}}, \mu_{x_{j+1}}, \dots, \mu_{x_n}] \quad \textbf{(6-31)}$$

$$\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i,j)+} = \left[\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_r} + h_r^+ \sigma_{x_r}, \dots, \mu_{x_{i-1}}, \mu_{x_{i+1}}, \dots, \mu_{x_{j-1}}, \mu_{x_{j+1}}, \dots, \mu_{x_n}\right], \quad r \neq i,j \quad \textbf{(6-32)}$$

$$\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i,j)-} = \left[\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_r} + h_r^- \sigma_{x_r}, \dots, \mu_{x_{i-1}}, \mu_{x_{i+1}}, \dots, \mu_{x_{j-1}}, \mu_{x_{j+1}}, \dots, \mu_{x_n}\right], \quad r \neq i,j \quad \textbf{(6-33)}$$

Now taking X_i, X_j as input variables, $V_{x_{\sim i,j}}\left(\left. y\right|_{x_i=X_i,x_j=X_j}\right)$ becomes a function of X_i, X_j:

$$g_2(X_i,X_j)=V_{x_{\sim i,j}}\left(\left. y\right|_{x_i=X_i,x_j=X_j}\right) \tag{6-34}$$

Equation (5-1) from URQ could now be used,

$$\begin{aligned} E_{X_{i,j}}\left(V_{x_{\sim i,j}}\left(\left. y\right|_{x_i=X_i,x_j=X_j}\right)\right) &= E_{X_{i,j}}\left(g_2(X_i,X_j)\right) \\ &= W_0^{(i,j)} g_2\left(\mu_{x_i},\mu_{x_j}\right) \\ &+ W_i\left[\frac{g_2(\mu_{x_i}+h_i^+\sigma_{x_i},\mu_{x_j})}{h_i^+}-\frac{g_2(\mu_{x_i}+h_i^-\sigma_{x_i},\mu_{x_j})}{h_i^-}\right] \\ &+ W_j\left[\frac{g_2(\mu_{x_i},\mu_{x_j}+h_j^+\sigma_{x_j})}{h_j^+}-\frac{g_2(\mu_{x_i},\mu_{x_j}+h_j^-\sigma_{x_j})}{h_j^-}\right], \end{aligned} \tag{6-35}$$

where W_i, W_j, h_i^+, h_j^+, h_i^- and h_j^- are defined in Table 5-1, and $W_0^{(i,j)}$ is defined by:

$$W_0^{(i,j)}=1+\frac{1}{h_i^+h_i^-}+\frac{1}{h_j^+h_j^-} \tag{6-36}$$

Substituting equations (6-36) and (5-2) into equation (6-7), the second order Sobol' index is obtained:

$$\begin{aligned} S_{ij} \approx \Bigg\{ & V_{URQ}(y)-W_0^{(i,j)} g_2\left(\mu_{x_i},\mu_{x_j}\right) \\ & + W_i\left[\frac{g_2\left(\mu_{x_i}+h_i^+\sigma_{x_i},\mu_{x_j}\right)}{h_i^+}-\frac{g_2\left(\mu_{x_i}+h_i^-\sigma_{x_i},\mu_{x_j}\right)}{h_i^-}\right] \\ & + W_j\left[\frac{g_2\left(\mu_{x_i},\mu_{x_j}+h_j^+\sigma_{x_j}\right)}{h_j^+}-\frac{g_2\left(\mu_{x_i},\mu_{x_j}+h_j^-\sigma_{x_j}\right)}{h_j^-}\right]\Bigg\} \\ & /V_{URQ}(y)-S_i-S_j \end{aligned} \tag{6-37}$$

6.3.2.3 Total Order Indices

Following equation (6-14), this time the inner loop $V_{x_i}\left(y|_{\boldsymbol{x_{\sim i}=X_{\sim i}}}\right)$ is calculated by varying only x_i. For this purpose, $f_3^{(i)}$ is defined by fixing all the other variables,

$$f_3^{(i)}(x_i) = y|_{\boldsymbol{x_{\sim i}=X_{\sim i}}} = f(X_1, X_2, \dots, X_{i-1}, x_i, X_{i+1}, \dots, X_n), \tag{6-38}$$

and $V_{x_i}\left(y|_{\boldsymbol{x_{\sim i}=X_{\sim i}}}\right)$ is obtained as:

$$\begin{aligned} V_{x_i}\left(y|_{\boldsymbol{x_{\sim i}=X_{\sim i}}}\right) &= V\left(f_3^{(i)}\right) \\ &= W_i^+ \left[\frac{f_3^{(i)}(x_i^+) - f_3^{(i)}(\mu_{x_i})}{h_i^+}\right]^2 + W_i^- \left[\frac{f_3^{(i)}(x_i^-) - f_3^{(i)}(\mu_{x_i})}{h_i^-}\right]^2 \\ &+ W_i^{\pm} \frac{\left[f_3^{(i)}(x_i^+) - f_3^{(i)}(\mu_{x_i})\right]\left[f_3^{(i)}(x_i^-) - f_3^{(i)}(\mu_{x_i})\right]}{h_i^+ h_i^-} \end{aligned} \tag{6-39}$$

Now $V_{x_i}\left(y|_{\boldsymbol{x_{\sim i}=X_{\sim i}}}\right)$ becomes a function of all the variables, except x_i

$$g_3(X_1, X_2, \dots X_{i-1}, X_{i+1}, \dots X_n) = V_{x_i}\left(y|_{\boldsymbol{x_{\sim i}=X_{\sim i}}}\right) \tag{6-40}$$

By applying Equation (5-1) on g_3, the nested expectation could be calculated:

$$\begin{aligned} E_{\boldsymbol{X_{\sim i}}}\left(V_{x_i}\left(y|_{\boldsymbol{x_{\sim i}=X_{\sim i}}}\right)\right) &= E_{\boldsymbol{X_{\sim i}}}\left(g_3(X_1, X_2, \dots X_{i-1}, X_{i+1}, \dots X_n)\right) \\ &= W_0^{(\sim i)} g_3\left(\boldsymbol{\mu_x}^{(\sim i)}\right) + \sum_{\substack{r=1 \\ r\neq i}}^{n} W_r \left[\frac{g_3(\boldsymbol{x_r}^{+(\sim i)})}{h_r^+} - \frac{g_3(\boldsymbol{x_r}^{-(\sim i)})}{h_r^-}\right], \end{aligned} \tag{6-41}$$

where $\boldsymbol{\mu_x}^{(\sim i)}$, $\boldsymbol{x_r}^{+(\sim i)}$, and $\boldsymbol{x_r}^{-(\sim i)}$ have been defined in equation (6-22), (6-23), and (6-24), respectively. $W_0^{(\sim i)}$ is given by:

$$W_0^{(\sim i)} = W_0 - \frac{1}{h_i^+ h_i^-} \tag{6-42}$$

Substituting equations (6-41) and (5-2) into equation (6-14), the total effect Sobol' indices is obtained,

$$S_i^T \approx \frac{W_0^{(\sim i)} g_3\left(\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i)}\right) + \sum_{\substack{r=1 \\ r\neq i}}^{n} W_r \left[\frac{g_3(\boldsymbol{x}_{\boldsymbol{r}}{}^{+}{}^{(\sim i)})}{h_r^+} - \frac{g_3(\boldsymbol{x}_{\boldsymbol{r}}{}^{-}{}^{(\sim i)})}{h_r^-}\right]}{V_{URQ}(y)} \quad \textbf{(6-43)}$$

6.3.3 Algorithm and Computational Cost

In this section, the algorithm of the proposed method is given. Once again to avoid repetition, only Option 1 from Table 6-1 will be discussed, as Option 2 and 3 could be implemented in a similar way. The algorithm for Option 4 is a straightforward application of URQ.

We start with the first order indices as presented in section 6.3.2.1; the overall process for calculating $E_{X_i}\left(V_{x_{\sim i}}\left(y|_{x_i=X_i}\right)\right)$ is illustrated in Figure 6-2, which could be considered as one specific realization of the general process in Figure 6-1

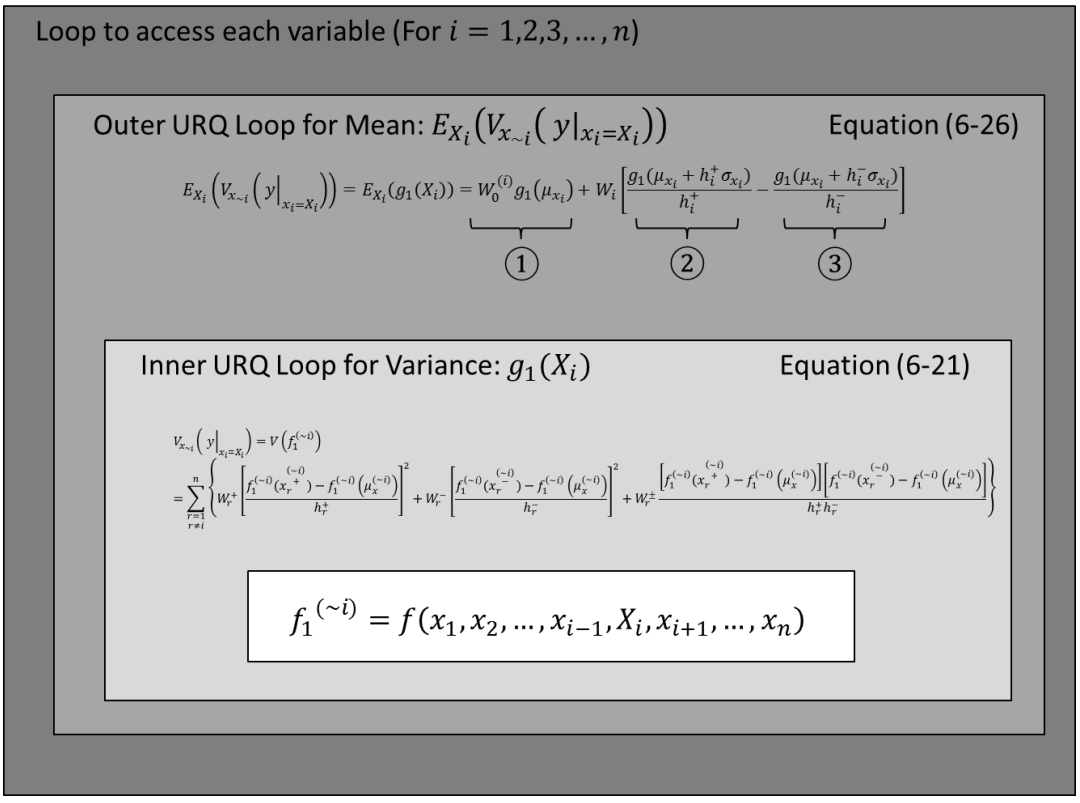

Figure 6-2. Illustration of the computational algorithm

To execute the outer URQ loop, X_i needs be sampled at μ_{x_i}, $\mu_{x_i} + h_i^+\sigma_{x_i}$, and $\mu_{x_i} + h_i^-\sigma_{x_i}$, for the term ①, ②, and ③, repectively. For each realization of X_i, the inner URQ loop need to sample the other $n-1$ variables at μ_{x_r}, $\mu_{x_r} + h_r^+\sigma_{x_r}$, and $\mu_{x_r} + h_r^-\sigma_{x_r}$, where $r \neq i$. It should be noted that in this nested loop, many samples could be reused. Therefore it is beneficial to evaluate the model at all the required points first, then to start the algebraic calculation of equation (6-21) and (6-26). The sampling includes:

- When all the variables are at their mean values: one evaluation
- When one variable is at its $\mu + h^+\sigma$ value: n evaluations
- When one variable is at its $\mu + h^-\sigma$ value: n evaluations
- When two variables are at their $\mu + h^+\sigma$ values: $\binom{n}{2} = n(n-1)/2$ evaluations
- When two variables are at their $\mu + h^-\sigma$ values: $\binom{n}{2} = n(n-1)/2$ evaluations
- When one variable is at its $\mu + h^+\sigma$ value and another variable is at its $\mu + h^-\sigma$ value: $n(n-1)$ evaluations

The total number of evaluations is therefore:

$$N_{total} = 2n^2 + 1 \quad \textbf{(6-44)}$$

For the second order indices, the process is similar. In Figure 6-2, replace equation (6-21) and (6-26) by equation (6-30) and (6-35), respectively. The outer URQ loop now has five terms. For each $(x_i, x_j) = (\mu_{x_i}, \mu_{x_j})$, $(\mu_{x_i} + h_i^+\sigma_{x_i}, \mu_{x_j})$, $(\mu_{x_i} + h_i^-\sigma_{x_i}, \mu_{x_j})$, $(\mu_{x_i}, \mu_{x_j} + h_j^+\sigma_{x_j})$, and $(\mu_{x_i}, \mu_{x_j} + h_j^-\sigma_{x_j})$, the inner URQ loop is calculated while the other $n-2$ variables are set to be μ_{x_r}, $\mu_{x_r} + h_r^+\sigma_{x_r}$, and $\mu_{x_r} + h_r^-\sigma_{x_r}$, where $r \neq i, j$, once at a time. It can be shown that the same set of samples used for the first order indices can also be used for the second order indices.

For total effect indices, the outer URQ loop now has $2n-1$ terms, in these terms, each $x_r \neq x_i$ is set as μ_{x_r}, $\mu_{x_r} + h_r^+\sigma_{x_r}$, and $\mu_{x_r} + h_r^-\sigma_{x_r}$ one a time. In the inner URQ loop, only x_i is sampled, at the values of μ_{x_i}, $\mu_{x_i} + h_i^+\sigma_{x_i}$, and $\mu_{x_i} + h_i^-\sigma_{x_i}$. Similar to the second order indices, no further sampling is required for the total effect indices.

The implementations of the Option 2 & 3 are similar, and also the same sampling points are required. Option 4, as discussed previously is a straightforward application of single-loop URQ for the mean of $f_+(\boldsymbol{x}_{+i})$ defined by Eq. (6-5). This requires even less points as there is no nested structure in calculation. As $f_+(\boldsymbol{x}_{+i})$ has $2n-1$ input variables, the number of model evaluations for Option 4 is:

$$N_{total} = 4n - 1 \tag{6-45}$$

A pseudo code for the first order indices using Option 1 is attached in Appendix A.2. A full Matlab code for the entire index using all the four options is attached in Appendix A.3.

Because Option 1, 2, and 3 require the same set of sampling points (and Option 4 requires a subset of these points), for each application, it is recommended to calculate four sets of indices using all the options. A potential way to compare the accuracy of these results (without knowing the true values) is to sum all the first order and high order indices from each option. Theoretically, for a set of indices from one option, the sum should be one, or slightly smaller if some higher order indices are neglected in implementation (assuming that the high order interactions are not significant). The option which gives the sum closest to one should be selected.

6.4 Evaluation

The evaluation is conducted via a series of numerical experiments. The proposed method is applied on a number of test-cases and the results are compared with theoretical values (where available) or with estimations from the traditional (benchmark) MCS approach. These include 18 analytical examples and one practical design case study.

The analytical examples are based on single line algebraic equations so that the representative mathematical properties could be explored explicitly. By contrast, the practical design case study is based on a complex engineering model. The purpose is to demonstrate the scaling potential of the proposed method.

6.4.1 Analytical Test Cases

6.4.1.1 Problem Statement

The typical mathematical properties, which may affect and limit the applicability of the proposed method include:

- Non-linearity of the model;
- Non-monotonicity of the model;
- Type of the distribution;
- Multi-modality of the distribution;
- Interaction of variables;
- Combinations of the above.

In order to take into account these properties, the test-cases are intended to form a "control variable experiment", which involves five models and various settings of several probability distributions. Although a large amount of experiments have been used for testing and evaluation, only the most representative ones are shown in this thesis. A summary of the test-cases are shown in Table 6-2, where the model and the corresponding distribution settings are indicated by the first and second number in the notation of the test-case. For example T1-3 indicates the combination of the first model and the third distribution settings.

The first three test-cases (T1-1, T1-2, and T1-3) are based on a linear model, defined by equation (6-46). In T1-1, all the inputs follow the uniform distribution $U(0,1)$. This is the simplest case which serves as a reference. In T1-2, the distributions are changed into Gaussian, triangular and uniform. T1-3 is designed to investigate the influence of multimodal distributions. The selected distribution is the mixed-Gaussian (McLachlan and Peel, 2000), which is represented as $\mathcal{MG}(\mu_1, \mu_2, \sigma_1, \sigma_2, \lambda_1, \lambda_2)$, where μ, σ, and λ are the mean, standard deviation, and proportions of two Gaussian distributions, respectively. It should be noted that in T1-2 and T1-3, the parameters of the distributions are arbitrarily selected. Also, as mentioned above, various other settings have been tested beside the ones reported here, and no influence on accuracy has been found for these linear cases.

Test cases T2-1, T2-2, T2-3, T2-4, T2-5, and T2-6 are based on the non-linear model defined by equation (6-47). In this model, there is no interaction effect between the variables. The influence of non-monotonicity is explored in T2-1, T2-2, and T2-3, in which the distributions of the input variables are gradually expanded from [0,1] to more non-monotonic regions. T2-4 and T2-5 are used to demonstrate the combination of non-monotonicity and multi-modality. T2-6 covers other PDF shapes under non-linear cases.

Test cases T3-1, T3-2, and T3-3 are proposed to evaluate the method on high order indices (interaction effects). In these three cases, equation (6-48) is used and the expected interaction effects are dependent on the ranges of input variables. In T3-1, first order effect will be the main contributor of the total variance, while in T3-2 and T3-3 more interactions are involved.

Apart from the proposed functions, experiments have also been conducted on the Sobol' G-Function specified in equation (6-49) (the original source is (Saltelli and Sobol', 1995), while the analytical solution is presented in (Saltelli et al., 2010)) and the Ishigami-Function in equation (6-50) (original source is (Ishigami and Homma, 1990), analytical solution available in (Sobol' and Levitan, 1999)). These two functions are considered as classic test-cases for sensitivity analysis.

$$f_{T1}(x_1, x_2, x_3) = x_1 + 2x_2 + 3x_3 \quad \textbf{(6-46)}$$

$$f_{T2}(x_1, x_2, x_3) = x_1^2 + \sin\left(\frac{\pi}{2}x_2\right) + e^{|x_3|} \quad \textbf{(6-47)}$$

$$f_{T3}(x_1, x_2, x_3) = x_1 x_3 + x_1 \sin\left(\frac{\pi}{2}x_2\right) + x_2 e^{|x_3|} + x_1 x_2 x_3 \quad \textbf{(6-48)}$$

$$f_{T4}(x_1, x_2, x_3) = \prod_{i=1}^{3} \frac{|4x_i - 2| + \rho_i}{1 + \rho_i} \quad \textbf{(6-49)}$$

$$f_{T5}(x_1, x_2, x_3) = sin(x_1) + \theta_1 \cdot sin^2(x_2) + \theta_2 \cdot x_3^4 sin(x_1) \quad \textbf{(6-50)}$$

Table 6-2. Summary of the test-cases

Notation	Equation	Distribution Settings	Justification
T1-1:	Linear Model (6-46)	$x_1 \sim \mathcal{U}(0,1)$ $x_2 \sim \mathcal{U}(0,1)$ $x_3 \sim \mathcal{U}(0,1)$	Simplest reference case
T1-2:	Linear Model (6-46)	$x_1 \sim \mathcal{N}(1,1)$ $x_2 \sim \mathcal{T}ri(0,2,1.5)$ $x_3 \sim \mathcal{U}(0,2)$	Linearity and mixture of different distributions
T1-3:	Linear Model (6-46)	$x_1 \sim \mathcal{MG}(0,5,0.5,0.7,0.25,0.75)$ $x_2 \sim \mathcal{MG}(1,3,0.5,0.5,0.5,0.5)$ $x_3 \sim \mathcal{MG}(5,7,0.7,0.5,0.75,0.25)$	Linear and multi-modality
T2-1:	Non-Linear Model (6-47)	$x_1 \sim \mathcal{U}(0,1)$ $x_2 \sim \mathcal{U}(0,1)$ $x_3 \sim \mathcal{U}(0,1)$	Non-linearity and monotonicity
T2-2:	Non-Linear Model (6-47)	$x_1 \sim \mathcal{U}(-1,1)$ $x_2 \sim \mathcal{U}(0,2)$ $x_3 \sim \mathcal{U}(-1,1)$	Non-linearity and non-monotonicity
T2-3:	Non-Linear Model (6-47)	$x_1 \sim \mathcal{U}(-1,1)$ $x_2 \sim \mathcal{U}(-2.5,2.5)$ $x_3 \sim \mathcal{U}(-1,1)$	Non-linearity and non-monotonicity
T2-4:	Non-Linear Model (6-47)	$x_1 \sim \mathcal{MG}(-0.5,0.5,0.2,0.2,0.75,0.25)$ $x_2 \sim \mathcal{MG}(-2.2,2.2,0.5,0.5,0.5,0.5)$ $x_3 \sim \mathcal{MG}(-0.5,0.5,0.2,0.2,0.25,0.75)$	Non-Linear and multi-modality
T2-5:	Non-Linear Model (6-47)	$x_1 \sim \mathcal{MG}(-0.5,0.5,0.2,0.2,0.75,0.25)$ $x_2 \sim \mathcal{MG}(-1.8,1.8,0.6,0.6,0.5,0.5)$ $x_3 \sim \mathcal{MG}(-0.5,0.5,0.2,0.2,0.25,0.75)$	Non-Linear and multi-modality
T2-6:	Non-Linear Model (6-47)	$x_1 \sim \mathcal{N}(1,1)$ $x_2 \sim \mathcal{T}ri(-2,3,1)$ $x_3 \sim \mathcal{U}(-2,2)$	Non-Linear and mixture of different distributions
T3-1:	Non-Linear Model (6-48)	$x_1 \sim \mathcal{U}(0,1)$ $x_2 \sim \mathcal{U}(0,1)$ $x_3 \sim \mathcal{U}(0,1)$	Non-linearity and moderate interaction effect
T3-2:	Non-Linear Model (6-48)	$x_1 \sim \mathcal{U}(-0.5,1.5)$ $x_2 \sim \mathcal{U}(-0.5,1.5)$ $x_3 \sim \mathcal{U}(-0.5,1.5)$	Non-linearity and strong interaction effect
T3-3:	Non-Linear Model (6-48)	$x_1 \sim \mathcal{U}(-1,1)$ $x_2 \sim \mathcal{U}(-1,1)$ $x_3 \sim \mathcal{U}(-1,1)$	Non-linearity, interaction effect
T4-1:	Sobol' G Function (6-49)	$x_1 \sim \mathcal{U}(0,1); \rho_1 = 0$ $x_2 \sim \mathcal{U}(0,1); \rho_2 = 0$ $x_3 \sim \mathcal{U}(0,1); \rho_3 = 0$	Classic Test Case
T4-2:	Sobol' G Function (6-49)	$x_1 \sim \mathcal{U}(0,1); \rho_1 = 0$ $x_2 \sim \mathcal{U}(0,1); \rho_2 = 3$	Classic Test Case

		$x_3 \sim \mathcal{U}(0,1);\ \rho_3 = 5$	
T4-3:	Sobol' G Function (6-49)	$x_1 \sim \mathcal{U}(0,1);\ \rho_1 = 10$ $x_2 \sim \mathcal{U}(0,1);\ \rho_2 = 30$ $x_3 \sim \mathcal{U}(0,1);\ \rho_3 = 50$	Classic Test Case
T5-1:	Ishigami Function (6-50)	$x_1 \sim \mathcal{U}(0,1);\ \theta_1 = 7$ $x_2 \sim \mathcal{U}(0,1);\ \theta_2 = 0.1$ $x_3 \sim \mathcal{U}(0,1)$	Classic Test Case: (θ_1 and θ_2 from (Ishigami and Homma, 1990))
T5-2:	Ishigami Function (6-50)	$x_1 \sim \mathcal{U}(0,1);\ \theta_1 = 7$ $x_2 \sim \mathcal{U}(0,1);\ \theta_2 = 0.05$ $x_3 \sim \mathcal{U}(0,1)$	Classic Test Case: (θ_1 and θ_2 from(Sobol' and Levitan, 1999))
T5-3:	Ishigami Function (6-50)	$x_1 \sim \mathcal{U}(0,1);\ \theta_1 = 7$ $x_2 \sim \mathcal{U}(0,1);\ \theta_2 = 0.01$ $x_3 \sim \mathcal{U}(0,1)$	Classic Test Case:

6.4.1.2 Results

For each test-case, all the four options (as summarized in Table 6-1) have been tested and the results are plotted in Figure 6-3 to Figure 6-7, with comparison to the theoretical/reference values.

For most of the cases, theoretical values were obtained by using the equations from (Sobol', 1993). Some parts of the integral were numerically solved using Matlab (and validated by MCS). In T1-3, T2-4, and T2-5, the theoretical values were difficult to obtain due to the distributions being used; therefore values from the MCS are used as references. The theoretical results of the adopted test-cases are obtained from relevant papers referred to above.

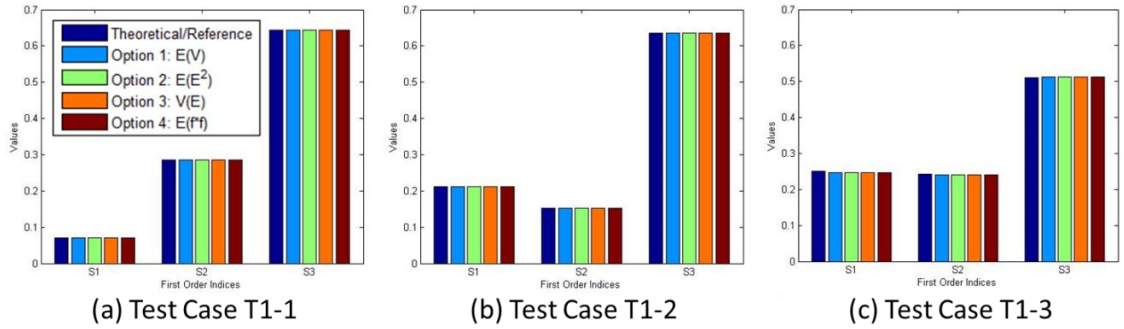

Figure 6-3. Results of the linear model under different settings

It can be seen from Figure 6-3 that the proposed method worked well on the linear model, regardless of the distributions used (including multi-modal distributions). The

results from the four options are almost the same and are very close to the theoretical/reference values.

Regarding the first non-linear model, the proposed method was able to deliver accurate results when all the inputs are inside the monotonic region of the function, as shown by Figure 6-4(a). In test-case T2-2 (Figure 6-4(b)), as the input distributions are expanded to the non-monotonic regions, the errors increase. It can be seen that, Option 4 is the most affected one, resulting in negative values for S_1 and S_2. The results of the other three options are still relatively close to the theoretical/reference values. Figure 6-4 (c) shows an extreme case when S_2 is not detectable by the proposed method.

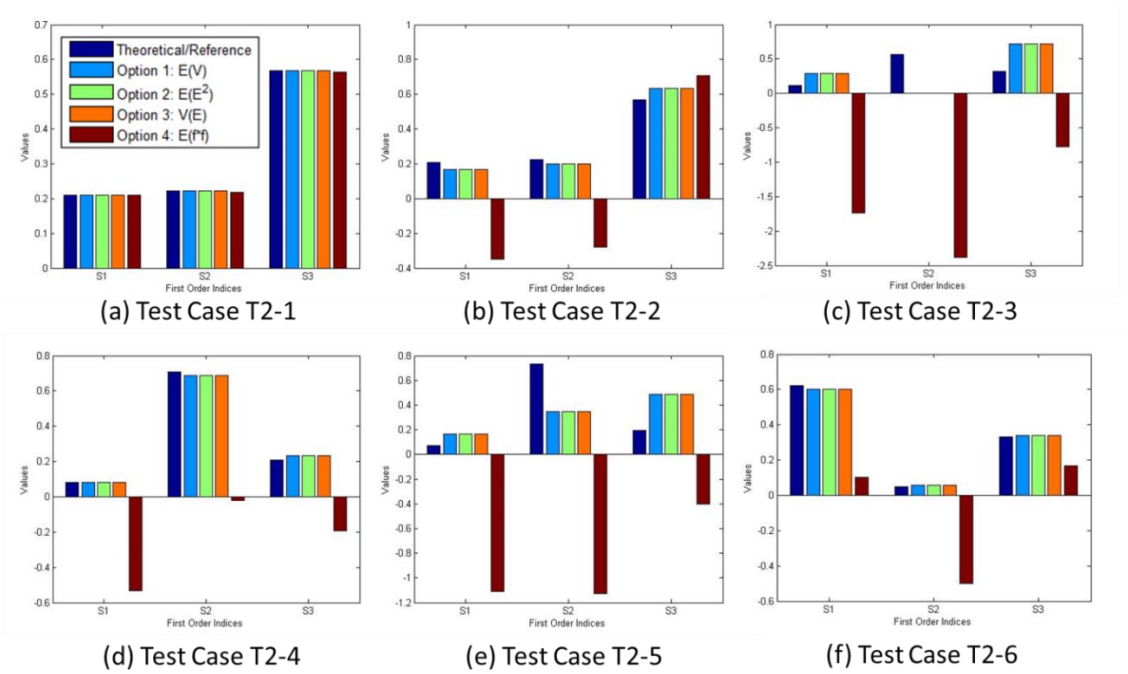

Figure 6-4. Results of the non-linear model 1 under different settings

Figure 6-4 (d) and (e) represent the results of test-case T2-4 and T2-5. It can be seen that the accuracy of the former is considerably higher, especially regarding S_2. This difference was initially unexpected, because in these two cases, the same model is used and the distribution settings are very close. The only slight difference is the distributions of x_2, which still covers roughly the same region (from -4 to 4) in both cases. A further investigation on this effect will be shown in Section 6.4.1.3.

Figure 6-4 (f) shows the results of test-case T2-6, in which a number of different distributions are used. The results from the first three options are close and relatively accurate.

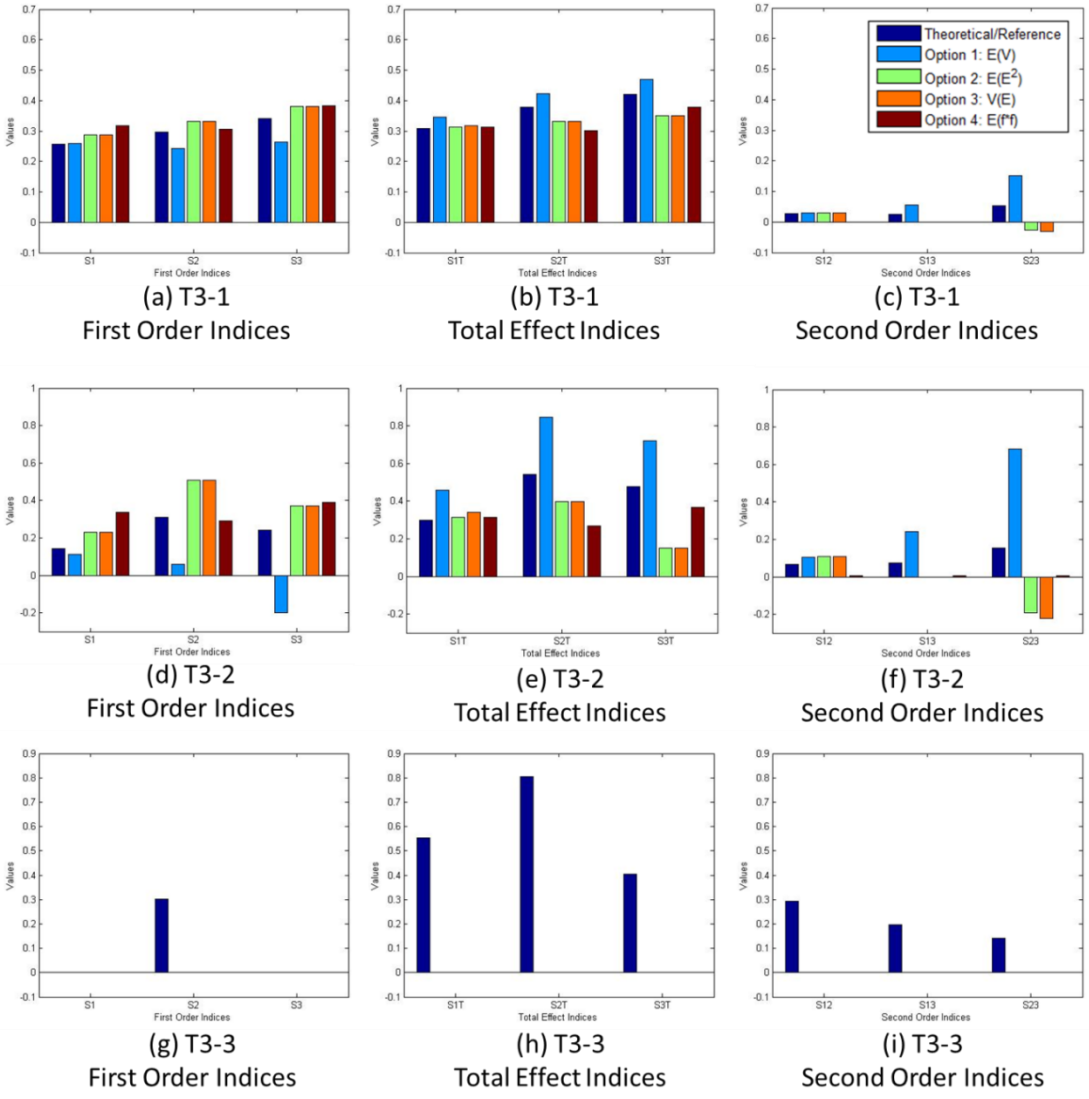

Figure 6-5. Results of test cases with interaction effects

The results of T3-1, T3-2 and T3-3 are shown in Figure 6-5. It could be seen that, the predicted first order and total effect indices are relatively accurate in test case T3-1, as plotted in Figure 6-5, (a) and (b). In T3-2, when the region of input is expanded to $(-0.5,1.5)$, the accuracy has been reduced, especially for option 2 (Figure 6-5, (d) and (e)). In both T3-1 and T3-2, the second order indices do not come very close to the benchmark (Figure 6-5, (c) and (f)). T3-3 is an extreme case, when the proposed method

fails altogether. It should be noted that in Figure 6-5 (g), the reference values of S_1 and S_3 are also zeros.

The results of test-cases T4 and T5 are shown in Figure 6-6 and Figure 6-7, respectively. The quality of results is dependent on the selected coefficients in equations (6-49) and (6-50). Regarding the Sobol' G-Function, when the ρ_i's are all zeros, the proposed method could not produce any results as shown in Figure 6-6 (a), (b), and (c). As the values of ρ_i's are increased, the first order and total order indices from option 1 and option 3 become more accurate (Figure 6-6 (d), (e), (g), and (h)). Regarding the Ishigami Function, the smaller value of θ_2, leads to better predictions (Figure 6-7). For all the cases, the second order indices are still quite irrelevant.

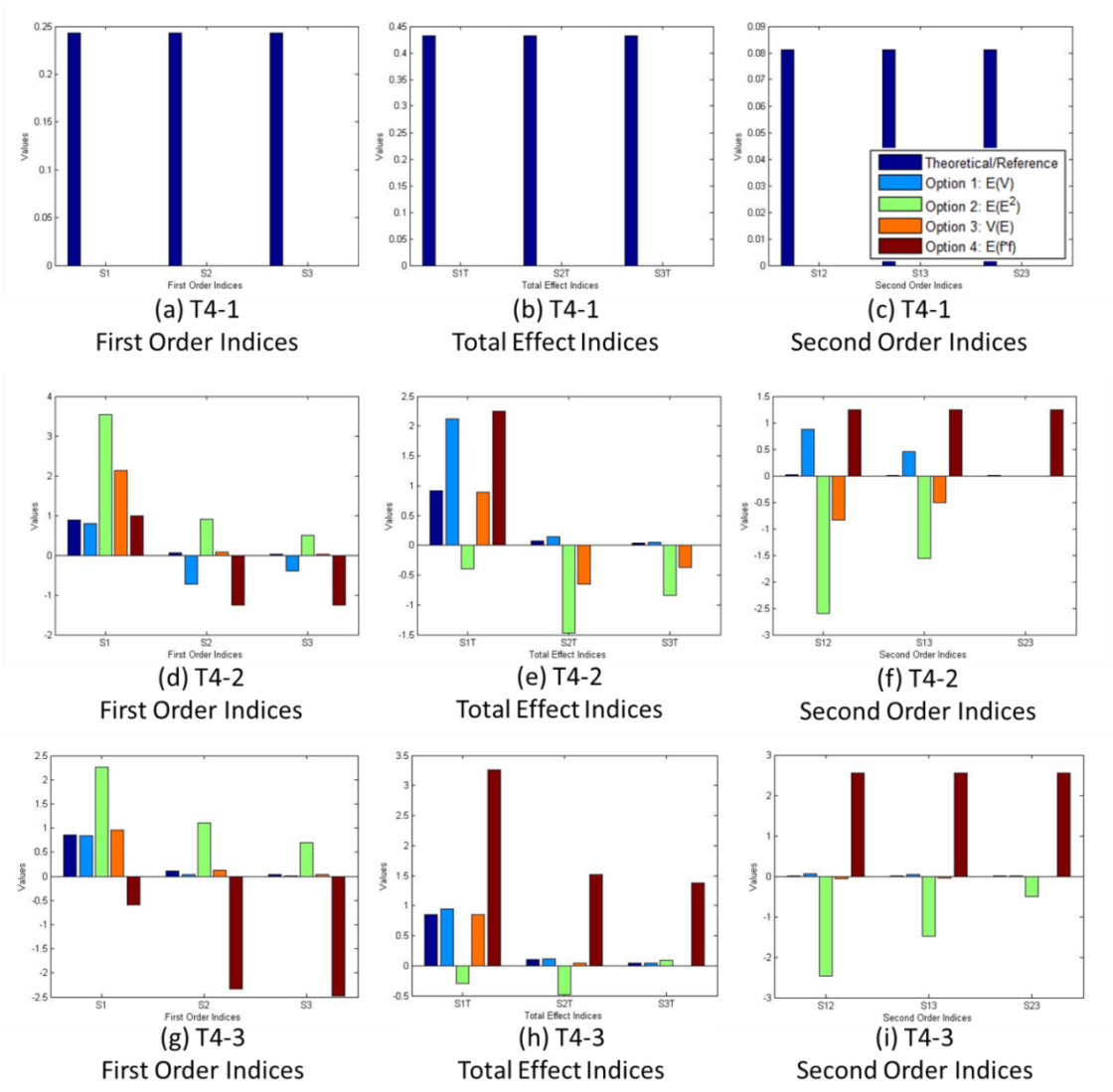

Figure 6-6. Results for test on the Sobol' G-function

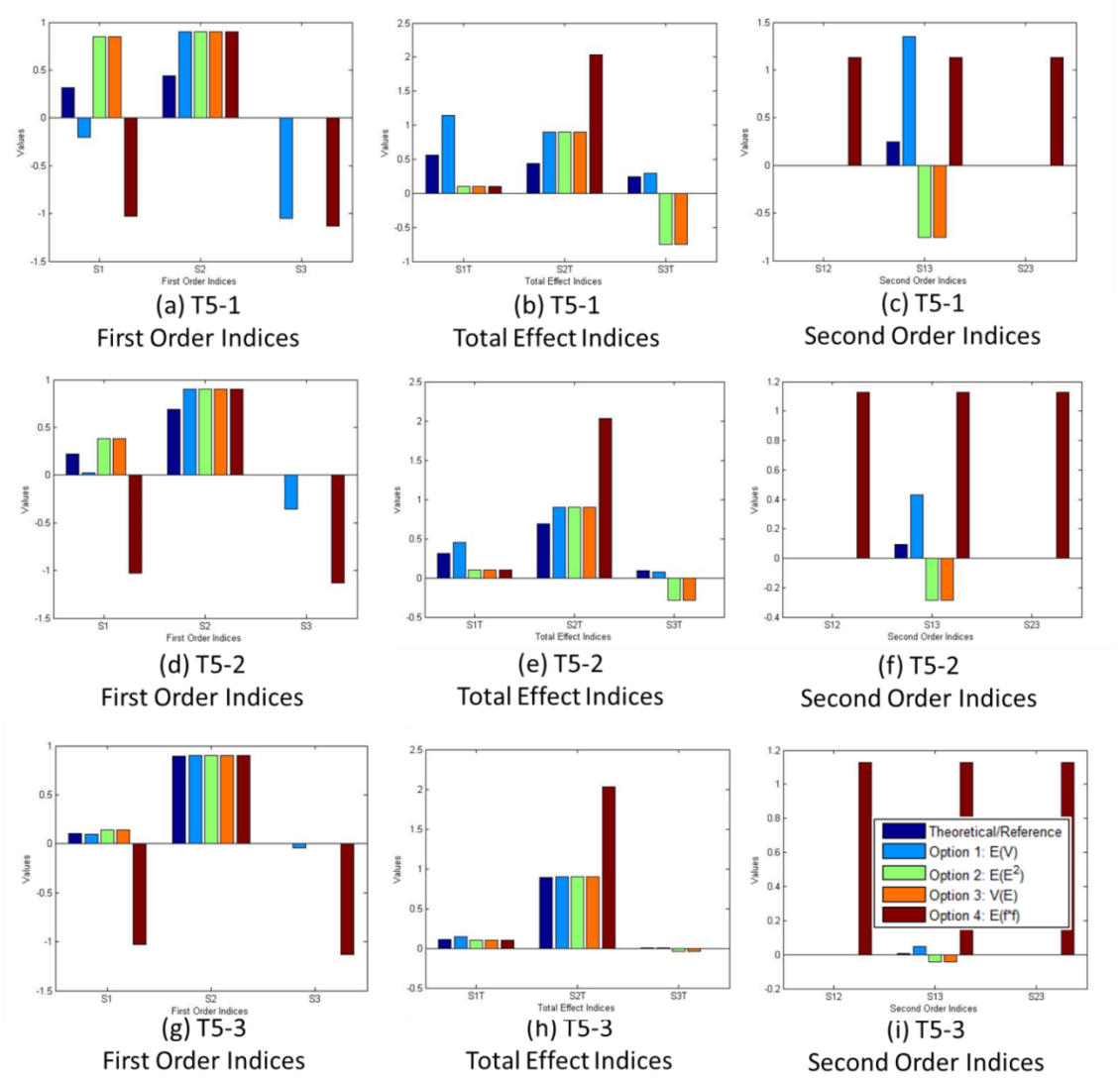

Figure 6-7. Results for tests on Ishigami Function

6.4.1.3 Discussion

The evaluation confirmed at large the effectiveness of the method, but also revealed some limitations. It was discovered that these are largely limitations of the URQ technique itself. The first limitation is due to aliasing, which could explain the low accuracy in T2-3 and T2-5.

Considering T2-3, due to the selected distribution for $x_2 \sim \mathcal{U}(-2.5, 2.5)$, URQ will sample x_2 at the points of -1.9365, 0, and 1.9365 (corresponding to $\mu_{x_2} + h_2^- \sigma_{x_2}$, μ_{x_2}, and $\mu_{x_2} + h_2^+ \sigma_{x_2}$, respectively). At these three points, the variation of $sin\left(\frac{\pi}{2} x_2\right)$ are grossly underestimated, as illustrated in Figure 6-8. This directly leads to the disregard

of S_2 and a reduction in the calculated total variance. The latter will also cause over-estimation of S_1 and S_3 (Figure 6-4 (c)).

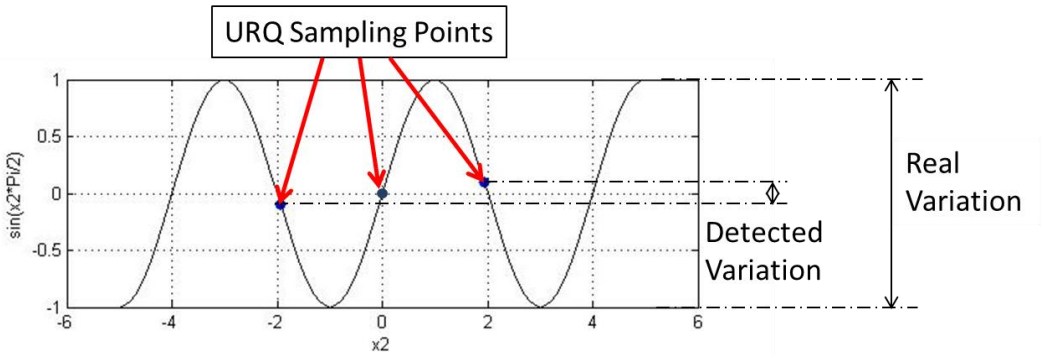

Figure 6-8. Under-estimated variation for $sin\left(\frac{\pi}{2}x_2\right)$

For T2-4 and T2-5, although in both cases, the distributions of x_2 have covered roughly the same region from -4 to 4, the subtle difference between the two PDFs has shifted the sampling points accordingly, as illustrated in Figure 6-9. Due to this shift, the accuracy of results in T2-4 (Figure 6-4(d)) is higher than that of T2-5 (Figure 6-4(e)).

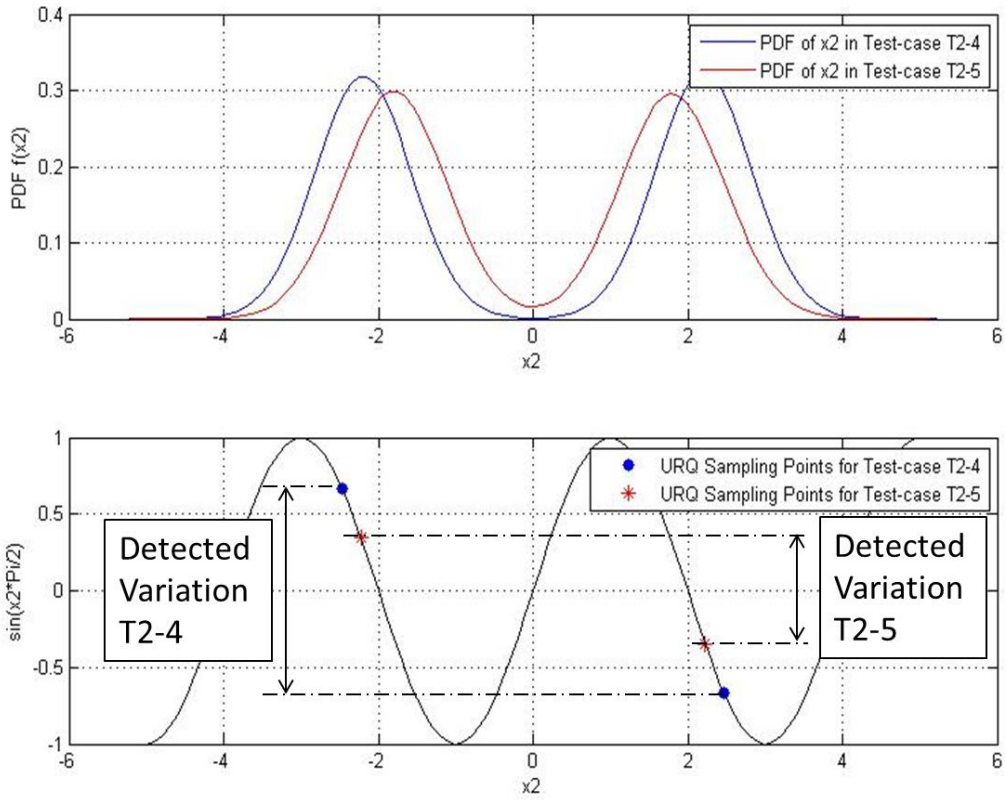

Figure 6-9. Comparison of the two PDFs and the relevant sampling points

The second limitation is that URQ cannot handle one specific feature of models. Figure 6-10 illustrates a 2-dimensional example of $y = sin(x_1) \cdot sin(x_2)$ to visualise the problem. The model output will be zero, on the x_1 or x_2 axis. However URQ would perturb only one variable at a time and keep the others at their mean values. If the distributions happen to have the mean values at zero, all the URQ sampling points (as indicated by the red dots in Figure 6-10) will return zeros. Therefore the variance will be totally disregarded, which cannot be used as denominator to calculate the indices (the estimated mean is zero as well). In test case T3-3 and T4-1 the models have similar (multiplicative) features and the distributions of inputs are symmetric to the zero point. This is the reason why no results could be given by this method for these two cases. In the Ishigami function, this effect is caused by the component $\theta_2 \cdot x_3^4 sin(x_1)$, therefore by using smaller values of coefficient θ_2, the effect will be weakened, which gives more accurate predictions (as shown by T5-1, T5-2, and T5-3).

As the aforementioned limitations are largely due to the URQ technique itself, it is expected that the proposed approach could be improved by adopting alternative propagation techniques. For instance, by using the propagation technique proposed in (Seo and Kwak, 2002), additional sampling points will be used as indicated by the green dots in Figure 6-10. This will capture more information of the model response surface, and avoid disregarding the output variance. Another potential solution is to introduce coordinate transformations in future developments.

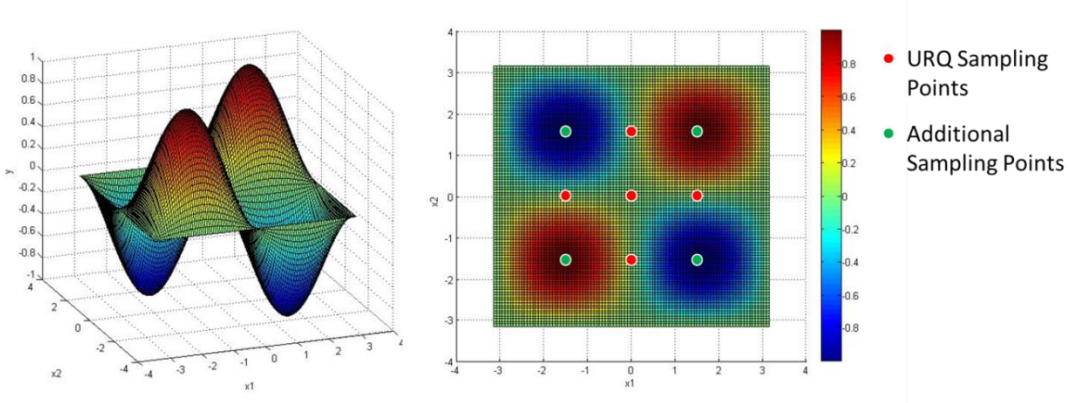

Figure 6-10. Example of a 2D case

6.4.2 Practical Test Case

6.4.2.1 Problem Statement

The proposed method was applied on an example of an aircraft environmental control system (ECS), adopted from Pérez-Grande and Leo (Pérez-Grande and Leo, 2002) as shown in Figure 6-11. The objective is to demonstrate a practical application of realistic size and complexity.

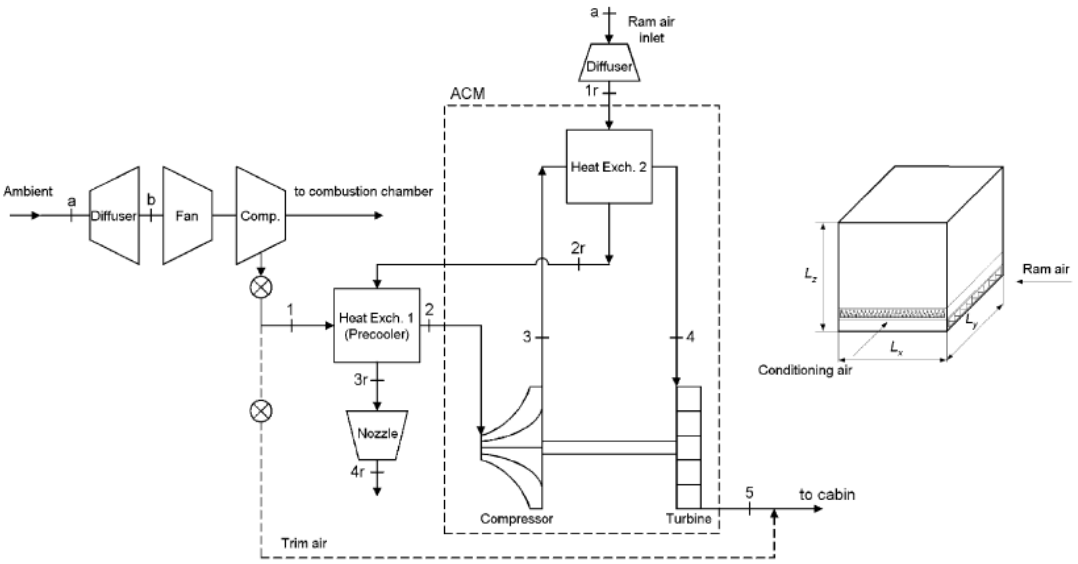

Figure 6-11. Schematic View of the ECS and core of a compact cross-flow heat exchanger (Adopted from (Pérez-Grande and Leo, 2002))

In this example, thermal characteristics of an ECS such as pressures, temperatures at different locations and the overall entropy generation rate are calculated using designer-specified geometry properties. The latter include the heat exchanger dimensions, fin characteristics, and efficiencies of turbines, compressors, nozzles, and diffusers. There are 119 models in total, adopted from (Pérez-Grande and Leo, 2002), including 59 independent (inputs and parameters) and 120 dependent variables (intermediate variables and outputs), respectively. The details of the inputs and outputs of interest are listed in Table 6-3 and Table 6-4, respectively.

While the original work of Pérez-Grande and Leo (Pérez-Grande and Leo, 2002) was focused on ECS optimization, the current study is concerned with sensitivity analysis of their optimal solution. Uncertainties have been considered for all the 42 input variables,

listed in Table 6-3. In the original paper (Pérez-Grande and Leo, 2002), some of these inputs were considered fixed parameters, while in the current test case, their variations are also taken into account. All the nominal values are adopted from the original paper and the uncertainty distributions are arbitrarily assigned, with combinations of uniform, Gaussian, triangular, and mixture-Gaussian (multi-modal) distributions. It should be emphasised that the purpose of this study is to validate the proposed method, rather than conduct an investigation into the merits of the particular ECS design. In a real design scenario, these distributions can be based on historical data or expert elicitations.

Table 6-3. Input Variables and Associated Uncertainties

Components	Inputs	Symbols	Probability Distribution
Flight Condition	Altitude (m)	h	$\mathcal{N}(11000, 550)$
	Mach Number	$Mach$	$\mathcal{U}(0.78, 0.82)$
Engine	Bleed Pressure (kPa)	P_1	$\mathcal{U}(250, 12.5)$
	Engine Fan/Compressor Efficiency	η_{ec}	$\mathcal{T}ri(0.8, 0.98, 0.9)$
Overall Parameters	Conditioning Mass Flow Rate (kg/s)	$\dot{m}$	$\mathcal{U}(0.65, 0.75)$
	Cooling Ratio	r	$\mathcal{U}(0.315, 0.385)$
Pre-cooler	Length (m)	L_{x_1}	$\mathcal{U}(0.09, 0.11)$
	Width (m)	L_{y_1}	$\mathcal{U}(0.27, 0.33)$
	Height (m)	L_{z_1}	$\mathcal{U}(0.3852, 0.4708)$
	Wall Thickness (m)	t_{W1}	$\mathcal{U}(5.4e-4, 6.6e-4)$
	Sheet Fin Thermal Conductivity (W/(m·K))	k_{W1}	$\mathcal{MG}(20,22,0.5,0.5,0.6,0.4)$
Pre-cooler Heat Transfer Surface (Main Stream side)	Plate spacing (m)	b_1	$\mathcal{U}(4.69e-3, 5.73e-3)$
	Hydraulic Diameter (m)	$4r_{h1}$	$\mathcal{U}(1.38e-3, 1.69e-3)$
	Fin Thickness (m)	δ_1	$\mathcal{U}(0.92e-4, 1.12e-4)$
	Heat Transfer Area/Volume Between Plates (m^2/m^3)	β_1	$\mathcal{N}(2231, 111.55)$
	Fin Area/Total Area	$(A_f/A)_1$	$\mathcal{MG}(0.8,0.9,0.03,0.03,0.59,0.41)$
Pre-cooler Heat Transfer Surface (Ram Air side)	Plate Spacing (m)	b_{1_r}	$\mathcal{U}(11.07e-3, 13.53e-3)$
	Hydraulic Diameter (m)	$4r_{h1_r}$	$\mathcal{U}(3.07e-3, 3.75e-3)$
	Fin Thickness (m)	δ_{1_r}	$\mathcal{U}(0.92e-4, 1.12e-4)$
	Heat Transfer Area/Volume Between Plates (m^2/m^3)	β_{1_r}	$\mathcal{N}(1115, 55.75)$
	Fin Area/Total Area	$(A_f/A)_{1_r}$	$\mathcal{MG}(0.8,0.9,0.03,0.03,0.38,0.62)$
Main Heat Exchanger	Length (m)	L_{x_2}	$\mathcal{U}(0.09,0.11)$
	Width (m)	L_{y_2}	$\mathcal{U}(0.315, 0.385)$
	Height (m)	L_{z_2}	$\mathcal{U}(0.2781, 0.3399)$
	Wall Thickness (m)	t_{W2}	$\mathcal{U}(5.4e-4, 6.6e-4)$
	Sheet Fin Thermal Conductivity (W/(m·K))	k_{W2}	$\mathcal{MG}(20,22,0.5,0.5,0.6,0.4)$
Main Heat Exchanger Heat Transfer Surface (Main Stream side)	Plate spacing (m)	b_2	$\mathcal{U}(4.69e-3, 5.73e-3)$
	Hydraulic Diameter (m)	$4r_{h2}$	$\mathcal{U}(1.38e-3, 1.69e-3)$
	Fin Thickness (m)	δ_2	$\mathcal{U}(0.92e-4, 1.12e-4)$
	Heat Transfer Area/Volume Between Plates (m^2/m^3)	β_2	$\mathcal{N}(2231, 111.55)$
	Fin Area/Total Area	$(A_f/A)_2$	$\mathcal{MG}(0.8,0.9,0.03,0.03,0.59,0.41)$
Main Heat Exchanger Heat Transfer	Plate spacing (m)	b_{2_r}	$\mathcal{U}(11.07e-3, 13.53e-3)$
	Hydraulic Diameter (m)	$4r_{h2_r}$	$\mathcal{U}(3.07e-3, 3.75e-3)$
	Fin Thickness (m)	δ_{2_r}	$\mathcal{U}(0.92e-4, 1.12e-4)$

Surface (Ram Air side)	Heat Transfer Area/Volume Between Plates (m^2/m^3)	β_{2_r}	$\mathcal{N}(1115, 55.75)$
	Fin Area/Total Area	$(A_f/A)_{2_r}$	$\mathcal{MG}$(0.8,0.9,0.03,0.03,0.38,0.62)
Miscellaneous	Main Stream Diffuser Efficiency	η_d	$\mathcal{T}ri(0.85, 0.97, 0.95)$
	Air Cycle Machine Compressor Efficiency	η_c	$\mathcal{T}ri(0.7, 0.8, 0.75)$
	Air Cycle Machine Turbine Efficiency	η_t	$\mathcal{T}ri(0.7, 0.9, 0.8)$
	Ram Air Section (m^2)	A_i	$\mathcal{U}(1.08e-2, 1.32e-2)$
	Ram Air Diffuser Efficiency	η_{rd}	$\mathcal{T}ri(0.85, 0.97, 0.95)$
	Ram Air Nozzle Efficiency	η_n	$\mathcal{T}ri(0.85, 0.97, 0.95)$

Table 6-4. Part of output variables

Locations	**Outputs**	**Symbols**
Main stream	Bleed Temperature (K)	T_1
	Temperature at the Exit of Pre-cooler (K)	T_2
	Temperature after the Compressor (K)	T_3
	Temperature at the Exit of Main Heat Exchanger (K)	T_4
	Temperature after the Turbine (to the Cabin) (K)	T_5
	Pressure at the Exit of Pre-cooler (kPa)	P_2
	Pressure after the Compressor (kPa)	P_3
	Pressure at the Exit of Main Heat Exchanger (kPa)	P_4
	Pressure after the Turbine (to the Cabin) (kPa)	P_5
Ram Air	Temperature at the Exit of Ram Air Diffuser (K)	T_{1r}
	Temperature at the Exit of Main Heat Exchanger (K)	T_{2r}
	Temperature at the Exit of Pre-cooler (K)	T_{3r}
	Temperature at the Exit of Ram Air Nozzle (K)	T_{4r}
	Pressure at the Exit of Ram Air Diffuser (kPa)	P_{1r}
	Pressure at the Exit of Main Heat Exchanger (kPa)	P_{2r}
	Pressure at the Exit of Pre-cooler (kPa)	P_{3r}
	Pressure at the Exit of Ram Air Nozzle (kPa)	P_{4r}
Overall Performance	Entropy Generation Rate	N_S
	Total Volume of Heat Exchangers (m^3)	V_T

6.4.2.2 Results

Given the complexity of the test case, it was deemed infeasible to calculate the theoretical Sobol' indices. Thus the results of the proposed method are compared with reference values obtained from the traditional (benchmark) Monte-Carlo approach.

Also as it is impractical to show the different sets of Sobol' indices for all the 19 output variables, only N_s is chosen for illustration purposes. In the original paper (Pérez-

Grande and Leo, 2002), this output was used as the objective function in the design optimization.

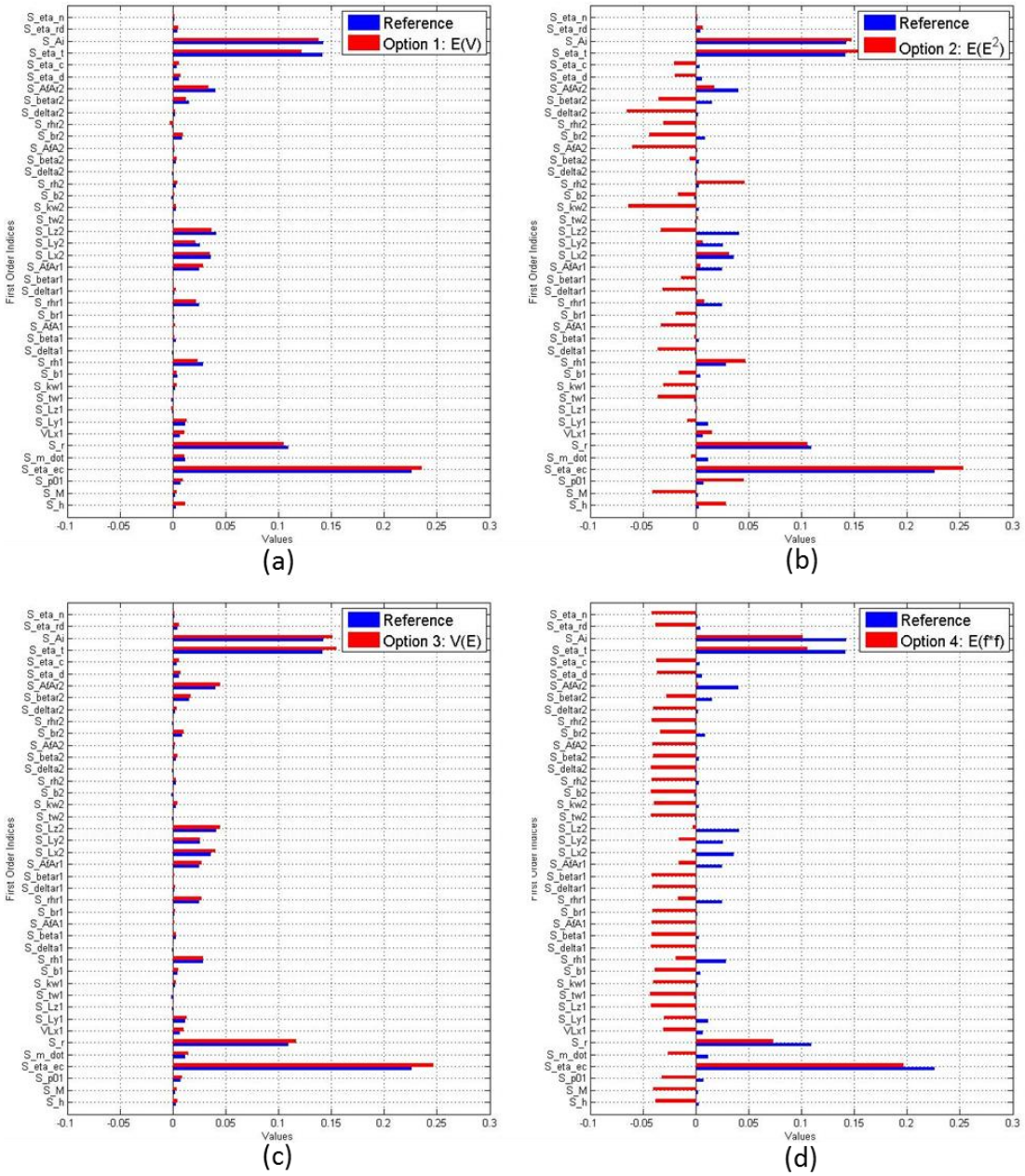

Figure 6-12. First order indices calculated from the four options, compared with the reference values

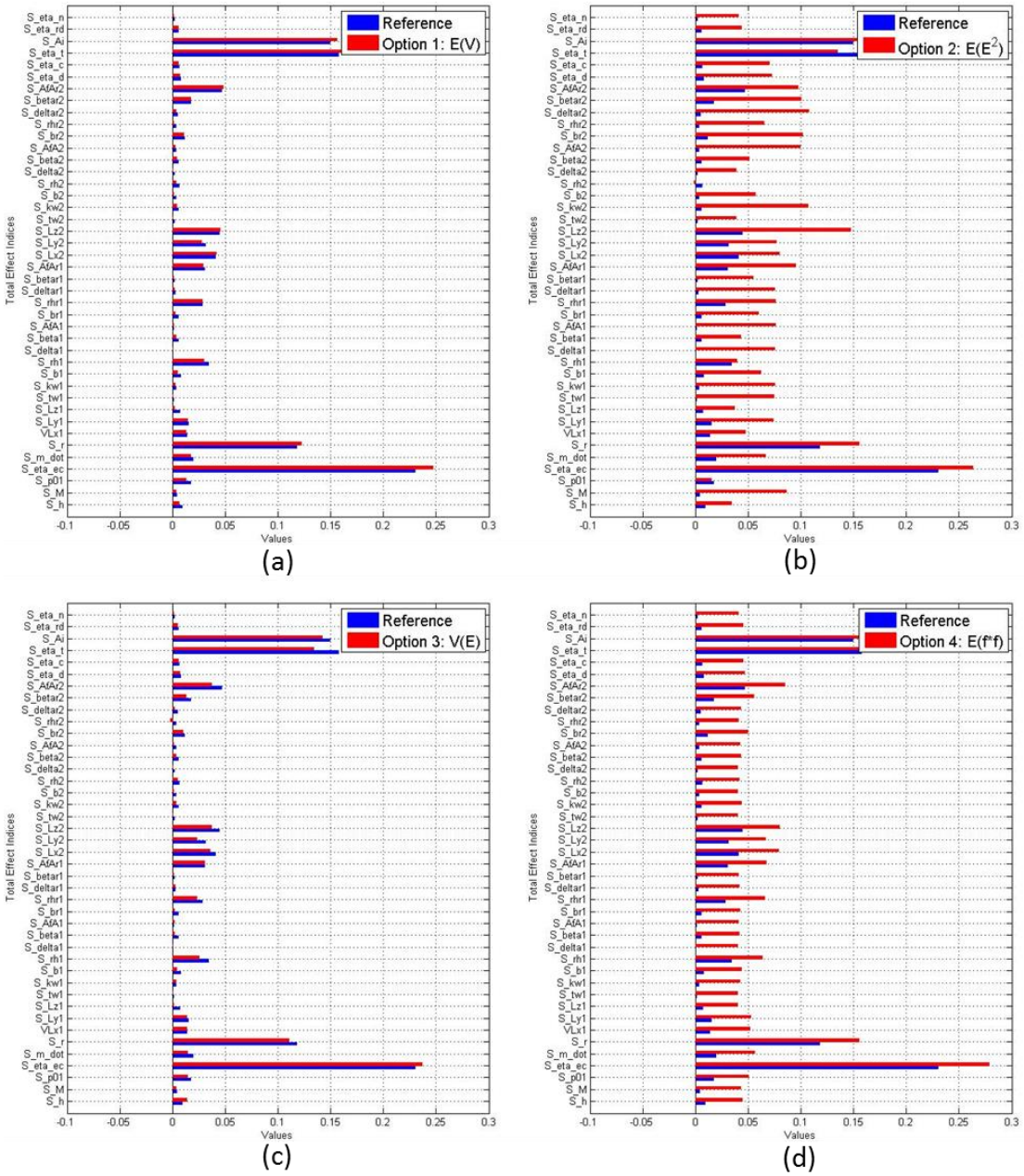

Figure 6-13. Total indices calculated from the four options, compared with the reference values

Shown in Figure 6-12 and Figure 6-13 are the first order and the total effect indices calculated as per the four options from Table 6-1 (marked as red in each subplot), compared with the reference values from MCS (marked as blue). It can be seen that Option 1 and 3 produced relatively good matches with the reference values, for both first order (Figure 6-12 (a) & (c)) and total effect indices (Figure 6-13 (a) & (c)). While Option 2 and 4 would sometimes give negative values for S_i (Figure 6-12 (b) & (d)) and

over-estimated values for S_i^T (Figure 6-13 (b) & (d))). The discrepancies are related mostly to the estimation of the insignificant factors, while the most influential ones are captured well.

The interaction effects in this model turned out to be insignificant. Comparison of the results are not shown here, because all the values (both the calculated and reference ones) are very close to zero.

6.4.2.3 Discussion

Figure 6-14 illustrates the efficiency and the effectiveness of the proposed method with regard to the practical test case. The solid blue line represents the first order index for cooling ratio (S_r) calculated from the traditional (benchmark) MCS approach, at 43000, 215000, 430000, 860000, and 2150000 model evaluations. The red error bars indicate 95% confidence intervals of the estimated values, obtained by using a boot-strap method from (Archer, Saltelli and Sobol, 1997). The straight black dash line is the value calculated by the proposed method (Option 1) at a computational cost of 3529 model evaluations. This is 623 times faster (compared with 2150000), at similar accuracy.

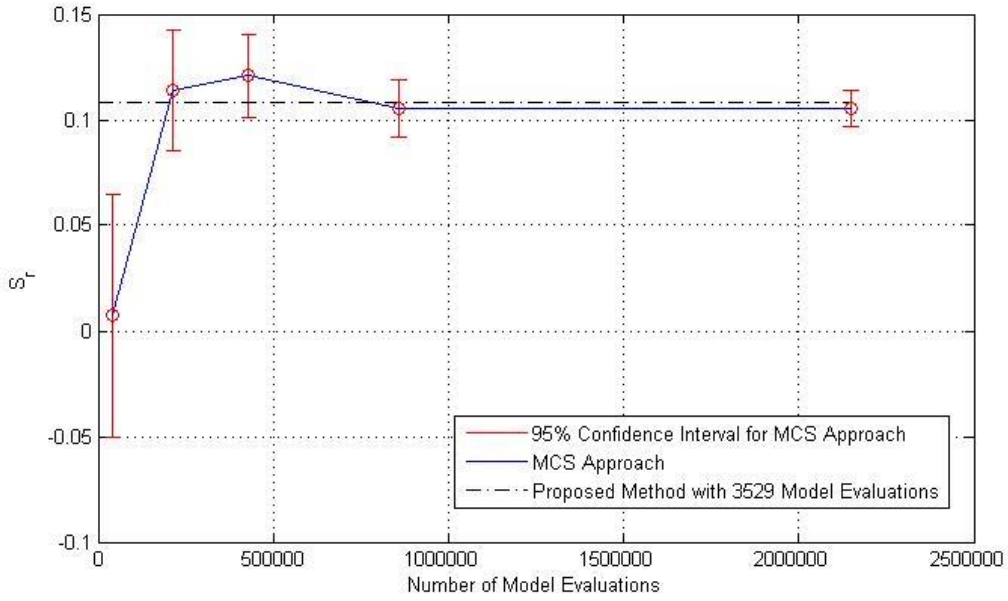

Figure 6-14. Convergence of the first order index of cooling ratio

Comparison with other available techniques has not been conducted yet. However it can be seen from Figure 6-14 that the result of the proposed method is already within the

95% confidence interval of the benchmark MCS study. This means that the other techniques should also provide similar results (assuming that all the methods are accurate).

On the other hand, the cost of each technique can be predicted analytically using the equations in Table 3-1. The computational costs of various methods are plotted again in Figure 6-15, with comparison to the proposed method. In general, the proposed method is comparable to the RBD (Mara and Rakoto Joseph, 2008), Bayesian approach (Oakley and O'Hagan, 2004), intrusive PCE (Sudret, 2008) with low truncation order, and non-intrusive PCE (using LSA (Cuneo, Traverso and Shahpar, 2017), low-rank tensor (Konakli and Sudret, 2016) and sparse grid (Xiu and Hesthaven, 2005)). One known limitation of the intrusive PCE is that it requires the modification of the original model, which will become prohibitive if the model is complex or even a black-box (Hosder, Walters and Perez, 2006). With regard to other methods, as their exact computational costs are also dependent on the required accuracy, a further comparison is planned for future work.

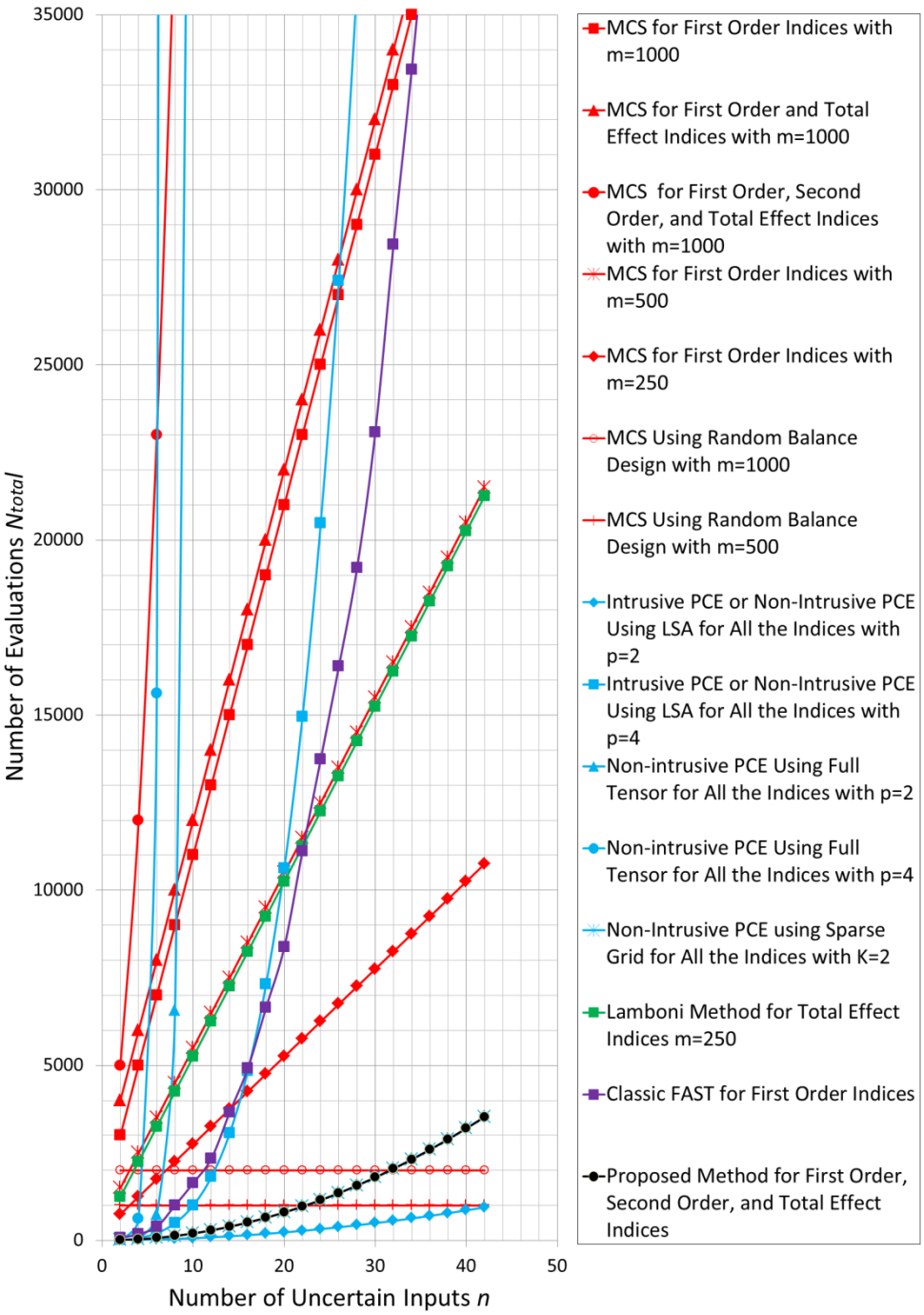

Figure 6-15. Number of evaluations required for different techniques

6.5 Summary and Conclusions

Presented in this paper is a method for efficient variance based sensitivity analysis. The main contribution can be considered in two levels. The first level is a general approach to transform the sensitivity analysis into an uncertainty propagation process, so that

various existing approximation techniques can be implemented to speed up the calculation. In the second level, formulations are proposed to implement the specific URQ technique, with four different options.

As demonstrated through numerical experiments, the proposed method is very accurate in all linear or monotonic cases, under various probability distributions. For most non-linear and non-monotonic conditions, this method still provides fairly accurate estimations of the first order and total effect indices. In general, Option 3 has the highest accuracy and robustness, followed by Option 1 and Option 2 which are close to each other. Option 4 is the least robust to non-linearity. Compared with other techniques, the particular strength is on the low computational cost, especially for large scale problems. This would enable the decision maker to quickly identify the most influential factors.

On the other hand, there are limitations of the specific method (implemented with URQ). Interaction effects in the computational models will cause a decrease in the accuracy. The higher order indices are not yet usable and their computation needs further improvement. Apart from that, certain combinations of non-linearity, non-monotonicity, and shapes of the distributions will restrict the application of the current method. In particular:

- Periodicity of the model, plus input distributions with span larger than half of the model period, will cause aliasing.
- Multiplicative models with constant output values along the axes, plus the sampling points corresponding to the input distributions happen to be on the axes, will lead to disregard of the variance.

In general, the proposed method could be used as a first assessment to identify the most influential factors, which will reduce the dimensionality of the original problem. If needed, further sensitivity analysis with more accurate, but computationally expensive methods such as MSC or FAST could be conducted within the reduced set of most influential factors.

Future work will focus on three tasks: the first is to further develop the specific method implemented with URQ. This will involve more rigorous and explicit representation of the numerical errors. The second is to implement the general approach with other

approximation techniques. Some of the candidates may include: the unscented transformation [19], Gaussian Quadrature [20], generalised Taguchi method [21], etc. These implementations will also be compared with current techniques for sensitivity analysis. The third task is to organise these specific technique as a comprehensive system to perform sensitivity analysis, adapting to different types of computational models.

7 Summary and Conclusions

7.1 Summary of Research

The presented research was motivated by the industrial need for novel and improved UQ&M techniques, in response to the development of computational design for aerospace systems. Three specific problems were identified, as a result from the literature review and requirements of the European research project of TOICA.

Objective 1 is motivated by the TOICA project, for an improved margin management policy. Margins have already been applied extensively in engineering practice. The project need is to integrate margins with UQ&M techniques, and to enable systematic trade-off between margins, performances, and probabilities of constraint satisfaction.

The corresponding result is a comprehensive method presented in Chapter 4. In this method, activities are classified from the perspective of the designer and the uncertainty expert, so that analysis could be conducted simultaneously. From the UQ&M perspective, sensitivity analysis and propagation are used to support the allocation of deterministic margins. On the other hand, different types of trade-off (decision making) are identified and organised as three categories: margins versus margins, margins versus probability of constraint satisfaction, and margins versus performance. Several enablers have been developed to facilitate each of the trade-off studies. The overall approach is constructed with the philosophy of set-based design, which enables the designer to explore multiple solutions. In this process, design knowledge is gained regarding the design/margin space by assess explicitly the interaction between margins and other quantities.

The method is demonstrated with an aircraft sizing problem which has two levels of models (aircraft level and subsystem level). In the test-case, the designer can interactively explore and trace the influence of each margin, and gradually narrow-down the set of solutions based on the available information. This test-case was implemented in AirCADia, and partially evaluated in the TOICA project.

Objective 2 also initiates from the TOICA project, for a method to ‘allocate’ uncertainty in the design process. This is based on the assumption that part of the uncertainty involved in the calculation could be reduced by advanced but more expensive design

studies. Therefore it is beneficial for the designer to know how much uncertainty could be tolerated from different sources, given the acceptable level of uncertainty associated with the system outputs.

Regarding this objective, a method is proposed for inverse uncertainty propagation as presented in Chapter 5. The method is enabled by constructing an optimization loop over a low-cost forward propagation process. Specifically, it is implemented with the URQ technique and the workflow reversal capability of AirCADia. By applying inverse propagation, the designer can specify the target variances of output variables, and inversely calculate those of the inputs.

For validation of the method, the inverse propagation was tested with both analytical examples and an aircraft sizing case study. The results of these test-cases have proven the accuracy of the proposed method. In the TOICA project, the proposed method was also applied on several other industrial test-cases, with which, the usefulness was evaluated and confirmed through the project review process.

Objective 3 is regarding the general need in this field, to reduce the computational cost of implementing UQ&M techniques. Specifically, sensitivity analysis is chosen as the problem area of interest. The reason is that sensitivity analysis is one of the most computationally demanding treatments of uncertainty, which requires more efficient methods, especially for high dimensional problems. Furthermore, by applying sensitivity analysis, computational cost of other UQ&M practice can be reduced as well, due to the reduction of problem dimensions.

The method proposed in Chapter 6 corresponds to achieving this objective. The development can be divided into two levels: The first level is a general approach to transfer the sensitivity analysis into an uncertainty propagation process, so that various existing approximation techniques could be implemented to speed up the calculation. In the second level, specific formulations are proposed to integrate the URQ technique with variance based sensitivity indices, as one realization of the aforementioned general approach.

For evaluation, the proposed method (URQ version) was applied on several test-cases, which include a series of analytical examples and an aircraft ECS case study. The

former were used to explore the applicability of the proposed method, regarding models with representative mathematical properties. The latter was applied to demonstrate the efficiency, on a practical design problem. For all the linear or monotonic cases, the proposed method achieved very high accuracy, under various probability distributions. For most non-linear and non-monotonic conditions, this method still provides fairly accurate estimations of the first order and total effect indices. In the practical test-case, the result is very close to the benchmark, and the efficiency of the proposed method is demonstrated by comparing the computational cost with other sensitivity analysis techniques.

7.2 Contribution to Knowledge

The contribution to knowledge is reflected on the novel methods developed in this PhD research, which include:

- A method for margin allocation and trade-off, which enhances the management of traditional margins, in a more explicit, systematic, and interactive manner.
- A method for inverse uncertainty propagation, which enables the designer to manage uncertainty interactively.
- A method for sensitivity analysis, which has significantly reduced the relevant computational cost. Furthermore, a general approach is provided, which enables various other implementations to improve the efficiency.

7.3 Current Limitations

The limitations of this PhD research can be considered in two levels. In the narrow sense, the limitations are related to detailed implementation of each proposed method:

- In margin allocation, full factorial design of experiment is applied for its simplicity. However, computational cost would grow significantly if the designer wants to treat a large number of margins simultaneously.
- The URQ technique is restricted by highly non-linearity of computational model. As a result, the accuracy of the proposed methods for sensitivity analysis and inverse propagation can be affected.

In a broader sense, it has been pointed out that uncertainty exists both *Within the Problem* and *About the Problem* as illustrated by Figure 7-1 (Padulo and Guenov, 2012). This research has been focused on the computational perspective which is largely within the problem. Behind the scenes of all the computation, it is the design knowledge, which reflects not only on the models, but also in the way that design problems are formulated. For instance, in the research and development (R&D) of novel concepts and technologies, (such as blended wind body airliners, more electric aircraft systems, etc.), due to the lack of experience and routines to follow, uncertainty about problem might be more influential. Such types of uncertainty might be non-parametric in representation, therefore are beyond the current probabilistic approach.

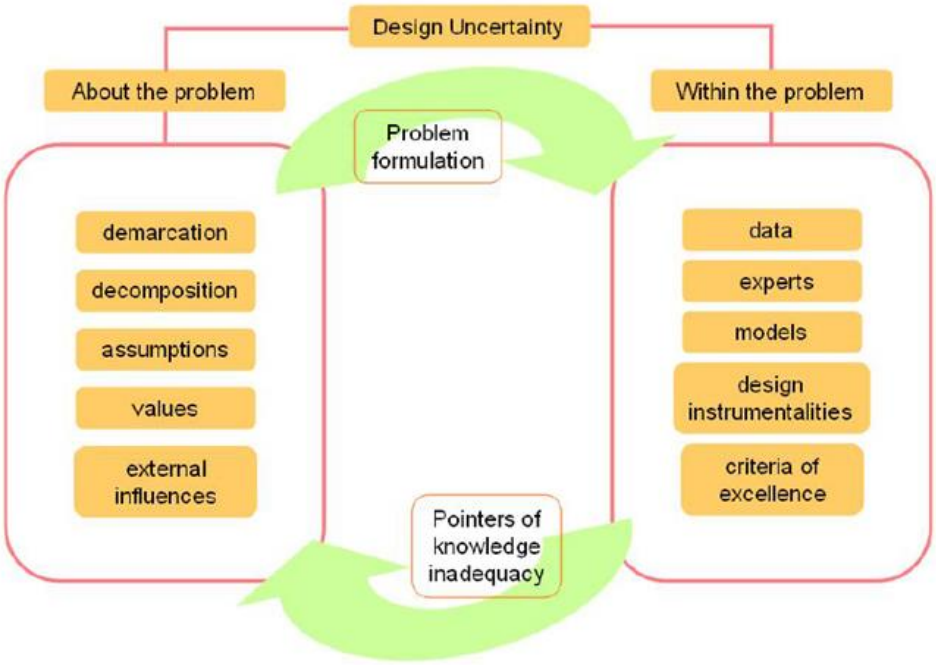

Figure 7-1. A classification of uncertainty in Engineering Design (Padulo and Guenov, 2012)

7.4 Future Work

Future research needs to consider not only the limitations as discussed in the previous section, but also development of new capabilities. Specifically:

- Regarding margin allocation, more efficient sampling techniques (e.g. Latin Hyper Cube or Sobol' sequence) can be used, to improve the scalability of the current method. Additionally, the method will consider the margin

decomposition from the top-down in the design hierarchy, and the evolution of margins during the design process.

- Regarding inverse propagation, further investigation should be conducted on the solver robustness. Also, methods can be developed to handle the under-determined and over-determined cases, in response to the broader problem of uncertainty allocation.
- The method of sensitivity analysis can be improved, by implementing the approach with other approximation techniques. In addition, investigation will be conducted on more rigorous and explicit representation of the numerical errors.

Regarding the broader picture of uncertainty *Within the Problem* and *About the Problem*, more advance mathematical theories should be adopted, for example the Generalized Information Theory (GIT) (Helton, Johnson and Oberkampf, 2004; Klir, 2005; Klir and Smith, 2001; Klir and Wierman, 1999). To implement this mathematical framework, some problem areas may include:

- Formalised representation of design knowledge, including assumptions, expert judgements, design routines, etc.
- Quantifying the uncertainty associated with these pieces of information, using elicitation method.
- Selecting proper mathematical tools for epistemic uncertainty analysis, for instance Possibility Theory (Zadeh, 1999), Evidence Theory (Dempster, 1967; Shafer, 1976; Zadeh, 1986), etc.
- Aggregation of the uncertainty with a unified measure.

References

Abusam, A. et al. (2003) ‘Forward and backward uncertainty propagation: an oxidation ditch modelling example’, *Water Research*, 37(2), pp. 429–435.

Antony, J. (2003) *Design of Experiments for Engineers and Scientists*.1st edn. London: Butterworth-Heinemann.

Archer, G.E.B. et al. (1997) ‘Sensitivity measures, anova-like techniques and the use of bootstrap’, *Journal of Statistical Computation and Simulation*, 58(2), pp. 99–120.

Aughenbaugh, J.M. (2006) *Managing Uncertainty in Engineering Design Using Imprecise Probabilities and Principles of Information Economics*. PhD thesis. Georgia Institute of Technology.

Ayyub, B.M. (2001) *Elicitation of expert opinions for uncertainty and risks*. Boca Raton: CRC Press.

Ayyub, B.M. and Klir, G.J. (2006) *Uncertainty modeling and analysis in engineering and the sciences*. Boca Raton: Chapman & Hall/CRC.

Balachandran, L.K. (2007) *Computational workflow management for conceptual design of complex systems: an air-vehicle design perspective*. PhD thesis. Cranfield University.

Balachandran, L.K. and Guenov, M.D. (2010) ‘Computational Workflow Management for Conceptual Design of Complex Systems’, *Journal of Aircraft*, 47(2), pp. 699–703.

Baumgärtel, P. et al. (2014) ‘Inverse uncertainty propagation for demand driven data acquisition’, *Proceedings of the Winter Simulation Conference 2014*. Savannah, GA, USA, 7-10 December 2014. Institute of Electrical and Electronics Engineers, pp. 710–721.

Bedford, T. (1998) ‘Sensitivity Indices for (Tree-) Dependent Variables’, *Proceedings of the Second International Symposium on Sensitivity Analysis of Model Output (SAMO98)*. Venice, Italy, 19–22 April 1998. Luxembourg: Commission of the European Committee, pp. 17–20.

Birman, J. (2013) *Uncertainty quantification and propagation in Conceptual Aircraft*

Design: from deterministic optimisation to chance constrained optimisation. PhD Thesis. University of Toulouse.

Blatman, G. and Sudret, B. (2010) 'Efficient computation of global sensitivity indices using sparse polynomial chaos expansions', *Reliability Engineering & System Safety*, 95(11), pp. 1216–1229.

Blochwitz, T. et al. (2011) 'The Functional Mockup Interface for Tool independent Exchange of Simulation Models', *Proceedings of the 8th International Modelica Conference*. Dresden, Germany, 20-22 March 2011. Linköping: Linköping University Electronic Press, pp. 105–114.

Browne, T. et al. (2017) *hal-01450891: Estimate of quantile-oriented sensitivity indices*. Hyper Articles en Ligne (HAL). Available at: https://hal.archives-ouvertes.fr/hal-01450891/document (Accessed: 25 February 2018).

Chakraborty, I. et al. (2014) 'A Requirements-driven Methodology for Integrating Subsystem Architecture Sizing and Analysis into the Conceptual Aircraft Design Phase', *14th AIAA Aviation Technology, Integration, and Operations Conference, AIAA AVIATION Forum*. Atlanta, GA, USA, 16-20 June 2014. Reston: American Institute of Aeronautics and Astronautics, AIAA 2014-3012.

Chase, K.W. (1999) *ADCATS Report No. 99-6: Tolerance allocation methods for designers*. Association for the Development of Computer-Aided Tolerancing Systems (*ADCATS*). Available at: http://adcats.et.byu.edu/Publication/99-6/ToleranceAlloc.pdf (Accessed: 25 February 2018).

Chen, W. et al. (1999) 'Quality Utility—A Compromise Programming Approach to Robust Design', *Journal of Mechanical Design*, 121(2), pp. 179–187.

Cherry, H.H. and Croshere, A.B. (1948) *SAE Technical Paper 480169: An Approach to the ANALYTICAL Design of Aircraft*. Society of Automotive Engineers (SAE).

Congedo, P.M. et al. (2012) 'Backward uncertainty propagation method in flow problems: Application to the prediction of rarefaction shock waves', *Computer Methods in Applied Mechanics and Engineering*, 213–216, pp. 314–326.

Cooke, R.M. (1991) *Experts in Uncertainty : Opinion and Subjective Probability in Science: Opinion and Subjective Probability in Science*. New York: Oxford University Press.

Cooke, R.M. et al. (2015) 'Sculpting: A Fast, Interactive Method for Probabilistic Design Space Exploration and Margin Allocation', *16th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*. Dallas, TX, USA, 22-26 June 2015. Reston: American Institute of Aeronautics and Astronautics, AIAA 2015-3440

Cramer, E.J. (1995) 'Airplane performance optimization - A design case study', *33rd Aerospace Sciences Meeting and Exhibit*. Reno, NV, USA, 9-12 January 1995. American Institute of Aeronautics and Astronautics, AIAA-95-0465

Crestaux, T. et al. (2009) 'Polynomial chaos expansion for sensitivity analysis', *Reliability Engineering and System Safety*, 94(7), pp. 1161–1172.

Cross, N. (1993) 'Science and design methodology: A review', *Research in Engineering Design*, 5(2), pp. 63–69.

Cukier, R.I. (1973) 'Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory', *The Journal of Chemical Physics*, 59(8) AIP Publishing, p. 3873.

Cukier, R.I. et al. (1978) 'Nonlinear sensitivity analysis of multiparameter model systems', *Journal of Computational Physics*, 26(1), pp. 1–42.

Cukier, R.I. et al. (1975) 'Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. III. Analysis of the approximations', *The Journal of Chemical Physics*, 63(3), pp. 1140–1149.

Cuneo, A. et al. (2017) 'Comparative Analysis of Methodologies for Uncertainty Propagation and Quantification', *ASME Turbo Expo 2017: Turbomachinery Technical Conference and Exposition*. Charlotte, NC, USA, 26-30 June 2017. American Society of Mechanical Engineers, GT2017-63238.

Datta, V.C. (2014) *Interactive Computational Model-Based Design: A Blackbox Perspective*. PhD Thesis. Cranfield University.

de Tenorio, C. (2010) Methods for collaborative conceptual design of aircraft power architectures. PhD Thesis. Georgia Institute of Technology.

Dempster, A.P. (1967) 'Upper and Lower Probabilities Induced by a Multivalued Mapping', *The Annals of Mathematical Statistics*, 38(2), pp. 325–339.

Driggs, I.H. (1950) 'Aircraft Design Analysis', *Journal of the Royal Aeronautical Society*, 54(470), pp. 65–116.

Du, X. and Chen, W. (2002) 'Efficient Uncertainty Analysis Methods for Multidisciplinary Robust Design', *AIAA Journal*, 40(3), pp. 545–552.

Eldred, M. and Burkardt, J. (2009) 'Comparison of Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Quantification', *47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition*. Orlando, FL, USA, 05-08 January 2009. Reston: American Institute of Aeronautics and Astronautics, AIAA 2009-976.

Evans, D.H. (1972) 'An Application of Numerical Integration Techniques to Statistical Tolerancing, III—General Distributions', *Technometrics*, 14(1), pp. 23–35.

Fonseca, J.R. et al. (2005) 'Uncertainty identification by the maximum likelihood method', *Journal of Sound and Vibration*, 288(3), pp. 587–599.

Fort, J.-C. et al. (2016) 'New sensitivity analysis subordinated to a contrast', *Communications in Statistics - Theory and Methods*, 45(15), pp. 4349–4364.

Ghanem, R.G. and Spanos, P.D. (1991) *Stochastic Finite Elements: A Spectral Approach*. Revised ed. New York: Springer-Verlag New York.

Girard, A. and Murray-Smith, R. (2005) 'Gaussian Processes: Prediction at a Noisy Input and Application to Iterative Multiple-Step Ahead Forecasting of Time-Series', in Murray-Smith, R. and Shorten, R. (eds.) *Switching and Learning in Feedback Systems*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 158–184.

Guenov, M.D. et al. (2017) 'Margin Allocation and Trade-off in Complex Systems Design and Optimization', *(paper accepted for publication in AIAA Journal)*.

Guenov, M.D. et al. (2010) ‘MDO at Predesign Stage’, in Guenov, M. D. and Kesseler, E. (eds.) *Advances in Collaborative Civil Aeronautical Multidisciplinary Design Optimization*. Reston: American Institute of Aeronautics and Astronautics, pp. 17–71.

Guenov, M.D. et al. (2006) ‘Computational design process modeling’, *25th Congress of the International Council of Aeronautical Sciences conferences (ICAS)*. Hamburg, Germany, 03-08 September 2006. International Council of The Aeronautical Sciences. Available at: http://www.icas.org/ICAS_ARCHIVE/ICAS2006/PAPERS/172.PDF (Accessed: 25 February 2018).

Guenov, M.D. et al. (2011) ‘Comparing Design Margins in Robust and Deterministic Design Optimization’, *Eurogen Conference*. Capua, Italy, 14-16 September 2011.

Guenov, M.D. et al. (2014a) ‘Aircadia - an interactive tool for the composition and exploration of aircraft computational studies at early design stage’, *29th Congress of the International Council of the Aeronautical Sciences (ICAS)*. St. Petersburg, Russia, 07-12 Septemper 2014. International Council of The Aeronautical Sciences. Available at: http://studygroup.riskinstitute.org.uk/docs/ICAS2014Last.pdf (Accessed: 25 February 2018).

Guenov, M.D. et al. (2014b) ‘Composition, Management, and Exploration of Computational Studies at Early Design Stage’, in Vasile, M. and Becerra, V. M. (eds.) *Computational Intelligence in Aerospace Sciences, Progress in Astronautics and Aeronautics*. Reston: American Institute of Aeronautics and Astronautics, pp. 415–460.

Haupt, U. (1977) ‘Case studies in aircraft design’, *Aircraft Systems and Technology Meeting*. Seattle, WA, USA, 22-24 August 1977. Reston: American Institute of Aeronautics and Astronautics.

Helton, J.C. (2009) *SAND2009-3055: Conceptual and computational basis for the quantification of margins and uncertainty*. Sandia National Laboratories. Available at: http://prod.sandia.gov/techlib/access-control.cgi/2009/093055.pdf (Accessed: 25 February 2018).

Helton, J.C. and Davis, F.J. (2003) ‘Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems’, *Reliability Engineering & System Safety*,

81(1), pp. 23–69.

Helton, J.C. et al. (2004) ‘An exploration of alternative approaches to the representation of uncertainty in model predictions’, *Reliability Engineering & System Safety*, 85(1–3) Elsevier, pp. 39–71.

Hicks, R. and Henne, P. (1978) ‘Wing design by numerical optimization’, *Journal of Aircraft,*15(7), pp. 407-412.

Homma, T. and Saltelli, A. (1996) ‘Importance measures in global sensitivity analysis of nonlinear models’, *Reliability Engineering & System Safety*, 52(1), pp. 1–17.

Hora, S.C. and Iman, R.L. (1986) *SAND-85-2839: Comparison of Maximus/Bounding and Bayes/Monte Carlo for fault tree uncertainty analysis*. Sandia National Laboratories.

Hosder, S. et al. (2006) ‘A Non-Intrusive Polynomial Chaos Method For Uncertainty Propagation in CFD Simulations’, *44th AIAA Aerospace Sciences Meeting and Exhibit, Aerospace Sciences Meetings*. Reno, NV, USA, 09-12 January 2006. Reston: American Institute of Aeronautics and Astronautics, AIAA 2006-891.

Huyse, L. and Bushnell, D.M. (2001) *NASA/CR-2001-211020: Free-form Airfoil Shape Optimization Under Uncertainty Using Maximum Expected Value and Second-order Second-moment Strategies*. National Aeronautics and Space Administration (NASA). Available at: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20010067587.pdf (Accessed: 25 February 2018).

Iman, R.L. and Hora, S.C. (1990) ‘A Robust Measure of Uncertainty Importance for Use in Fault Tree System Analysis’, *Risk Analysis*, 10(3), pp. 401–406.

Inselberg, A. (2009) *Parallel coordinates : visual multidimensional geometry and its applications*. New York: Springer.

Ishigami, T. and Homma, T. (1990) ‘An importance quantification technique in uncertainty analysis for computer models’, *Proceedings. First International Symposium on Uncertainty Modeling and Analysis*. Maryland, USA, 03-05 December 1990. IEEE Comput. Soc. Press, pp. 398–403.

Jansen, M.J. et al. (1994) ‘Monte carlo estimation of uncertainty contributions from several independent multivariate sources’, in Grasman, J. and van Straten, G. (eds.) *Predictability and Nonlinear Modelling*. Dordrecht: Springer Netherlands, pp. 334–343.

Jansen, M.J.W. (1999) ‘Analysis of variance designs for model output’, *Computer Physics Communications*, 117(1–2), pp. 35–43.

Jenkinson, L.R. et al. (1999) *Civil jet aircraft design*.1st edn. London: Arnold.

Jiang, C. et al. (2008) ‘A Novel Method for Uncertainty Inverse Problems and Application to Material Characterization of Composites’, *Experimental Mechanics*, 48(4), pp. 539–548.

Kaykobad, M. (1985) ‘Positive solutions of positive linear systems’, *Linear Algebra and its Applications*, 64, pp. 133–140.

Keane, A. and Nair, P. (2005) *Computational approaches for aerospace design: the pursuit of excellence*. Chichester: John Wiley & Sons.

Kennedy, J. and Eberhart, R. (1995) ‘Particle swarm optimization’, *Proceedings of ICNN’95 - International Conference on Neural Networks*. Perth, WA, Australia, 27 November-01 December 1995. Institute of Electrical and Electronics Engineers, Vol.4, pp. 1942–1948.

Kennedy, M.C. and O’Hagan, A. (2001) ‘Bayesian calibration of computer models’, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3), pp. 425–464.

Kiureghian, A. Der and Ditlevsen, O. (2009) ‘Aleatory or epistemic? Does it matter?’, *Structural Safety*, 31(2), pp. 105–112.

Klir, G.J. (2005) *Uncertainty and Information: Foundations of Generalized Information Theory*. Hoboken: John Wiley & Sons.

Klir, G.J. and Smith, R.M. (2001) ‘On Measuring Uncertainty and Uncertainty-Based Information: Recent Developments’, *Annals of Mathematics and Artificial Intelligence*, 32, pp. 5–33.

Klir, G.J. and Wierman, M.J. (1999) *Uncertainty-Based Information: Elements of*

Generalized Information Theory. 2nd edn. Heidelberg: Physica-Verlag.

Konakli, K. and Sudret, B. (2016) 'Global sensitivity analysis using low-rank tensor approximations', *Reliability Engineering & System Safety*, 156, pp. 64–83.

Kucherenko, S. et al. (2017) 'Sobol' indices for problems defined in non-rectangular domains', *Reliability Engineering & System Safety*, 167, pp. 218–231.

Kucherenko, S. et al. (2012) 'Estimation of global sensitivity indices for models with dependent variables', *Computer Physics Communications*, 183(4), pp. 937–946.

Kullback, S. and Leibler, R.A. (1951) 'On Information and Sufficiency', *The Annals of Mathematical Statistics*, 22(1), pp. 79–86.

Lamboni, M. (2016) 'Global sensitivity analysis: an efficient numerical method for approximating the total sensitivity index', *International Journal for Uncertainty Quantification*, 6(1), pp. 1–17.

Lamboni, M. et al. (2013) 'Derivative-based global sensitivity measures: General links with Sobol' indices and numerical tests', *Mathematics and Computers in Simulation*, 87, pp. 45–54.

Lammering, T. (2014) *Integration of aircraft systems into conceptual design synthesis*. PhD Thesis. RWTH Aachen University.

Lee, S.H. and Chen, W. (2009) 'A comparative study of uncertainty propagation methods for black-box-type problems', *Structural and Multidisciplinary Optimization*, 37(3), pp. 239–253.

Litvinenko, A. and Matthies, H.G. (2013) *arXiv:1312.5048: Inverse problems and uncertainty quantification*. ArXiv. Available at: https://arxiv.org/abs/1312.5048 (Accessed: 25 February 2018).

Lööf, J. et al. (2007) 'An efficient solution to the discrete least-cost tolerance allocation problem with general loss functions', in Davidson, J. K. (ed.) *Models for Computer Aided Tolerancing in Design and Manufacturing*. Dordrecht: Springer Netherlands, pp. 115–124.

Lööf, J. and Söderberg, R. (2011) 'Discrete tolerance allocation for product families',

Engineering Optimization, 44(1), pp. 75–85.

Mangeant, F. (2011) 'Joined initiatives around uncertainty management Generic methodologies, mathematical challenges, and numerical implementations', *Annals of Telecommunications*, 66(7–8), pp. 397–407.

Mara, T.A. and Rakoto Joseph, O. (2008) 'Comparison of some efficient methods to evaluate the main effect of computer model factors', *Journal of Statistical Computation and Simulation*, 78(2), pp. 167–178.

Mara, T.A. and Tarantola, S. (2012) 'Variance-based sensitivity indices for models with dependent inputs', *Reliability Engineering & System Safety*, 107, pp. 115–121.

Mares, C. et al. (2002) 'Review of Parameter Uncertainty Propagation Methods in Structural Dynamic Analysis', *Proceedings of ISMA2002*. Leuven, Belgium, 16-18 September 2002. pp. 1853–1860.

Massey, F.J. and Jr. (1951) 'The Kolmogorov-Smirnov Test for Goodness of Fit', *Journal of the American Statistical Association*, 46(253), pp. 68–78.

Maume-Deschamps, V. and Niang, I. (2018) 'Estimation of quantile oriented sensitivity indices', *Statistics & Probability Letters*, 134, pp. 122–127.

Mccullers, L.A. (1984) 'Aircraft configuration optimization including optimized flight profiles', *Symposium on Recent Experiences in Multidisciplinary Analysis and Optimization*. NASA CP-2327, pp. 395-412

McLachlan, G.J. and Peel, D. (2000) *Finite mixture models*. New York: John Wiley & Sons.

McManus, H. and Hastings, D. (2005) 'A Framework for Understanding Uncertainty and its Mitigration and Exploitation in Complex Systems', *INCOSE International Symposium*, 15(1), pp. 484–503.

McRae, G.J. et al. (1982) 'Global sensitivity analysis—a computational implementation of the Fourier Amplitude Sensitivity Test (FAST)', *Computers & Chemical Engineering*, 6(1), pp. 15–25.

Miles Jr., R.F. (1993) 'Specifying design conservatism: Worst case versus probabilistic

analysis', *Proceedings of the Second International Symposium on Ground Data Systems for Space Mission Operations*. Pasadena, CA, USA. NASA, pp. 703–709.

Molina-Cristóbal, A. et al. (2018) 'Interactive Uncertainty Allocation and Trade-off at Early-stage Aircraft Computational Design', *2018 AIAA Non-Deterministic Approaches Conference, AIAA SciTech Forum.* Kissimmee, FL, USA, 08-12 January 2018. Reston: American Institute of Aeronautics and Astronautics, AIAA 2018-2170.

Molina-Cristóbal, A. et al. (2014) 'Black-Box Model Epistemic Uncertainty at Early Design Stage. An Aircraft Power-Plant Integration Case Study', *29th Congress of the International Council of the Aeronautical Sciences (ICAS)*. St. Petersburg, Russia, 07-12 Septemper 2014. International Council of The Aeronautical Sciences.

Narayanan, V.A.B. and Zabaras, N. (2004) 'Stochastic inverse heat conduction using a spectral approach', *International Journal for Numerical Methods in Engineering*, 60(9), pp. 1569–1593.

Ngnepieba, P. and Hussaini, M.Y. (2007) 'An efficient sampling method for stochastic inverse problems', *Computational Optimization and Applications*, 37(2), pp. 121–138.

Niang, I. (2016) *Quantification et méthodes statistiques pour le risque de modèle*. PhD Thesis. Université de Lyon.

Nunez, M. et al. (2012) 'Enabling exploration in the conceptual design and optimisation of complex systems', *Journal of Engineering Design*, 23(10–11) Taylor & Francis, pp. 852–875.

Nunez, M. et al. (2009) 'Integrated Exploration and Visualization of Optimal Aircraft Conceptual Designs', *50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. Palm Springs, CA, USA, 04-07 May 2009. Reston: American Institute of Aeronautics and Astronautics, AIAA 2009-2204.

O'Hagan, A. et al. (2006) *Uncertain Judgements: Eliciting Experts' Probabilities*. Chichester: John Wiley & Sons.

Oakley, J.E. and O'Hagan, A. (2004) 'Probabilistic sensitivity analysis of complex models: a Bayesian approach', *Journal of the Royal Statistical Society: Series B*

(Statistical Methodology), 66(3), pp. 751–769.

Oberkampf, W.L. et al. (2002a) 'Error and uncertainty in modeling and simulation', *Reliability Engineering & System Safety*, 75(3), pp. 333–357.

Oberkampf, W.L. et al. (2002b) 'Estimation of total uncertainty in modeling and simulation', *Reliability Engineering & System Safety*, 75, pp. 333–357.

Opgenoord, M.M.J. (2016) *Uncertainty budgeting methods for conceptual aircraft design*. MSc Thesis. Massachusetts Institute of Technology.

Opgenoord, M.M.J. and Willcox, K.E. (2016) 'Sensitivity Analysis Methods for Uncertainty Budgeting in System Design', *AIAA Journal*, 54(10), pp. 3134–3148.

Padulo, M. (2009) *Computational engineering design under uncertainty: an aircraft conceptual design perspective*. PhD Thesis. Cranfield University.

Padulo, M. et al. (2007) 'Comparative analysis of uncertainty propagation methods for robust engineering design', *International Conference on Engineering Design*. Paris, France, 28-31 August 2007.

Padulo, M. et al. (2011) 'Novel Uncertainty Propagation Method for Robust Aerodynamic Design', *AIAA Journal*, 49(3), pp. 530–543.

Padulo, M. and Guenov, M.D. (2011) 'Worst-case robust design optimization under distributional assumptions', *International Journal for Numerical Methods in Engineering*, 88(8), pp. 797–816.

Padulo, M. and Guenov, M.D. (2012) 'A Methodological Perspective on Computational Engineering Design under uncertainty', *European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS)*. Vienna, Austria, 10-14 September 2012.

Park, G.-J. et al. (2006) 'Robust Design: An Overview', *AIAA Journal*, 44(1), pp. 181–191.

Parkinson, A. et al. (1993) 'A General Approach for Robust Optimal Design', *Journal of Mechanical Design*, 115(1), p. 74.

Pérez-Grande, I. and Leo, T.J. (2002) 'Optimization of a commercial aircraft environmental control system', *Applied Thermal Engineering*, 22(17), pp. 1885–1904.

Phadke, M.S. (1989) *Quality engineering using robust design*. New Jersey: Prentice Hall.

Pilch, M. et al. (2011) 'Ideas underlying the Quantification of Margins and Uncertainties', *Reliability Engineering & System Safety*, 96(9), pp. 965–975.

Rachdi, N. (2011) *Statistical Learning and Computer Experiments*. PhD Thesis. University Paul Sabatier of Toulouse.

Raymer, D.P. (2012) *Aircraft design : a conceptual approach*.5th edn. Reston: American Institute of Aeronautics and Astronautics.

Renard, B. et al. (2011) 'Toward a reliable decomposition of predictive uncertainty in hydrological modeling: Characterizing rainfall errors using conditional simulation', *Water Resources Research*, 47(11)

Riaz, A. (2015) *A set-based approach to passenger aircraft family design*. PhD Thesis. Cranfield University.

Riaz, A. et al. (2017) 'Set-Based Approach to Passenger Aircraft Family Design', *Journal of Aircraft*, 54(1), pp. 310–326.

Rocquigny, E. de. et al. (2008) *Uncertainty in industrial practice : a guide to quantitative uncertainty management*. Chichester: John Wiley & Sons.

Rouvreau, S., Mangeant, F., and Arbez, P. (2015) 'TOICA Innovations in Aircraft Architecture Selection, Uncertainty Management in Collaborative Trade-off', *6th International Conference on Experiments/Process/System Modeling/Simulation/Optimization*. Athen, Greece, 08-11 July 2015.

Saltelli, A. (2002) 'Making best use of model evaluations to compute sensitivity indices', *Computer Physics Communications*, 145(2), pp. 280–297.

Saltelli, A. et al. (1993) 'Sensitivity analysis of model output: An investigation of new techniques', *Computational Statistics & Data Analysis*, 15(2), pp. 211–238.

Saltelli, A. et al. (2010) ‘Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index’, *Computer Physics Communications*, 181(2), pp. 259–270.

Saltelli, A. and Homma, T. (1992) ‘Sensitivity analysis for model output’, *Computational Statistics & Data Analysis*, 13(1), pp. 73–94.

Saltelli, A. et al. (2008) *Global Sensitivity Analysis: The Primer*. Chichester: John Wiley & Sons.

Saltelli, A. and Sobol’, I.M. (1995) ‘Sensitivity analysis for nonlinear mathematical models: numerical experience’, *Matematicheskoe Modelirovanie*, 7(11), pp. 16–28.

Saltelli, A. et al. (2004) *Sensitivity analysis in practice : a guide to assessing scientific models*. Chichester: John Wiley & Sons.

Saltelli, A. et al. (1999) ‘A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output’, *Technometrics*, 41(1), pp. 39–56.

Satterthwaite, F.E. (1959) ‘Random Balance Experimentation’, *Technometrics*, 1(2), pp. 111–137.

Schaibly, J.H. and Shuler, K.E. (1973) ‘Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. II Applications’, *The Journal of Chemical Physics*, 59(8), pp. 3879–3888.

Schuëller, G.I. and Jensen, H.A. (2008) ‘Computational methods in optimization considering uncertainties – An overview’, *Computer Methods in Applied Mechanics and Engineering*, 198(1), pp. 2–13.

Seo, H.S. and Kwak, B.M. (2002) ‘Efficient statistical tolerance analysis for general distributions using three-point information’, *International Journal of Production Research*, 40(4), pp. 931–944.

Shafer, G. (1976) *A mathematical theory of evidence*. New Jersey: Princeton University Press.

Shibata, K. et al. (2014) *SAE Technical Paper 2014-01-2199: Aircraft Secondary Power System Integration into Conceptual Design and Its Application to More Electric*

System'. Society of Automotive Engineers (SAE).

Simon, H.A. (1973) 'The structure of ill structured problems', *Artificial Intelligence*, 4(3–4), pp. 181–201.

Singer, D.J. et al. (2009) 'What Is Set-Based Design?', *Naval Engineers Journal*, 121(4), pp. 31–43.

Smith, R.C. (2013) *Uncertainty Quantification: Theory, Implementation, and Applications*. Society for Industrial and Applied Mathematics.

Sobek, D.K.D.K.I. et al. (1999) 'Toyota's Principles of Set-Based Concurrent Engineering', *MIT Sloan Management Review*, 40(2), pp. 67–84.

Sobol', I.M. (1990) 'On sensitivity estimation for nonlinear mathematical models (in Russian)', *Matematicheskoe Modelirovanie*, 2(1), pp. 112–118.

Sobol', I.M. (1993) 'Sensitivity analysis for nonlinear mathematical models', *Mathematical Modeling and Computational Experiment*, 1(4), pp. 407–414.

Sobol', I.M. (2003) 'Theorems and examples on high dimensional model representation', *Reliability Engineering & System Safety*, 79(2), pp. 187–193.

Sobol', I.M. and Kucherenko, S. (2009) 'Derivative based global sensitivity measures and their link with global sensitivity indices', *Mathematics and Computers in Simulation*, 79(10), pp. 3009–3017.

Sobol', I.M. and Levitan, Y.L. (1999) 'On the use of variance reducing multipliers in Monte Carlo computations of a global sensitivity index', *Computer Physics Communications*, 117(1–2), pp. 52–61.

Sobol', I.M. et al. (2007) 'Estimating the approximation error when fixing unessential factors in global sensitivity analysis', *Reliability Engineering & System Safety*, 92(7), pp. 957–960.

Sobol′, I.M. (2001) 'Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates', *Mathematics and Computers in Simulation*, 55(1–3), pp. 271–280.

Sudret, B. (2008) ‘Global sensitivity analysis using polynomial chaos expansions’, *Reliability Engineering & System Safety*, 93(7), pp. 964–979.

Taguchi, G. (1986) *Introduction to quality engineering: designing quality into products and processes*. Tokyo: Asian Productivity Organization.

Tarantola, A. (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics.

Tarantola, S. et al. (2006) ‘Random balance designs for the estimation of first order global sensitivity indices’, *Reliability Engineering & System Safety*, 91(6), pp. 717–727.

Thunnissen, D.P. (2003) ‘Uncertainty Classification for the Design and Development of Complex Systems’, *Proc. Third Annual Predictive Methods Conf. Veros Software*., pp. 1–16.

Thunnissen, D.P. (2004) ‘Method for Determining Margins in Conceptual Design’, *Journal of Spacecraft and Rockets*, 41(1), pp. 85–92.

Thunnissen, D.P. (2005) *Propagating and mitigating uncertainty in the design of complex multidisciplinary systems*. PhD Thesis. California Institute of Technology.

Thunnissen, D.P. and Tsuyuki, G.T. (2004) ‘Margin Determination in the Design and Development of a Thermal Control System’, *34th International Conference on Environmental Systems (ICES)*. Colorado Springs, CO, USA, 19-22 July 2004. Society of Automotive Engineers, SAE 04ICES-239.

van Heerden, A.S.J. et al. (2016) ‘Enhancing the exploration of aircraft changeability during conceptual design’, *30th Congress of the International Council of the Aeronautical Sciences (ICAS)*. Daejeon, Korea, 26-29 September 2016. International Council of The Aeronautical Sciences. Available at: http://www.icas.org/ICAS_ARCHIVE/ICAS2016/data/papers/2016_0525_paper.pdf (Accessed: 25 February 2018).

Wang, J. and Zabaras, N. (2004) ‘A Bayesian inference approach to the inverse heat conduction problem’, *International Journal of Heat and Mass Transfer*, 47(17–18), pp. 3927–3941.

Williams, G. (1978) *Computational linear algebra with models*. Boston: Allyn and Bacon.

Xia, X. and Lawson, C.P. (2013) 'The development of a design methodology for dynamic power distribution management on a civil transport all electric aircraft', *Aerospace Science and Technology*, 25(1), pp. 125–131.

Xiu, D. and Hesthaven, J.S. (2005) 'High-Order Collocation Methods for Differential Equations with Random Inputs', *SIAM Journal on Scientific Computing*, 27(3), pp. 1118–1139.

Xiu, D. and Karniadakis, G.E. (2002) 'The Wiener--Askey Polynomial Chaos for Stochastic Differential Equations', *SIAM Journal on Scientific Computing*, 24(2), pp. 619–644.

Xiu, D. et al. (2002) 'Stochastic Modeling of Flow-Structure Interactions Using Generalized Polynomial Chaos', *Journal of Fluids Engineering*, 124(1), pp. 51–59.

Xu, C. and Gertner, G. (2007) 'Extending a global sensitivity analysis technique to models with correlated parameters', *Computational Statistics & Data Analysis*, 51(12), pp. 5579–5590.

Xu, C. and Gertner, G.Z. (2008) 'Uncertainty and sensitivity analysis for models with correlated parameters', *Reliability Engineering & System Safety*, 93(10), pp. 1563–1573.

Yao, W. et al. (2011) 'Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles', *Progress in Aerospace Sciences*, 47(6), pp. 450–479.

Yuan, J. et al. (2016) 'Probabilistic margin assessment of aircraft conceptual design using a modified reliability based design optimization methodology', *Proceedings of ISMA 2016 International Conference on Noise and Vibration Engineering and USD 2016 International Conference on Uncertainty in Structural Dynamics*. pp. 4463–4477.

Zabaras, N. and Ganapathysubramanian, B. (2008) 'A scalable framework for the solution of stochastic inverse problems using a sparse grid collocation approach',

Journal of Computational Physics, 227(9), pp. 4697–4735.

Zadeh, L.A. (1999) 'Fuzzy sets as a basis for a theory of possibility', *Fuzzy Sets and Systems*, 100(1), pp. 9–34.

Zadeh, L.A. (1986) 'A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination', *AI Magazine*, 7(2), pp. 85–90.

Zang, T.A. et al. (2002) *NASA/TM-2002-211462: Needs and opportunities for uncertainty-based multidisciplinary design methods for aerospace vehicles*. National Aeronautics and Space Administration (NASA). Available at: https://ntrs.nasa.gov/search.jsp?R=20020063596 (Accessed: 25 February 2018).

Zang, T.A. et al. (2015) 'A Strategy for Probabilistic Margin Allocation in Aircraft Conceptual Design', *16th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, AIAA AVIATION Forum*. Dallas, TX, USA, 22-26 June 2015. Reston: American Institute of Aeronautics and Astronautics, AIAA 2015-3443.

Appendices

Appendix A Efficient Method for Variance Based Sensitivity Analysis

A.1 Formulation

The formulation of the first order, second order, total effect indices, adopting Option 2, 3, and 4 are given in this section. Some formulas which can be reused are listed in Table A-1.

Table A-1 Equations reused from the previous sections

Variables	Equation
$f_1^{(\sim i)}$	(6-20)
$\boldsymbol{x_r}^{+(\sim i)}, \boldsymbol{x_r}^{-(\sim i)}, \boldsymbol{\mu_x}^{(\sim i)}$	(6-22), (6-23), (6-24)
$W_0^{(i)}$	(6-27)
$W_0^{(\sim i)}$	(6-42)
$f_2^{(\sim i,j)}$	(6-29)
$\boldsymbol{x_r}^{+(\sim i,j)}, \boldsymbol{x_r}^{-(\sim i,j)}, \boldsymbol{\mu_x}^{(\sim i,j)}$	(6-31), (6-32), (6-33)
$W_0^{(i,j)}$	(6-36)
$f_3^{(i)}$	(6-38)

A.1.1 Option 2

The first order index:

$$S_i \approx \frac{W_0^{(i)}\left(g_4\left(\mu_{x_i}\right)\right)^2 + W_i\left[\frac{\left(g_4\left(\mu_{x_i}+h_i^+\sigma_{x_i}\right)\right)^2}{h_i^+} - \frac{\left(g_4\left(\mu_{x_i}+h_i^-\sigma_{x_i}\right)\right)^2}{h_i^-}\right] - E_{URQ}^2(y)}{V_{URQ}(y)}, \tag{A-1}$$

where:

$$g_4(X_i) = W_0^{(\sim i)} f_1^{(\sim i)}\left(\boldsymbol{\mu_x}^{(\sim i)}\right) + \sum_{\substack{r=1 \\ r\neq i}}^{n} W_r\left[\frac{f_1^{(\sim i)}(\boldsymbol{x_r}^{+(\sim i)})}{h_r^+} - \frac{f_1^{(\sim i)}(\boldsymbol{x_r}^{-(\sim i)})}{h_r^-}\right] \tag{A-2}$$

The second order index:

$$S_{ij} = \frac{W_0^i\left(g_5\left(\mu_{x_i},\mu_{x_j}\right)\right)^2 + W_i\left[\frac{\left(g_5\left(\mu_{x_i}+h_i^+\sigma_{x_i},\mu_{x_j}\right)\right)^2}{h_i^+} - \frac{\left(g_5\left(\mu_{x_i}+h_i^-\sigma_{x_i},\mu_{x_j}\right)\right)^2}{h_i^-}\right] + W_j\left[\frac{\left(g_5\left(\mu_{x_i},\mu_{x_j}+h_j^+\sigma_{x_j}\right)\right)^2}{h_j^+} - \frac{\left(g_5\left(\mu_{x_i},\mu_{x_j}+h_j^-\sigma_{x_j}\right)\right)^2}{h_j^-}\right] - E_{URQ}^2(y)}{V_{URQ}(y)} - S_i - S_j, \quad \textbf{(A-3)}$$

where:

$$g_5\left(X_i, X_j\right) = W_0^{(\sim i,j)} f_2^{(\sim i,j)}\left(\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i,j)}\right) + \sum_{\substack{r=1 \\ r\neq i,j}}^{n} W_r\left[\frac{f_2^{(\sim i,j)}\left(\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i,j)+}\right)}{h_r^+} - \frac{f_2^{(\sim i,j)}\left(\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i,j)-}\right)}{h_r^-}\right] \quad \textbf{(A-4)}$$

The total effect index:

$$S_i^T = \frac{V_{URQ}(y) + E_{URQ}^2(y) - W_0^{(\sim i)}\left(g_6\left(\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i)}\right)\right)^2 + \sum_{\substack{r=1 \\ r\neq i}}^{n} W_r\left[\frac{\left(g_6\left(\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i)+}\right)\right)^2}{h_r^+} - \frac{\left(g_6\left(\boldsymbol{x}_{\boldsymbol{r}}^{(\sim i)-}\right)\right)^2}{h_r^-}\right]}{V_{URQ}(y)}, \quad \textbf{(A-5)}$$

where:

$$g_6(X_1, X_2, \ldots X_{i-1}, X_{i+1}, \ldots X_n) = W_0^{(i)} f_3^{(i)}\left(\mu_{x_i}\right) + W_i\left[\frac{f_3^{(i)}(x_i^+)}{h_i^+} - \frac{f_3^{(i)}(x_i^-)}{h_i^-}\right] \quad \textbf{(A-6)}$$

A.1.2 Option 3

The first order index:

$$S_i \approx \frac{\left\{\begin{array}{c} W_i^+\left[\frac{g_4\left(\mu_{x_i}+h_i^+\sigma_{x_i}\right)-g_4\left(\mu_{x_i}\right)}{h_i^+}\right]^2 + W_i^-\left[\frac{g_4\left(\mu_{x_i}+h_i^-\sigma_{x_i}\right)-g_4\left(\mu_{x_i}\right)}{h_i^-}\right]^2 \\ + W_i^{\pm}\frac{\left[g_4\left(\mu_{x_i}+h_i^+\sigma_{x_i}\right)-g_4\left(\mu_{x_i}\right)\right]\left[g_4\left(\mu_{x_i}+h_i^-\sigma_{x_i}\right)-g_4\left(\mu_{x_i}\right)\right]}{h_i^+ h_i^-} \end{array}\right\}}{V_{URQ}(y)}, \quad \textbf{(A-7)}$$

where $g_4(\cdot)$ is defined by Equation (A-2).

The second order index:

$$S_{i,j} = \frac{\sum_{r=i,j}\left\{\begin{array}{c} W_r^+\left[\frac{g_5(\mu_{x_r}+h_r^+\sigma_{x_r})-g_5(\mu_{x_r})}{h_r^+}\right]^2 \\ +W_r^-\left[\frac{g_5(\mu_{x_r}+h_r^-\sigma_{x_r})-g_5(\mu_{x_r})}{h_r^-}\right]^2 \\ +W_r^{\pm}\frac{[g_5(\mu_{x_r}+h_r^+\sigma_{x_r})-g_5(\mu_{x_r})][g_5(\mu_{x_r}+h_r^-\sigma_{x_r})-g_5(\mu_{x_r})]}{h_r^+h_r^-}\end{array}\right\}}{V_{URQ}(y)} - S_i - S_j, \quad \textbf{(A-8)}$$

where $g_5(\cdot)$ is defined by Equation (A-4).

The total effect index:

$$S_i^T = \frac{V_{URQ}(y)-\sum_{\substack{r=1\\r\neq i}}^{n}\left\{\begin{array}{c} W_r^+\left[\frac{g_6\left(\boldsymbol{x_r}^{+}{}^{(\sim i)}\right)-g_6\left(\boldsymbol{\mu_x}^{(\sim i)}\right)}{h_r^+}\right]^2 \\ +W_r^-\left[\frac{g_6\left(\boldsymbol{x_r}^{-}{}^{(\sim i)}\right)-g_6\left(\boldsymbol{\mu_x}^{(\sim i)}\right)}{h_r^-}\right]^2 \\ +W_r^{\pm}\frac{\left[g_6(\boldsymbol{x_r}^{+}{}^{(\sim i)})-g_6\left(\boldsymbol{\mu_x}^{(\sim i)}\right)\right]\left[g_6(\boldsymbol{x_r}^{-}{}^{(\sim i)})-g_6\left(\boldsymbol{\mu_x}^{(\sim i)}\right)\right]}{h_r^+h_r^-}\end{array}\right\}}{V_{URQ}(y)}, \quad \textbf{(A-9)}$$

where $g_6(\cdot)$ is defined by Equation (A-6).

A.1.3 Option 4

The first order index:

$$S_i = \frac{W_0^{+i}f^2(\boldsymbol{\mu_x})+\sum_{\substack{r=1\\r\neq i}}^{n}2W_r\left[\frac{f(x_r^+)f(\boldsymbol{\mu_x})}{h_r^+}-\frac{f(x_r^-)f(\boldsymbol{\mu_x})}{h_r^-}\right]+W_i\left[\frac{f(x_i^+)f\left(x_i^+\right)}{h_i^+}-\frac{f(x_i^-)f(x_i^-)}{h_i^-}\right]-E_{URQ}^2(y)}{V_{URQ}(y)}, \quad \textbf{(A-10)}$$

where:

$$W_0^{+i} = W_0 + \sum_{\substack{r=1\\r\neq i}}^{n}\frac{1}{h_r^+h_r^-} = 2W_0 - \frac{1}{h_i^+h_i^-} - 1 \quad \textbf{(A-11)}$$

The second order index:

$$S_{ij} = \frac{\begin{matrix} W_0^{++ij} f^2(\boldsymbol{\mu_x}) + \sum_{\substack{r=1 \\ r\neq i,j}}^{n} 2W_r \left[\frac{f(x_r^+)f(\mu_x)}{h_r^+} - \frac{f(x_r^-)f(\mu_x)}{h_r^-}\right] \\ + \sum_{r=i,j} \left[\frac{f(x_r^+)f(x_r^+)}{h_r^+} - \frac{f(x_r^-)f(x_r^+)}{h_r^-}\right] - E_{URQ}^2(y) \end{matrix}}{V_{URQ}(y)} - S_i - S_j, \quad \textbf{(A-12)}$$

where:

$$W_0^{++ij} = W_0 + \sum_{\substack{r=1 \\ r\neq i,j}}^{n} \frac{1}{h_r^+ h_r^-} = 2W_0 - \frac{1}{h_i^+ h_i^-} - \frac{1}{h_j^+ h_j^-} - 1 \quad \textbf{(A-13)}$$

The total effect index:

$$S_i^T = \frac{\begin{matrix} V_{URQ}(y) + E_{URQ}^2(y) - W_0^{-+i} f^2(\boldsymbol{\mu_x}) \\ - \sum_{\substack{r=1 \\ r\neq i}}^{n} W_r \left[\frac{f(x_r^+)f(x_r^+)}{h_r^+} - \frac{f(x_r^-)f(x_r^-)}{h_r^-}\right] - 2W_i \left[\frac{f(x_i^+)f(\mu_x)}{h_i^+} - \frac{f(x_i^-)f(\mu_x)}{h_i^-}\right] \end{matrix}}{V_{URQ}(y)}, \quad \textbf{(A-14)}$$

where:

$$W_0^{-+i} = W_0 + \frac{1}{h_i^+ h_i^-} \quad \textbf{(A-15)}$$

A.2 Pseudo Code

Input: First four moments of all the input variables, μ_{x_i}, σ_{x_i}, γ_{x_i}, Γ_{x_i}, $i = 1,2,3,\dots,n$;
Output: First order indices, $S_i, i = 1,2,3,\dots,n$;

do Calculate all the coefficients;
do Evaluate the model at all the required points

for $i = 1{:}\,n$ **do**

for $r = 1{:}\,n; r \neq i$ **do**
$f_1^{(\sim i)}(\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i)}) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i}, \dots, \mu_{x_r}, \dots, \mu_{x_n})$;
$f_1^{(\sim i)}\left(\boldsymbol{x}_{\boldsymbol{r}}^{+(\sim i)}\right) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i}, \dots, \mu_{x_r} + h_r^+ \sigma_{x_r}, \dots, \mu_{x_n})$;
$f_1^{(\sim i)}(\boldsymbol{x}_{\boldsymbol{r}}^{-(\sim i)}) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i}, \dots, \mu_{x_r} + h_r^- \sigma_{x_r}, \dots, \mu_{x_n})$;
Execute Eq. (6-21);
end

for $r = 1{:}\,n; r \neq i$ **do**
$f_1^{(\sim i)}(\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i)}) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i} + h_i^+ \sigma_{x_i}, \dots, \mu_{x_r}, \dots, \mu_{x_n})$;
$f_1^{(\sim i)}\left(\boldsymbol{x}_{\boldsymbol{r}}^{+(\sim i)}\right) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i} + h_i^+ \sigma_{x_i}, \dots, \mu_{x_r} + h_r^+ \sigma_{x_r}, \dots, \mu_{x_n})$;
$f_1^{(\sim i)}(\boldsymbol{x}_{\boldsymbol{r}}^{-(\sim i)}) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i} + h_i^+ \sigma_{x_i}, \dots, \mu_{x_r} + h_r^- \sigma_{x_r}, \dots, \mu_{x_n})$;
Execute Eq. (6-21);
end

for $r = 1{:}\,n; r \neq i$ **do**
$f_1^{(\sim i)}(\boldsymbol{\mu}_{\boldsymbol{x}}^{(\sim i)}) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i} + h_i^- \sigma_{x_i}, \dots, \mu_{x_r}, \dots, \mu_{x_n})$;
$f_1^{(\sim i)}\left(\boldsymbol{x}_{\boldsymbol{r}}^{+(\sim i)}\right) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i} + h_i^- \sigma_{x_i}, \dots, \mu_{x_r} + h_r^+ \sigma_{x_r}, \dots, \mu_{x_n})$;
$f_1^{(\sim i)}(\boldsymbol{x}_{\boldsymbol{r}}^{-(\sim i)}) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_i} + h_i^- \sigma_{x_i}, \dots, \mu_{x_r} + h_r^- \sigma_{x_r}, \dots, \mu_{x_n})$;
Execute Eq. (6-21);
end

Execute Eq. (6-26);
Execute Eq. (6-28);

end

A.3 Matlab Code

```
function [S_1, S_2, S_T]=urqsobol(f,in)


%f is the model
%in is the input information in the following format:
%   {   mean, std, skew, kurtosis;
%       mean, std, skew, kurtosis;
%       mean, std, skew, kurtosis;
%       ...
%       mean, std, skew, kurtosis;  }

n=size(in,1); %number of var
S_1=zeros(1,n);  %first order
S_2=zeros(n,n); %Second order
S_T=zeros(1,n);  %Total

%URQ coefficient
h_plus=zeros(1,n);
h_minus=zeros(1,n);
w=zeros(1,n);
w_plus=zeros(1,n);
w_minus=zeros(1,n);
w_pm=zeros(1,n);
x_mu=in(:,1);

E=0;
V=0;
w0=1;

f0=f(x_mu');  %first run    +1
f_p=zeros(1,n);     % to store sampling points: one plus all others remain at mean
f_m=zeros(1,n);     % to store sampling points: one minus all others remain at mean
f_pp=zeros(n,n);
f_mm=zeros(n,n);
f_pm=zeros(n,n);


%normal URQ
for i=1:n
    h_plus(i)=(in(i,3)/2)+sqrt(in(i,4)-3*(in(i,3)^2)/4);
    h_minus(i)=(in(i,3)/2)-sqrt(in(i,4)-3*(in(i,3)^2)/4);

    w0=w0+1/(h_plus(i)*h_minus(i));

    w(i)=1/(h_plus(i)-h_minus(i));
    w_plus(i)=(h_plus(i)^2-h_plus(i)*h_minus(i)-1)/((h_plus(i)-h_minus(i))^2);
    w_minus(i)=(h_minus(i)^2-h_plus(i)*h_minus(i)-1)/((h_plus(i)-h_minus(i))^2);
    w_pm(i)=2/((h_plus(i)-h_minus(i))^2);

    x_plus=in(:,1);
    x_plus(i)=x_plus(i)+h_plus(i)*in(i,2);

    x_minus=in(:,1);
    x_minus(i)=x_minus(i)+h_minus(i)*in(i,2);

    %run for single plus        +n
    f_p(i)=f(x_plus');
    %run for single minus       +n
    f_m(i)=f(x_minus');

    E=E+w(i)*(f_p(i)/h_plus(i)-f_m(i)/h_minus(i));
    V=V+(w_plus(i)*((f_p(i)-f0)/h_plus(i))^2+w_minus(i)*((f_m(i)-
f0)/h_minus(i))^2+w_pm(i)*((f_p(i)-f0)*(f_m(i)-f0)/(h_plus(i)*h_minus(i))));
end

E=E+f0*w0;   %w0 is not ready until the for loop finished
%normal URQ
```

```
for i=1:n %a seperate loop is used as all the h_plus and h_minus are needed, and they
are not ready until the previous loop finished
    for j=1:n
        if j~=i
            x_plus_minus=in(:,1);
            x_plus_minus(i)=x_plus_minus(i)+ h_plus(i)*in(i,2);
            x_plus_minus(j)=x_plus_minus(j)+ h_minus(j)*in(j,2);
            f_pm(i,j)= f(x_plus_minus');                                %run for plus and
minus n(n-1)
        end

        if j>i
            x_plus_plus=in(:,1);
            x_plus_plus(i)=x_plus_plus(i)+ h_plus(i)*in(i,2);
            x_plus_plus(j)=x_plus_plus(j)+ h_plus(j)*in(j,2);
            f_pp(i,j)= f(x_plus_plus');                                 %run for two plus
n(n-1)/2
            f_pp(j,i)=f_pp(i,j);                                        %for convenience


            x_minus_minus=in(:,1);
            x_minus_minus(i)=x_minus_minus(i)+ h_minus(i)*in(i,2);
            x_minus_minus(j)=x_minus_minus(j)+ h_minus(j)*in(j,2);
            f_mm(i,j)= f(x_minus_minus');                               %run for two minus
n(n-1)/2
            f_mm(j,i)=f_mm(i,j);                                        %for convenience
        end
    end
end

option = 4; %or 2,3,4

switch option
    case 1
        %First order
        for i=1:n
            G1_1=V-(w_plus(i)*((f_p(i)-f0)/h_plus(i))^2+w_minus(i)*((f_m(i)-
f0)/h_minus(i))^2+w_pm(i)*((f_p(i)-f0)*(f_m(i)-f0)/(h_plus(i)*h_minus(i))));
            % V~xi

            G1_2=0;
            G1_3=0;

            for j=1:n
                if j~=i
                    G1_2=G1_2 ...
                        +w_plus(j)  *   ( ( f_pp(i,j)-f_p(i) ) / h_plus(j) )^2 ...
                        +w_minus(j) *   ( ( f_pm(i,j)-f_p(i) ) / h_minus(j) )^2 ...
%ith always plus
                        +w_pm(j)    *   ( ( f_pp(i,j)-f_p(i) )*( f_pm(i,j)-f_p(i) ) / (
h_plus(j)*h_minus(j) ) );

                    G1_3=G1_3 ...
                        +w_plus(j)  *   ( ( f_pm(j,i)-f_m(i) ) / h_plus(j) )^2 ...
                        +w_minus(j) *   ( ( f_mm(i,j)-f_m(i) ) / h_minus(j) )^2 ...
%ith always minus
                        +w_pm(j)    *   ( ( f_pm(j,i)-f_m(i) )*( f_mm(i,j)-f_m(i) ) / (
h_plus(j)*h_minus(j) ) );
                end
            end
            EV=(1+1/(h_plus(i)*h_minus(i)))*G1_1+w(i)*(G1_2/h_plus(i)-G1_3/h_minus(i));
            S_1(i)=(V-EV)/V;
        end
        %Second Order
        for i=1:n-1
            for j=i+1:n
                %(1)
                H1_1=V-(w_plus(i)*((f_p(i)-f0)/h_plus(i))^2+w_minus(i)*((f_m(i)-
f0)/h_minus(i))^2+w_pm(i)*((f_p(i)-f0)*(f_m(i)-f0)/(h_plus(i)*h_minus(i))))...
                    -(w_plus(j)*((f_p(j)-f0)/h_plus(j))^2+w_minus(j)*((f_m(j)-
f0)/h_minus(j))^2+w_pm(j)*((f_p(j)-f0)*(f_m(j)-f0)/(h_plus(j)*h_minus(j))));
```

```
                H1_2_i_plus=0;
                H1_2_i_minus=0;
                H1_3_j_plus=0;
                H1_3_j_minus=0;
                for k=1:n
                    if ( (k~=i) && (k~=j) )
                    H1_2_i_plus=H1_2_i_plus...
                          +w_plus(k)  *   ( ( f_pp(i,k)-f_p(i) ) / h_plus(k) )^2 ...
                          +w_minus(k) *   ( ( f_pm(i,k)-f_p(i) ) / h_minus(k) )^2 ...
%ith always plus
                          +w_pm(k)    *   ( ( f_pp(i,k)-f_p(i) )*( f_pm(i,k)-f_p(i) ) /
( h_plus(k)*h_minus(k) ) );

                    H1_2_i_minus=H1_2_i_minus...
                          +w_plus(k)  *   ( ( f_pm(k,i)-f_m(i) ) / h_plus(k) )^2 ...
                          +w_minus(k) *   ( ( f_mm(i,k)-f_m(i) ) / h_minus(k) )^2 ...
%ith always minus
                          +w_pm(k)    *   ( ( f_pm(k,i)-f_m(i) )*( f_mm(i,k)-f_m(i) ) /
( h_plus(k)*h_minus(k) ) );

                    H1_3_j_plus=H1_3_j_plus...
                          +w_plus(k)  *   ( ( f_pp(j,k)-f_p(j) ) / h_plus(k) )^2 ...
                          +w_minus(k) *   ( ( f_pm(j,k)-f_p(j) ) / h_minus(k) )^2 ...
%jth always plus
                          +w_pm(k)    *   ( ( f_pp(j,k)-f_p(j) )*( f_pm(j,k)-f_p(j) ) /
( h_plus(k)*h_minus(k) ) );

                    H1_3_j_minus=H1_3_j_minus...
                          +w_plus(k)  *   ( ( f_pm(k,j)-f_m(j) ) / h_plus(k) )^2 ...
                          +w_minus(k) *   ( ( f_mm(j,k)-f_m(j) ) / h_minus(k) )^2 ...
%jth always minus
                          +w_pm(k)    *   ( ( f_pm(k,j)-f_m(j) )*( f_mm(j,k)-f_m(j) ) /
( h_plus(k)*h_minus(k) ) );
                    end
                end
                w0_ij=1+1/(h_plus(i)*h_minus(i))+1/(h_plus(j)*h_minus(j));
                V_alao=w0_ij*H1_1...   %V as long as others
                    +w(i)*(H1_2_i_plus/h_plus(i)-H1_2_i_minus/h_minus(i))...
                    +w(j)*(H1_3_j_plus/h_plus(j)-H1_3_j_minus/h_minus(j));
                S_2(i,j)=(V-V_alao)/V-S_1(i)-S_1(j);
            end
        end
        %Total effect
        for i=1:n
            %all X~i = mu
            G_mu=w_plus(i) *((f_p(i)-f0)/h_plus(i) )^2 ...
                +w_minus(i)*((f_m(i)-f0)/h_minus(i))^2 ...
                +w_pm(i)*(f_p(i)-f0)*(f_m(i)-f0)/(h_plus(i)*h_minus(i));
            EV=(w0-1/(h_plus(i)*h_minus(i)))*G_mu;

            for j=1:n  %outter loop
                if j~=i
                    %Xj +
                    G_j_plus=w_plus(i) *((f_pp(i,j)-f_p(j))/h_plus(i) )^2 ...
%h_plus(i) is used, as in G_j_plus, xi is the variable
                           +w_minus(i)*((f_pm(j,i)-f_p(j))/h_minus(i))^2 ...   %note
that f_pm(j,i) is not a mistake, it means f(xi-,xj+)
                           +w_pm(i)*(f_pp(i,j)-f_p(j))*(f_pm(j,i)-
f_p(j))/(h_plus(i)*h_minus(i));

                    %Xj -
                    G_j_minus=w_plus(i) *((f_pm(i,j)-f_m(j))/h_plus(i) )^2 ...
%h_plus(i) is used, as in G_j_minus, xi is the variable
                            +w_minus(i)*((f_mm(i,j)-f_m(j))/h_minus(i))^2 ...
                            +w_pm(i)*(f_pm(i,j)-f_m(j))*(f_mm(i,j)-
f_m(j))/(h_plus(i)*h_minus(i));

                    EV=EV+w(j)*(G_j_plus/h_plus(j)-G_j_minus/h_minus(j));
                end
            end
        S_T(i)=EV/V;
```

```
        end
    case 2
        %First order
        for i=1:n

            G2_1=(E-f0/(h_plus(i)*h_minus(i))-w(i)*(f_p(i)/h_plus(i)-
f_m(i)/h_minus(i)))^2;
            % (E~xi)^2

            G2_2=(w0-1/(h_plus(i)*h_minus(i)))*f_p(i);
            G2_3=(w0-1/(h_plus(i)*h_minus(i)))*f_m(i);
            for j=1:n
                if j~=i
                    G2_2=G2_2 ...
                        +w(j)  *   ( f_pp(i,j) / h_plus(j) - f_pm(i,j) / h_minus(j) );

                    G2_3=G2_3 ...
                        +w(j)  *   ( f_pm(j,i) / h_plus(j) - f_mm(i,j) / h_minus(j) );
                end
            end
            G2_2=G2_2^2;
            G2_3=G2_3^2;

            EE=(1+1/(h_plus(i)*h_minus(i)))*G2_1+w(i)*(G2_2/h_plus(i)-G2_3/h_minus(i));
            S_1(i)=(EE-E^2)/V;
        end
        %Second Order
        for i=1:n-1
            for j=i+1:n
                H2_1=E-( 1/( h_plus(i)*h_minus(i) ) + 1/( h_plus(j)*h_minus(j) ))*f0...
                    -w(i)*(f_p(i)/h_plus(i)-f_m(i)/h_minus(i))...
                    -w(j)*(f_p(j)/h_plus(j)-f_m(j)/h_minus(j));
                H2_1=H2_1^2;

                H2_2_i_plus = ( w0 -( 1/( h_plus(i)*h_minus(i) ) + 1/(
h_plus(j)*h_minus(j) ) ) )*f_p(i);
                H2_2_i_minus = ( w0 -( 1/( h_plus(i)*h_minus(i) ) + 1/(
h_plus(j)*h_minus(j) ) ) )*f_m(i);
                H2_3_j_plus = ( w0 -( 1/( h_plus(i)*h_minus(i) ) + 1/(
h_plus(j)*h_minus(j) ) ) )*f_p(j);
                H2_3_j_minus = ( w0 -( 1/( h_plus(i)*h_minus(i) ) + 1/(
h_plus(j)*h_minus(j) ) ) )*f_m(j);

                for k=1:n
                    if ( (k~=i) && (k~=j) )
                        H2_2_i_plus=H2_2_i_plus+w(k)*(f_pp(i,k)/h_plus(k)-
f_pm(i,k)/h_minus(k));
                        H2_2_i_minus=H2_2_i_minus+w(k)*(f_pm(k,i)/h_plus(k)-
f_mm(i,k)/h_minus(k));
                        H2_3_j_plus=H2_3_j_plus+w(k)*(f_pp(j,k)/h_plus(k)-
f_pm(j,k)/h_minus(k));
                        H2_3_j_minus=H2_3_j_minus+w(k)*(f_pm(k,j)/h_plus(k)-
f_mm(j,k)/h_minus(k));
                    end
                end

                H2_2_i_plus=H2_2_i_plus^2;
                H2_2_i_minus=H2_2_i_minus^2;
                H2_3_j_plus=H2_3_j_plus^2;
                H2_3_j_minus=H2_3_j_minus^2;

                EEo2=(1+( 1/( h_plus(i)*h_minus(i) ) + 1/( h_plus(j)*h_minus(j) )
))*H2_1...
                    +w(i)*(H2_2_i_plus/h_plus(i)-H2_2_i_minus/h_minus(i))...
                    +w(j)*(H2_3_j_plus/h_plus(j)-H2_3_j_minus/h_minus(j));

                S_2(i,j)=(EEo2-E^2)/V-S_1(i)-S_1(j);
            end
        end
        %Total effect
        for i=1:n
```

```
            %all X~i = mu
            G_mu=(1+1/(h_plus(i)*h_minus(i)))*f0 ...
                +w(i)*(f_p(i)/h_plus(i)-f_m(i)/h_minus(i));
            G_mu=G_mu^2;
            EE=(w0-1/(h_plus(i)*h_minus(i)))*G_mu;
            for j=1:n  %outter loop
                if j~=i
                    %Xj +
                    G_j_plus=(1+1/(h_plus(i)*h_minus(i)))*f_p(j) ...  %xi = mu; xj = +
                            +w(i)*(f_pp(j,i)/h_plus(i)-f_pm(j,i)/h_minus(i));
                    G_j_plus=G_j_plus^2;

                    %Xj -
                    G_j_minus=(1+1/(h_plus(i)*h_minus(i)))*f_m(j) ... %xi = mu; xj = -
                            +w(i)*(f_pm(i,j)/h_plus(i)-f_mm(i,j)/h_minus(i));
                    G_j_minus=G_j_minus^2;

                    EE=EE+w(j)*(G_j_plus/h_plus(j)-G_j_minus/h_minus(j));
                end
            end

            S_T(i)=(V+E^2-EE)/V;
        end
    case 3
        %First order
        for i=1:n

            G2_1=E-f0/(h_plus(i)*h_minus(i))-w(i)*(f_p(i)/h_plus(i)-f_m(i)/h_minus(i));
            % (E~xi)

            G2_2=(w0-1/(h_plus(i)*h_minus(i)))*f_p(i);
            G2_3=(w0-1/(h_plus(i)*h_minus(i)))*f_m(i);
            for j=1:n   %j is the counter for inner loop. that is all the variables
other than the i_th
                if j~=i
                    G2_2=G2_2 ...
                        +w(j)  *   ( f_pp(i,j) / h_plus(j) - f_pm(i,j) / h_minus(j) );

                    G2_3=G2_3 ...
                        +w(j)  *   ( f_pm(j,i) / h_plus(j) - f_mm(i,j) / h_minus(j) );
                end
            end

            VE= w_plus(i)*((G2_2-G2_1)/h_plus(i))^2+w_minus(i)*((G2_3-
G2_1)/h_minus(i))^2+w_pm(i)*(G2_2-G2_1)*(G2_3-G2_1)/(h_plus(i)*h_minus(i));
            S_1(i)=VE/V;
        end
        %Second Order
        for i=1:n-1
            for j=i+1:n
                H2_1=E-( 1/( h_plus(i)*h_minus(i) ) + 1/( h_plus(j)*h_minus(j) ))*f0...
                    -w(i)*(f_p(i)/h_plus(i)-f_m(i)/h_minus(i))...
                    -w(j)*(f_p(j)/h_plus(j)-f_m(j)/h_minus(j));

                H2_2_i_plus = ( w0 -( 1/( h_plus(i)*h_minus(i) ) + 1/(
h_plus(j)*h_minus(j) ) ) )*f_p(i);
                H2_2_i_minus = ( w0 -( 1/( h_plus(i)*h_minus(i) ) + 1/(
h_plus(j)*h_minus(j) ) ) )*f_m(i);
                H2_3_j_plus = ( w0 -( 1/( h_plus(i)*h_minus(i) ) + 1/(
h_plus(j)*h_minus(j) ) ) )*f_p(j);
                H2_3_j_minus = ( w0 -( 1/( h_plus(i)*h_minus(i) ) + 1/(
h_plus(j)*h_minus(j) ) ) )*f_m(j);

                for k=1:n
                    if ( (k~=i) && (k~=j) )
                        H2_2_i_plus=H2_2_i_plus+w(k)*(f_pp(i,k)/h_plus(k)-
f_pm(i,k)/h_minus(k));
                        H2_2_i_minus=H2_2_i_minus+w(k)*(f_pm(k,i)/h_plus(k)-
f_mm(i,k)/h_minus(k));
                        H2_3_j_plus=H2_3_j_plus+w(k)*(f_pp(j,k)/h_plus(k)-
f_pm(j,k)/h_minus(k));
```

```
                        H2_3_j_minus=H2_3_j_minus+w(k)*(f_pm(k,j)/h_plus(k)-
f_mm(j,k)/h_minus(k));
                    end
                end

                VE  =w_plus(i) *((H2_2_i_plus -H2_1)/h_plus(i))^2 ...
                    +w_minus(i)*((H2_2_i_minus-H2_1)/h_minus(i))^2 ...
                    +w_pm(i) * (H2_2_i_plus -H2_1)*(H2_2_i_minus-
H2_1)/(h_plus(i)*h_minus(i))...
                    +w_plus(j) *((H2_3_j_plus -H2_1)/h_plus(j))^2 ...
                    +w_minus(i)*((H2_3_j_minus-H2_1)/h_minus(j))^2 ...
                    +w_pm(i) * (H2_3_j_plus -H2_1)*(H2_3_j_minus-
H2_1)/(h_plus(j)*h_minus(j));


                S_2(i,j)=VE/V-S_1(i)-S_1(j);
            end
        end
        %Total effect
        for i=1:n
            %all X~i = mu
            G_j_mu=(1+1/(h_plus(i)*h_minus(i)))*f0 ...
                +w(i)*(f_p(i)/h_plus(i)-f_m(i)/h_minus(i));

            VE=0;
            for j=1:n  %outter loop
                if j~=i
                    %Xj +
                    G_j_plus=(1+1/(h_plus(i)*h_minus(i)))*f_p(j) ...  %xi = mu; xj = +
                            +w(i)*(f_pp(j,i)/h_plus(i)-f_pm(j,i)/h_minus(i));


                    %Xj -
                    G_j_minus=(1+1/(h_plus(i)*h_minus(i)))*f_m(j) ... %xi = mu; xj = -
                            +w(i)*(f_pm(i,j)/h_plus(i)-f_mm(i,j)/h_minus(i));


                    VE=VE  + w_plus(j) *((G_j_plus -G_j_mu)/h_plus(j) )^2 ...
                           + w_minus(j)*((G_j_minus-G_j_mu)/h_minus(j))^2 ...
                           +w_pm(j)*(G_j_plus -G_j_mu)*(G_j_minus-
G_j_mu)/(h_plus(j)*h_minus(j));
                end
            end

            S_T(i)=(V-VE)/V;
        end
    otherwise
        %First order
        for i=1:n

            w0_ff=2*w0-1/(h_plus(i)*h_minus(i))-1;
            E_ff=w0_ff*f0^2;
            for j=1:n
                if j~=i
                    E_ff=E_ff+2*w(j)*(f_p(j)*f0/h_plus(j)-f_m(j)*f0/h_minus(j));
                else
                    E_ff=E_ff+w(i)*(f_p(i)*f_p(i)/h_plus(i)-f_m(i)*f_m(i)/h_minus(i));
                end
            end

            S_1(i)=(E_ff-E^2)/V;
        end
        %Second Order
        for i=1:n-1
            for j=i+1:n
                w0_ff_2=2*w0-1/(h_plus(i)*h_minus(i))-1/(h_plus(j)*h_minus(j))-1;
                E_ff=w0_ff_2*f0^2;
                for k=1:n
                    if ( (k~=i) && (k~=j) )
                        E_ff=E_ff+2*w(k)*(f_p(k)*f0/h_plus(k)-f_m(k)*f0/h_minus(k));
                    else
```

```
                        E_ff=E_ff+w(k)*(f_p(k)*f_p(k)/h_plus(k)-
f_m(k)*f_m(k)/h_minus(k));
                    end
                end
        %         xxx =(1-2*w0)*f0^2-(  2*w(1)*(f_p(1)*f0/h_plus(1)-
f_m(1)*f0/h_minus(1))...
        %                              +2*w(2)*(f_p(2)*f0/h_plus(2)-
f_m(2)*f0/h_minus(2))...
        %                              +2*w(3)*(f_p(3)*f0/h_plus(3)-
f_m(3)*f0/h_minus(3)))
        % xxx is the fix value


                S_2(i,j)=(E_ff-E^2)/V-S_1(i)-S_1(j);
            end
        end
        %Total effect
        for i=1:n
            EF=(w0+1/(h_plus(i)*h_minus(i)))*f0^2;
            for j=1:n
                if j~=i
                    EF=EF+w(j)*(f_p(j)^2/h_plus(j)-f_m(j)^2/h_minus(j));
                end
            end
            EF=EF+2*w(i)*(f_p(i)*f0/h_plus(i)-f_m(i)*f0/h_minus(i));
            S_T(i)=(V+E^2-EF)/V;
        end
end
```

Appendix B Publications

Chen, X., Molina-Cristóbal, Guenov, M. D. (2015) 'A Novel Method for Inverse Uncertainty Propagation', *Proceeding of International Conference on Evolutionary and Deterministic Methods for Design, Optimization and Control with Applications to Industrial and Societal Problems, EUROGEN*. Glasgow, UK 04-16 September 2015. (accepted for publication in book: Computational Methods in Applied Sciences).

Guenov, M. D., Chen, X., Molina-Cristóbal, A., Riaz, A., van Heerden, A. S. J., and Padulo, M. (2017) 'Margin Allocation and Trade-off in Complex Systems Design and Optimization', (paper accepted for publication in AIAA Journal).

Molina-Cristóbal, A., Chen, X., Guenov, M. D., Riaz, A., and van Heerden, A. S. J. (2018) 'Interactive Uncertainty Allocation and Trade-off at Early-stage Aircraft Computational Design', *2018 AIAA Non-Deterministic Approaches Conference, AIAA SciTech Forum.* Kissimmee, FL, USA, 08-12 January 2018. Reston: American Institute of Aeronautics and Astronautics, AIAA 2018-2170.

Chen, X., Molina-Cristóbal, Guenov, M. D. (2017) 'An Efficient Method for Variance Based Sensitivity Analysis', (paper submitted to Reliability Engineering and System Safety).