On Credibility Assessment in Aircraft System Simulation

Magnus Eek

Division of Machine Design
Department of Management and Engineering
Linköping University, SE-581 83 Linköping, Sweden

Linköping 2016
To Martina, Selma, Maja & Ellen
Kliv, kliv, överlev

– Mathias Fredriksson

Pust, pust och pust

– Martin Johnsrud Sundby
Abstract

The aeronautical industry is becoming increasingly reliant on Modeling and Simulation (M&S) for use throughout all system development phases, for system verification, and end-user training. To justify and to utilize the full potential of today’s model-based approach and to enable the industrially requested transition towards virtual testing and virtual certification, the development of efficient and industrially applicable methods for M&S credibility assessment is a key challenge. Credibility assessment of an M&S effort using one individual simulation model is undoubtedly a challenging task involving several vital aspects like verification, validation, uncertainty quantification, and model use history. Building confidence in simulator applications in which several individual models are connected and simulated together is even more problematic.

This work addresses methods facilitating credibility assessment of simulation models and simulator applications used in aircraft system development. For models of individual aircraft subsystems, like liquid cooling or environmental control systems, an uncertainty aggregation method is proposed that facilitates early model validation through approximate uncertainty quantification. The central idea is to integrate information obtained during component level validation directly into the component equations, and to utilize this information in model level uncertainty quantification. The method is applicable in early development phases when the availability of system level measurement data for traditional model validation purposes is typically very limited. To clarify, the model may describe a system not yet realized, neither in a test rig nor in a test aircraft.

In addition to methods intended for models of individual subsystems, this work also proposes a method and an associated tool for credibility assessment of large-scale simulator applications. As a complement to traditional document-centric approaches, static and dynamic credibility information is here presented to end-users directly during simulation. This implies a support for detecting test plan deficiencies, or that a simulator configuration is not a suitable platform for the execution of a particular test. The credibility assessment tool has been implemented and evaluated in two large-scale system simulators for the Saab Gripen fighter aircraft. The work presented herein also includes an industrially applicable workflow for development, validation, and export of simulation models.

Thanks to the close connection to industry, some of the research results have already been successfully implemented in operations and are currently in industrial use.
LYGINDOSTRIN förlitar sig i allt högre utsträckning på modellering och simulering inom såväl systemutveckling som verifiering, samt för träning av piloter och flygtekniker. För att kunna styrka och dra maximal nytta av dagens modellbaserade arbetssätt krävs utveckling av effektiv och industriellt relevant metodik för bedömning av trovärdigheten i simuleringsresultaten. Detta är också en nyckelfråga i den pågående vidareutvecklingen av modellering och simulering för användning inom virtuell testning och virtuell certifiering. Trovärdighetsbedömning är ett vitt begrepp som inkluderar ett flertal viktiga aspekter så som verifiering, validering, osäkerhetskvantifiering och användningshistorik. Bedömning av en enskild modells trovärdighet kan vara nog så utmanande, och i simulatorfallet då ett flertal modeller kopplas ihop och simuleras tillsammans är det än mer problematiskt.

Detta arbete behandlar metodik för trovärdighetsbedömning av simuleringsmodeller och simulatorapplikationer som används vid systemutveckling inom flygindustrin. För att möjliggöra tidig validering av modeller för enskilda delsystem i flygplan, så som vätske- eller luftkylsystem, föreslås en metod för approximativ osäkerhetskvantifiering. Även i tidiga utvecklingsfas kan oftast någon form av modellvalidering på komponentnivå genomföras. Den centrala idén är här att integrera information från sådan komponentvalidering direkt i en modells komponenter, och att utnyttja denna information vid förenklad osäkerhetskvantifiering på modellnivå. Metodiken är tillämpbar i tidiga utvecklingsfas där tillgängligheten på mätdata på systemnivå normalt sett är mycket begränsad, d.v.s. modellen kan beskriva ett system som ännu inte är realiserat vare sig i form av en testrigg eller i ett provflygplan.

Uöver metodik avsedd för modeller av enskilda delsystem föreslås även metodik och tillhörande verktyg för trovärdighetsbedömning av storskaliga simulatorapplikationer. Som ett komplement till traditionella dokumentationscenterade arbetssätt presenterar verktyget både statisk och dynamisk trovärdighetsinformation till simulatoranvändaren direkt under pågående simulering. Detta möjliggör upptäckt av brister i provprogram eller insikt om att aktuell simulator inte är en lämplig plattform för den tänkta provningen. Verktyget är implementerat och utvärderat i två storskaliga systemsimulatorer för Saab.
39 Gripen. Resultatet från projektet inkluderar även en industriellt relevant process för utveckling, validering och export av simuleringsmodeller.

Den goda kopplingsen till industrin har möjliggjort att delar av forskningsresultaten redan har implementerats och kommit till industriell nytta.
Acknowledgements

The work presented in this thesis was carried out in the form of an industrial PhD project at the Division of Machine Design at Linköping University. The research was funded by Saab Aeronautics and the National Aviation Engineering Research Programme (NFFP) jointly driven by the Swedish Armed Forces, the Swedish Defence Materiel Administration (FMV), and the Swedish Governmental Agency for Innovation Systems (VINNOVA).

First of all, I’d like to thank my supervisor Prof. Johan Ölvander for his efforts in reviewing, discussing, and directing the research, and for excellent guidance through the academic world. I also want to thank my industrial co-supervisor Dr. Hampus Gavel for always providing rational and concise advice. Special thanks go to Lic.Eng. Sören Steinkellner who helped me get started with the research project, and to my line manager Peter Gotenstam for protecting the research from drowning in industrial assignments. I’ve been very fortunate to receive good response and commitment internally at Saab Aeronautics. Thus, there are a number of people from several departments who have contributed to this work, especially from Vehicle Systems, Flight Test & Verification, and Simulator, Training & Support.

In addition to the ones already mentioned, I’d like to thank all my colleagues at System Simulation and Thermal Analysis at Vehicle Systems. The Friday coffee break with its rewarding discussions is the peak of my working week. I also want to thank my colleagues at the divisions of Machine Design and Fluid and Mechatronic Systems at Linköping University. It has been a great experience working together during the last five years.

To my parents Gerd and Lennart, thanks for your continuous support and for letting me know that I can do whatever I want in life. My greatest gratitude goes to my wife Martina and our wonderful daughters Selma, Maja, and Ellen. Martina, thanks for your patience and for your broad offering of home renovation projects letting my mind rest from work. To my daughters, thanks for all play and fun, and for keeping my feet on the ground.

Magnus Eek
Linköping, April 2016
The following papers are appended and will be referred to by their Roman numerals. A contribution statement is provided for each paper. The papers are printed in their originally published state, except for minor changes in formatting.


*Eek (formerly Carlsson) conducted the main developments related to model validation and was the main contributor to the manuscript writing. Andersson contributed regarding software development standards and software product lines. Gavel and Ölvander provided feedback.*


*Eek managed and carried out the major part of the case study and results analysis, and was the main contributor to the manuscript writing. Kharrazi contributed data collection from VTI, and assisted in results analysis and manuscript writing. Gavel and Ölvander provided feedback.*


*Eek (formerly Carlsson) carried out the major part of the analysis and was the main contributor to the manuscript writing. Steinkellner assisted in the analysis*
and manuscript writing, especially regarding model form uncertainty. Gavel and Ölvander provided feedback.


_Eek developed, implemented, and applied the proposed uncertainty aggregation method, and was the main contributor to the manuscript writing. Gavel and Ölvander provided feedback._


_Eek was the main contributor to the manuscript writing. Karlén contributed results from his master thesis work, defined and supervised by Eek. Ölvander provided feedback._


_Eek coordinated the development and evaluation, and was the main contributor to the manuscript writing. Hållqvist contributed implementations and assisted significantly in the evaluation and manuscript writing. Gavel and Ölvander provided feedback._
The following papers are not included in the thesis but constitute an important part of the background.


### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACARE</td>
<td>Advisory Council for Aeronautics Research in Europe</td>
</tr>
<tr>
<td>BIT</td>
<td>Built-In Test</td>
</tr>
<tr>
<td>BVP</td>
<td>Boundary Value Problem</td>
</tr>
<tr>
<td>CAS</td>
<td>Credibility Assessment Scale</td>
</tr>
<tr>
<td>CFD</td>
<td>Computational Fluid Dynamics</td>
</tr>
<tr>
<td>CSM</td>
<td>Computational Solid Mechanics</td>
</tr>
<tr>
<td>DAE</td>
<td>Differential-Algebraic Equation</td>
</tr>
<tr>
<td>DoD</td>
<td>Department of Defense</td>
</tr>
<tr>
<td>DSM</td>
<td>Design Structure Matrix</td>
</tr>
<tr>
<td>EASA</td>
<td>European Aviation Safety Agency</td>
</tr>
<tr>
<td>ECS</td>
<td>Environmental Control System</td>
</tr>
<tr>
<td>EW</td>
<td>Electronic Warfare</td>
</tr>
<tr>
<td>FMI</td>
<td>Functional Mock-up Interface</td>
</tr>
<tr>
<td>GECU</td>
<td>General systems Electronic Control Unit</td>
</tr>
<tr>
<td>GM-VV</td>
<td>Generic Methodology for Verification and Validation</td>
</tr>
<tr>
<td>GSA</td>
<td>Global Sensitivity Analysis</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical User Interface</td>
</tr>
<tr>
<td>HIL</td>
<td>Hardware-In-the-Loop</td>
</tr>
<tr>
<td>H/W</td>
<td>Hardware</td>
</tr>
<tr>
<td>IVP</td>
<td>Initial Value Problem</td>
</tr>
<tr>
<td>LSA</td>
<td>Local Sensitivity Analysis</td>
</tr>
<tr>
<td>MFL</td>
<td>Modelica Fluid Light</td>
</tr>
<tr>
<td>M&amp;S</td>
<td>Modeling and Simulation</td>
</tr>
<tr>
<td>OBOGS</td>
<td>On-Board Oxygen Generating System</td>
</tr>
<tr>
<td>ODE</td>
<td>Ordinary Differential Equation</td>
</tr>
<tr>
<td>SA</td>
<td>Sensitivity Analysis</td>
</tr>
<tr>
<td>SoS</td>
<td>System-of-Systems</td>
</tr>
<tr>
<td>S/W</td>
<td>Software</td>
</tr>
<tr>
<td>TLM</td>
<td>Transmission Line Modeling</td>
</tr>
<tr>
<td>UQ</td>
<td>Uncertainty Quantification</td>
</tr>
<tr>
<td>V&amp;V</td>
<td>Verification and Validation</td>
</tr>
<tr>
<td>VV&amp;UQ</td>
<td>Verification, Validation and Uncertainty Quantification</td>
</tr>
</tbody>
</table>
# Contents

1 Introduction ................................................................. 1
   1.1 Background................................................................. 1
   1.2 Industrial Objectives ................................................... 2
   1.3 Research Questions and Research Method ...................... 5
   1.4 Delimitations ............................................................... 7
   1.5 Related Research Projects ............................................ 8
   1.6 Thesis Outline ............................................................ 9

2 Theoretical Background .................................................. 11
   2.1 Credibility Assessment .................................................. 11
   2.2 Verification & Validation ................................................. 13
   2.3 Uncertainty Quantification .............................................. 14
      2.3.1 Types of Uncertainty ............................................... 15
      2.3.2 Sources of Uncertainty ............................................. 16
      2.3.3 Propagation of Uncertainty ....................................... 19
      2.3.4 Uncertainty Quantification Challenges in System Simulation ... 19
   2.4 Test Worthiness ........................................................... 21
   2.5 Philosophical Aspects on Credibility Assessment .......... 22

3 Modeling of Aircraft Vehicle Systems .............................. 25
   3.1 ODE and DAE Fundamentals ......................................... 26
   3.2 Methods, Languages, and Tools ..................................... 27
      3.2.1 Signal-Flow Modeling ............................................ 29
      3.2.2 Power-Port Modeling ............................................. 31
      3.2.3 Modelica ............................................................. 33
   3.3 Closed-loop Simulation ................................................ 36
   3.4 Industrial Application Examples .................................... 37
      3.4.1 Environmental Control Systems ............................ 37
      3.4.2 Liquid Cooling Systems ........................................ 39
      3.4.3 Large-Scale Aircraft System Simulators .................... 41
4 Contributions

4.1 Paper [I]: Methodology for Development and Validation of Multipurpose Simulation Models .......................................................... 43

4.2 Paper [II]: Study of Industrially Applied Methods for VV&UQ of Simulator Models .................................................................................... 46

4.3 Paper [III]: Enabling Uncertainty Quantification of Large Aircraft System Simulation Models .......................................................... 49

4.4 Paper [IV]: Definition and Implementation of a Method for Uncertainty Aggregation in Component-Based System Simulation Models .......... 51

4.5 Paper [V]: A Framework for Early and Approximate UQ of Large System Simulation Models .......................................................... 54

4.6 Paper [VI]: Development and Evaluation of a Concept for Credibility Assessment of Aircraft System Simulators ............................................. 58

5 Discussion and Conclusions

5.1 Organization and Independence ................................................................. 64

5.2 Automation ............................................................................................... 65

5.3 Approximate Uncertainty Quantification ................................................. 65

5.4 End-User Support for M&S Credibility Assessment .............................. 66

5.5 Answers to Research Questions ............................................................... 67

6 Outlook

References
Introduction

To what extent can we trust a simulation model? How well does a simulation model represent the real-world system of interest? To what extent can we use simulation as a complement to physical testing? Questions like these relate to model validation, and in a broader scope credibility assessment, and have been relevant ever since the first computerized simulation model showed up. Taking it one or two steps further, the problem of model validation dates back to the first scientific theory or even the first simple hypothesis ever formulated. Not surprisingly, from this wide perspective the scope of this research project is sharply limited.

1.1 Background

The use of Modeling and Simulation (M&S) in aircraft system design has a long history and is nowadays extensive. M&S is typically an integral part of the system development process and is applied to support a wide range of engineering activities all the way from conceptual design to detailed design and further on to system operation. As the aeronautical industry is becoming increasingly reliant on the use of M&S, efficient methods for Verification and Validation (V&V) of simulation models and simulator applications are essential (Lehmann, 2013). V&V of an individual simulation model undoubtedly includes challenging tasks, and building confidence in simulator applications in which several models are integrated is even more problematic.

The importance of V&V of simulation models is well known and the V&V research field has a long history, see for example Naylor et al. (1967) who propose a method named multi-stage verification, and Sargent (1979) who provides an overview of the subject and describes a set of validation techniques. In the aeronautical industry’s endeavor to reduce the system development time and the cost of physical testing, the challenging task of assessing a model’s validity is nonetheless of greater importance than ever.
In a broader perspective, model validation is only one factor in the assessment of the credibility of an M&S effort. In addition to model validation, a credibility assessment may consider various aspects such as Uncertainty Quantification (UQ), M&S management, a model’s use history, or the qualifications of model developers and end-users (NASA, 2008). The aim of this research is to explore and to further develop industrially applicable methods facilitating credibility assessment of system simulation models and simulator applications. Methods that facilitate early credibility assessment are of special interest. That is, methods applicable during preliminary design or early detailed design when the availability of system level measurement data for traditional model validation typically is very limited.

To end this introduction, the author cannot resist sharing a quote by Page et al. (1997), who with precision introduce the reader to some challenges a researcher in the field of V&V may face:

“V&V is misunderstood at all levels: To some managers, V&V is a panacea for the rising costs of software. To some developers, V&V looks like another means to further bureaucratize software development, giving management a raison d’être and placing even more roadblocks in the way of doing the really important (and interesting) work of actually producing software. To some academics, V&V is an avenue for publication and funding through the introduction of yet another “methodology” – without regard to its practicality.”

As Page et al. (1997) conclude, each of the above perceptions is erroneous. Nevertheless, the author’s endeavor to achieve V&V methods with industrial applicability cannot be overstated. As the research is closely connected to the industry, rather than finding the most comprehensive and theoretically correct method the emphasis is on methods suitable for more or less immediate industrial implementation and usability in daily engineering work, given the ever-present limitations in organizational resources, project budgets, and time schedules. Thus, the work presented here is aimed to contribute an applied and pragmatic engineering perspective on credibility assessment of M&S efforts in aircraft system development.

1.2 Industrial Objectives

Major drivers for utilizing M&S in aircraft system development are to enable early model-based design decisions, early detection of design errors, and to reduce the cost of physical testing. Furthermore, in the effort to reduce the cost of physical testing specifically related to the certification process, the aeronautic industry strives to expand the usage of M&S further by introducing the concept of virtual testing. While no compact and broadly agreed definition has been found, the term virtual testing here refers to the structured use of
M&S to critically evaluate a product’s design against specified requirements. In the case of certification, the requirements are set by certification authorities, typically the Federal Aviation Administration (FAA)\(^1\) in the USA or the European Aviation Safety Agency (EASA)\(^2\) in Europe. When virtual testing is used as an Acceptable Means of Compliance (AMC) for certification, this may be termed *virtual certification*.

The primary motivator for research on credibility assessment of M&S efforts is risk reduction. Model validation, which is a vital component in credibility assessment, can be regarded as risk reduction in terms of ensuring that models are actually suitable for their intended use. Industrially applicable methods and tools for credibility assessment are therefore critical to justify the use of M&S, not least when it comes to aircraft certification. Related to the challenge of maintaining and extending the competitiveness of the European aviation industry, the Advisory Council for Aeronautics Research in Europe (ACARE) states the following vision for 2050 (ACARE, 2011):

> “Streamlined systems engineering, design, manufacturing, certification and upgrade processes have addressed complexity and significantly decreased development costs (including a 50% reduction in the cost of certification). A leading new generation of standards is created.”

Indeed, a main enabler identified for attaining this ambitious goal is virtual certification (ACARE, 2012). Reasonably, a central challenge will be to develop and establish uniform methods to convince one’s own company as well as the certification authority of the credibility of M&S efforts.

Like other aircraft developers, Saab Aeronautics is continuously facing challenges of increasing product complexity and market competition. With the aim to increase development efficiency and thereby respond to these challenges, Saab Aeronautics invests in processes, methods, and tools for Model-Based Systems Engineering (MBSE). For a successful deployment of MBSE, efficient and industrially applicable methods for credibility assessment in general and model validation in particular are essential prerequisites. To utilize the full potential of a model – in the extreme case for use as an AMC in certification – the credibility assessment method must consider the impact of uncertainties hidden in, for example, model inputs or parameters, or due to model simplifications or numerical approximations. It is also important that the credibility assessment method is easy to apply and iterate as new information becomes available. A major challenge is to find credibility assessment methods applicable for industry grade simulation models.

At Saab Aeronautics, system simulation is applied on several different scales; from models of individual components such as a valve, a pump, or a heat exchanger, to models

\(^1\) [http://www.faa.gov](http://www.faa.gov)

of subsystems like a hydraulic system, a fuel system or an Environmental Control System (ECS), and further on to large-scale aircraft system simulators in which a large number of individual models are connected and simulated together. The use of aircraft system simulators in system development and verification requires thorough documentation of each individual model, and at Saab Aeronautics there is an established process for test worthiness declaration. However, despite current efforts related to configuration management and model documentation, it is difficult for simulator users (typically system engineers and test engineers) to assess the credibility of the simulation results. As this is a critical capability for the successful use of MBSE, methods and tools facilitating a simplified and less document-centric credibility assessment are needed.

Development of new aircraft and subsystems typically takes several years, implying that simulation models are developed and used long before system level measurement data for traditional model validation purposes are available. Therefore, in early phases like conceptual design and preliminary design, UQ is seen as an essential – but currently poorly utilized – means to build confidence in aircraft system simulation. As the system development continues, the availability of measurement data increases – starting with measurement data for separate equipment, on to test rig data, and in later phases flight test data. However, also during late detail design, measurement data may be relatively scarce and it is often difficult to obtain satisfactory validation coverage of a model’s operational domain. In such cases, UQ is a useful support also in later development phases.

A conceptual view of the applicability of UQ and traditional model validation using measurement data is given in Figure 1-1. The faded gray area in the figure indicates an approximate interval during system development when a combination of UQ and traditional model validation using measurement data is seen as especially suitable. In this interval, there is typically measurement data available at component level (e.g. pumps, valves, etc.) but not on subsystem level (e.g. hydraulics or fuel systems). Reasonably, it should be possible to utilize the knowledge gained during component level validation to support credibility assessment on model level. Figure 1-1 is also intended to visualize the context of this research project by providing rough and generalized information of the different types of models and simulators used during system development and operation.
To summarize this section, the objective of this research project is to develop industrially applicable methods for credibility assessment, focusing on the model validation aspect in early development stages when system level measurement data is scarce. The research considers methods applicable for models representing individual subsystems as well as for aircraft system simulators in which several individual models are integrated.

1.3 Research Questions and Research Method

Based on the above industrial objectives and with further input from both industry and academia, the following research questions have been defined:

**RQ1** How can model validity be assessed in the absence of system level measurement data?
RQ2  How can uncertainty information obtained on component or sub-model level be used to assess uncertainty on model top level?

RQ3  Which model validation techniques are suitable for use in an industrial applicable process for development and integration of aircraft vehicle system models?

RQ4  How can credibility information for individual models and for a simulator as a whole be presented to simulator users to support the detection of deficiencies in model representativeness – prior to, during, and after a simulator run?

RQ5  Which credibility measures and which level of detail in the credibility assessment is industrially applicable for large-scale aircraft system simulators?

Regarding methods for early validation of individual simulation models, RQ1 is a central but yet very broad research question. The research has involved a continuous process of limiting the scope and refining the problem formulation, resulting in RQ2 and RQ3. Regarding methods for credibility assessment of simulator applications, RQ4 is a central research question from which RQ5 has been derived.

The research has been carried out using the industry-as-laboratory approach, as proposed by Potts (1993) and further described by Muller (2013). That is, the research problem is defined by the industry and solutions from research are applied and evaluated using an actual industrial setting as test environment. Figure 1-2 shows the principle of industry-as-laboratory (Muller, 2013).

Figure 1-2: The principle of industry-as-laboratory; i) research problems are provided by the industry, ii) hypotheses and new methods are developed and applied on an industrial setting, iii) results are observed and evaluated, and iv) new hypotheses are developed and methods are improved.

The test environment or “application playground” used for development and evaluation of solution alternatives to the above research questions consists of industry grade
simulation models of individual aircraft subsystems as well as large-scale aircraft system
simulators, as described in section 3.4. As part of the industry-as-laboratory approach, the
presented research also includes an exploratory case study with two units of analysis (Yin,
2009). Throughout this research project, the recurrent evaluation step in Figure 1-2 has
been carried out in various ways, such as analysis of simulation results, demonstrations,
workshops, interviews, and surveys. For analysis of open-ended survey questions, a
structured team approach has been applied, inspired by the Affinity Diagram Technique
(Kasunic, 2005). More details on the qualitative methods used can be found in papers [II]
and [VI].

As the research has close connections to the industry, the research is fairly close to
application and the method used has similarities with a typical engineering design process
used in product development. However, from an engineering perspective, science may also
be seen as product development, in which case the name of the product is “new knowledge”.
There are both deductive and inductive views of the research method used. To use an
identified scenario and industrial setting as a general starting point for the development
of specific credibility assessment methods would imply a deductive view. On the other
hand, the methods developed are evaluated in a specific scenario using a specific simulation
model or simulator in a specific engineering domain. Developed methods are conceived to
be useful within a wide range of applications, but no formal proofs of further applicability
are available. Since it is not only strict mathematics that leads to the research results, the
research is in general seen as inductive rather than deductive. For further discussions
related to research methodology, see section 2.5.

To benefit from the skills in other domains also dealing with development and use of
large-scale system simulators and to increase the likelihood of developing more generalized
methods applicable also outside aircraft system simulation, the research has been carried
out in collaboration between Saab Aeronautics, Linköping University, and the Swedish
National Road and Transport Research Institute (VTI).

1.4 Delimitations

Since M&S is a diverse field used in numerous scientific and engineering disciplines, and
may be implemented and executed using a vast amount of different techniques, it should
be noted that this work is focused on what is here referred to as system simulation. This
term is here understood as the simulation of mathematical models representing behavioral
aspects of physical systems with or without controlling software. Such systems are typically
described by Ordinary Differential Equations (ODEs) or Differential-Algebraic Equations
(DAEs). As a consequence, in the following the term model should be interpreted as system
simulation model.

The research field of V&V, and in a broader scope M&S credibility assessment, is
multifaceted and interdisciplinary, but in practice the methods used are often, and by
necessity, domain specific. Compared to the diverse field of system simulation, Computational Fluid Dynamics (CFD) and Computational Solid Mechanics (CSM) are two domains that have come far regarding V&V methodology. Thus, even though the work presented in this thesis is not focused on CFD or CSM, some methods from these domains are considered and in a few cases applied and evaluated on system simulation applications.

As already mentioned, model credibility assessment includes several vital aspects. This research is focused on methodology for model validation, which for early development phases also relates to UQ. To clarify, a proper model verification which in itself may be a very challenging task, is a prerequisite for a successful application of the methodology proposed herein. Definitions of terminology used throughout this thesis can be found in chapter 2.

Even though the credibility assessment methods proposed may be of interest for use in training simulators, they are primarily aimed towards models and simulator applications used in system development.

1.5 Related Research Projects

The research is mainly sponsored by Saab Aeronautics and the National Aviation Engineering Research Programme (NFFP). Most of the work presented in this thesis has been carried out in the NFFP5 project Validation of Complex Simulation Models and its continuation and extension NFFP6 Model Validation – from Concept to Product.

The NFFP4 project Modeling and Simulation for the 2010s Energy Management Systems can be seen as a predecessor that has formed the context and provided significant input to the research (Steinkellner, 2011).

There has also been a fruitful collaboration with the NFFP5 project Heterogenous Modeling and Simulation Techniques, focusing on reuse and configuration of simulation models through a product line approach (Andersson, 2012).

In addition, an active part has been taken in the European Union Seventh Framework Programme project CRESCENDO, a research project with more than 60 partners from European aerospace industry and academia. The project has focused on development of processes, methodologies and tools enabling collaborative design, covering a major part of the product lifecycle including Virtual Testing and Virtual Certification (CRESCENDO, 2012).

Related research continues in the EUREKA/ITEA3 project OPENCPS, which generally aims to establish interoperability between the modeling languages Modelica and UML by the use and possible extension of the Functional Mock-up Interface (FMI) standard (OPENCPS, 2016).
1.6 Thesis Outline

This work is presented in the form of a compilation thesis, consisting of an introductory summary and a number of appended papers. In general, the introductory summary is intended to provide a context to the appended papers and to summarize essential theory and results. For details, the reader is referred to the appended papers. The introductory summary also contains parts of the scientific basis not contained in the appended papers, for example regarding research questions and general research method. Table 1-1 provides an overview of the appended papers, how they are related to each other and to the research projects mentioned in the previous section. The appended papers are largely ordered in terms of methodological scope, i.e. from methods applicable on subsystem (model) level to methods applicable on system (simulator) level.

Table 1-1: Overview of appended papers, research methods, and research projects.

<table>
<thead>
<tr>
<th>Paper number</th>
<th>Content and relation to other appended papers</th>
<th>Research method</th>
<th>Contributing research projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>[I]</td>
<td>Describes a process for development, V&amp;V, and export of multipurpose simulation models, providing a context to the remaining papers.</td>
<td>Collaborative interdisciplinary effort, including workshops with model and simulator developers.</td>
<td>NFFP5 Heterogenous M&amp;S Techniques, NFFP5 Validation of Complex Simulation Models, and CRESCENDO.</td>
</tr>
<tr>
<td>[II]</td>
<td>Case study to identify industry’s current best practices as well as gaps/needs in VV&amp;UQ of simulation models and simulator applications. Identified needs are simplified UQ methods (see papers [III]-[V]) and enhanced user support for simulator credibility assessment (see paper [VI]).</td>
<td>Exploratory case study with two units of analysis, including surveys and interviews. Analysis of open-ended survey questions inspired by the Affinity Diagram Technique.</td>
<td>NFFP6 Model Validation – from Concept to Product</td>
</tr>
<tr>
<td>[III]</td>
<td>Describes an established UQ method from the CFD domain, and estimates the work effort required to apply the method on a large simulation model of an individual aircraft subsystem. Further demonstrates the need for simplified UQ methods and discusses possible solution/mitigation alternatives.</td>
<td>Industry-as-laboratory, including literature study and team-based estimation of required work effort using an industry grade simulation model as test environment.</td>
<td>NFFP5 Validation of Complex Simulation Models</td>
</tr>
</tbody>
</table>
Motivated by papers [II] and [III], this paper describes the definition, implementation, and application of an approximate uncertainty aggregation method for use in component-based system simulation. Industry-as-laboratory, focusing on modeling and implementation in Modelica. Evaluation using an aircraft liquid cooling system simulation model as test environment.

As a response to the needs identified in papers [II] and [III], a framework for approximate UQ of large simulation models representing several connected aircraft subsystems is proposed. The framework includes the aggregation method described in [IV]. Industry-as-laboratory, using a connected set of aircraft subsystem models as test environment.

Responds to the needs identified in paper [II] by describing the development and evaluation of a methodology and a related tool providing enhanced user support in credibility assessment of large-scale simulator applications. The tool makes use of information obtained during the validation of individual simulation models. Collaborative interdisciplinary effort including model developers, model integrators, simulator users, methodologists, and decision makers. Evaluation using simulator testing, surveys, and workshops.

The remaining part of the introductory summary is outlined as follows: Chapter 2 provides a brief theoretical background to M&S credibility assessment and its main constituents such as verification, validation, and UQ. Some philosophical aspects of credibility assessment are also discussed. In Chapter 3, techniques and languages commonly used for modeling of aircraft vehicle systems are discussed. The chapter also provides a context in terms of field of application, and describes the main industrial application examples used in the research. Chapter 4 clarifies the contributions of each appended paper, while Chapter 5 provides a discussion, conclusions, and a summarized answer to the research questions. Finally, chapter 6 suggests possible directions for future work.
2

Theoretical Background

There is an extensive amount of M&S literature proposing definitions, conceptual frameworks, or standards related to V&V. See, for example, the literature review given in (Bair and Tolk, 2013), or in (Oberkampf and Roy, 2012) which provides a historical overview of V&V developments in different communities such as AIAA, ASME, and IEEE. The intention of this chapter is not to provide a comprehensive documentation of the V&V field. Rather, this chapter provides a brief introduction to credibility assessment in M&S, and related concepts and terminology used in the appended papers.

The remaining part of this chapter is outlined as follows. First, the umbrella term credibility assessment is explained. Verification and validation are two areas essential in any credibility assessment framework and are therefore briefly described. As this research is directed towards validation rather than verification, the emphasis in the provided description is on validation. Furthermore, this research is focused on validation in early development phases when measurement data for traditional model validation purposes are scarce. In such situations, UQ is here regarded as a main enabler in building confidence in the simulation results, and is thus described on a somewhat more detailed level. The final section of the chapter contains some philosophical reflections on credibility assessment.

2.1 Credibility Assessment

Credibility assessment of an M&S effort may include several aspects such as verification, validation, UQ, M&S management, the qualifications of modelers and end-users, and the model’s use history. In the current taxonomy, credibility is a relatively wide and vague term. For example, a legacy model with deficient V&V documentation but with an extensive use history may be considered to have higher credibility than a new model that has gone through extensive and well-documented V&V. The following two definitions of credibility are given by NASA and ITOP, respectively:
• “The quality to elicit belief or trust in M&S results.” (NASA, 2008)
• “The subjective belief that a model or simulation of interest satisfies some requirement and serves its purpose.” (ITOP, 2004)

Both definitions indicate either directly or indirectly that credibility involves a degree of subjectivity. In 2004, Pace pointed out qualitative assessment in V&V as an area that needs to progress in terms of repeatability and credibility (of the assessment itself) (Pace, 2004). To succeed in this, subjectivity must be kept to a minimum. Challenges remain but important steps have been taken in the development of methods to assess M&S credibility. One of the more recent and well-known examples is the Credibility Assessment Scale (CAS) developed by NASA (NASA, 2008). In the application of the CAS, M&S results are evaluated on each of eight factors grouped into the following three categories: 1) M&S Development including the two factors Verification and Validation, 2) M&S Operations with three factors related to different aspects of UQ – namely Input Pedigree, Results Uncertainty, and Results Robustness, and 3) Supporting evidence focusing on Use History, M&S Management, and People Qualifications. For detailed information and guidance on how to apply the CAS, please refer to (NASA, 2008) and (NASA, 2013). An interpretation and modification of the CAS towards application in M&S for aircraft design can be found in (Vincent et al., 2012). Two additional methods related to credibility assessment are the Predictive Capability Maturity Model proposed by Sandia National Laboratories (Oberkampf et al. 2007), and the Validation Process Maturity Model proposed by Harmon and Youngblood (2005). A brief summary of these methods is provided in paper [I].

In general, each method mentioned above defines a set of aspects for consideration. These aspects are then rated to produce one or more overall credibility scores of the assessed M&S effort. Note that there is not a one-to-one correspondence between a credibility score and the accuracy of M&S results. An improved score would commonly imply greater accuracy, but this is not necessarily true. Typically, different credibility aspects also have different characteristics. Aggregating scores of different aspects into one overall score may thus be deceptive.

When managing a large number of models, e.g. for integration in simulator environments, one typically wants models checked-in to a model storage to have an assessed credibility. An attractive idea may be to perform a model level credibility assessment and attach the results to each model prior to check-in. However, this is only possible for the subset of credibility aspects related to the model itself. As an example, people qualification of end users is not a characteristic of the model. Major aspects possible to handle as model characteristics are verification, validation, and UQ, which are discussed in the following sections.
An interesting method identified in paper [II] and further developed in paper [VI] is the use of a credibility measure embedded in a model, i.e. an output signal describing one or more aspects of model credibility. It is an appealing idea that a model itself can provide information about selected aspects of its own credibility – and not only by static information prior to simulation, but also by dynamic information in real-time during simulation.

2.2 Verification & Validation

Several definitions of the terms verification and validation exist. As formulated by Balci (1997), verification concerns building the model right, i.e. determining whether the model is compliant with the model specification and if it accurately represents the underlying mathematical model, whereas validation concerns building the right model, i.e. determining whether the model is a sufficiently accurate representation of the real system of interest. Both verification and validation should be performed with respect to the intended uses of the model. Thus, well-defined intended uses significantly enhance the outcome of V&V efforts. Balci and Ormsby (2000) discuss the importance of intended uses and provide examples of hierarchically defined intended uses. Piersall III and Grange (2014) point out that well-defined intended uses are especially important in System-of-Systems (SoS) contexts. It should be noted that a simulator application, in which several individual models are connected, may be considered to be an SoS. The brief description of V&V terminology provided here is in line with definitions used in e.g. (NASA, 2008), (DoD, 2007), and (ITOP, 2004). A recent development in which a number of established V&V related definitions are collected is the Generic Methodology for Verification and Validation (GM-VV) (SISO, 2013). The GM-VV provides a generic and comprehensive framework for V&V of M&S assets. See (Roza et al., 2012) for an introduction to the GM-VV and for illustrative examples of how the GM-VV may be tailored to specific M&S applications.

Given the above description of validation, any validation effort needs information on i) the level of realism required by the model and ii) the level of realism provided by the model. This “level of realism”, or in other words how well the model represents the real-world system, is a short and vague description of fidelity – a disputed term surrounded by considerable confusion in the V&V field. In line with (Gross, 1999), the following definition of fidelity is given in GM-VV (SISO, 2013): “The degree to which a model or simulation reproduces the state and behavior of a real world object or the perception of a real world object, feature, condition, or chosen standard in a measurable or perceivable manner; a measure of the realism of a model or simulation; faithfulness. Fidelity should generally be described with respect to the measures, standards or perceptions used in assessing or stating it”. See Roza (2004) who provides an extensive literature review on the subject and proposes a unified fidelity framework.
Balci (1997) lists more than 75 V&V techniques divided into four groups; informal, formal, static, and dynamic. These are further described in (Balci, 1998). Another well-established set of validation techniques is provided by Sargent; see (Sargent, 2013) for an up-to-date version. Informal techniques like face validation and reviews are generic and may concern both verification and validation. When applying face validation, the model is handed over to an expert or a potential user for subjective comparison of model and real-world system behavior. Informal techniques are mainly based on human reasoning, and are of great importance and often easy to apply. Note that the word informal does not prevent these techniques being well-structured. Formal techniques based on mathematical proof of correctness may also cover both verification and validation aspects. However, as indicated by Balci (1998), formal methods are rarely applicable where complex simulation models are concerned.

Static techniques concern analysis of the static model design and source code, and do not require execution of the model. Static techniques like interface analysis and structural analysis are directed more towards verification than validation. A fundamental tool for static verification is the model language compiler itself. What remains is the group of dynamic techniques that require model execution.

Dynamic techniques commonly used are predictive validation, sensitivity analysis, and regression testing. In predictive validation, the model is fed with measurement data of the input to the real-world system and the predictive ability of the model is assessed by comparing model output against measurement data of the quantities of interest of the real-world system. Sensitivity Analysis (SA) is the study of how variation in the output of a model can be apportioned to different sources of variation, and how the given model depends upon the information fed into it (Saltelli et al., 2000). Regression testing is a method for identifying differences between simulation results from an updated model and an earlier version of the same model, with the purpose of verifying that the model update only affects the output signals that are intended to be affected, i.e. that the update has no adverse side-effects.

It is not always easy to determine whether a specific technique is directed towards verification or towards validation. It should also be noted that the taxonomy of informal, formal, static, and dynamic in (Balci, 1997) is not necessarily orthogonal. There may, for example, exist informal dynamic techniques. Actually, face validation, which is found in the group of informal techniques, can be seen as an example of this.

### 2.3 Uncertainty Quantification

UQ refers to the process of identifying, quantifying, and assessing the impact of uncertainty sources embedded along the development and usage of simulation models. UQ may be seen as an integral part of model validation, but sometimes the term VV&UQ is used to explicitly point out that UQ is considered. UQ is an interdisciplinary field which has
attracted increased interest in recent years. As summarized by Smith (2014); “The present novelty lies in the synthesis of probability, statistics, model development, mathematical and numerical analysis, large-scale simulations, experiments, and disciplinary sciences to provide a computational framework for quantifying input and response uncertainties in a manner that facilitates predictions with quantified and reduced uncertainty”.

Typically, UQ is closely linked to SA and there are numerous and sometimes overlapping combinations of methods for UQ and SA. It is therefore not an easy task to build a proper taxonomy of methods. As an attempt to separate the two, a main difference is that SA does not necessarily require a quantitative characterization of the input uncertainties. Therefore, even though a model output may be very sensitive to perturbations in a specific input, the uncertainty in the output is not necessarily large if the input uncertainty is very small, and vice versa.

Local SA (LSA) is normally used to study first-order effects of small variations in input, i.e. how the model output varies considering one input variation at a time. Normally, LSA is carried out by calculation of output derivatives with respect to inputs, and the word local indicates that the analysis is carried out locally around a nominal value of an input. For studying higher-order effects, non-linear models, and larger input variations, each input is typically treated as a random variable and assigned a Probability Density Function (PDF). In this case, the analysis is referred to as global SA (GSA), and is often carried out using Monte Carlo based sampling techniques (Weiße, 2009).

For fairly linear models, LSA may also be used for larger variations in input. In this case, the inputs may be treated as bounded intervals. Similar studies may be carried out using simplified GSA, only considering the bounds of the input intervals. Below, such an approach is referred to as interval-based SA. Saltelli et al. (2008) provide a comprehensive introduction to methods for GSA, and de Rocquigny et al. (2008) give an overview of methods for UQ in an industrial context. See also Zhang et al. (2014) who provide a concise introduction to SA.

Precise, standardized, and cross-domain established definitions of the terms error and uncertainty are hard to find. In this work, error primarily regards implementation deficiencies or mistakes (commonly known as “bugs”) while uncertainty is related to the model’s representation of the physics of the real-world system.

2.3.1 Types of Uncertainty

Commonly, a distinction is made between aleatory uncertainty (due to statistical variations, also referred to as variability, inherent uncertainty, irreducible uncertainty, or stochastic uncertainty), and epistemic uncertainty (due to lack of information, also referred to as reducible uncertainty or subjective uncertainty), (Helton, 1994, Helton, 1996). Epistemic uncertainty is normally represented by an interval, and aleatory uncertainty by a PDF. Mixed epistemic and aleatory uncertainty may be represented by an imprecise
PDF, which here refers to when the parameters of the PDF, for example the mean and standard deviation, are given as intervals or PDFs themselves (Roy and Oberkampf, 2011).

In practice, it may be hard to conclude whether a particular uncertainty should be characterized as epistemic or aleatory. Defining an uncertainty as purely aleatory, in the meaning of *irreducible*, may require a great deal of certain information. One interpretation is that the UQ analyst needs all information necessary to conclude that the uncertainty cannot be reduced by means of any further investigations or physical testing. Another viewpoint is that the characterization depends on the given context, and may change depending on the type of UQ to be performed as well as at which point in time the UQ is carried out (Kiureghian and Ditlevsen, 2009).

### 2.3.2 Sources of Uncertainty

Pace (2009) stresses the importance of a comprehensive treatment of simulation uncertainty and proposes a seven-part paradigm. Pace claims that this paradigm addresses all possible aspects of uncertainty in simulation use. The seven uncertainty aspects are listed and briefly exemplified below. For further descriptions, see (Pace, 2013).

1) *Simulation application domain*: e.g. variability of physical entities or boundary conditions represented by the model, or lack of understanding of underlying physical principles.

2) *Simulation purpose*: e.g. imprecise, incomplete, or contradicting intended uses.

3) *Simulation concept to design*: e.g. uncertainties due to decisions taken during the modeling process, like assumptions or approximations.

4) *Simulation implementation*: e.g. software errors ("bugs").

5) *Simulation inputs and use*: e.g. uncertain input signals or model parameters, or due to user effects when setting up and running a simulation.

6) *Validation referent information*: e.g. uncertainty in an experimental setup or in measurement data used as validation referent.

7) *Interpretation of simulation results*: e.g. unrecognized uncertainty or implicit assumptions made when analyzing simulation results.

Roy and Oberkampf provide an example of a well-structured framework for UQ (Roy and Oberkampf, 2011). According to Roy and Oberkampf, all uncertainties originate from three key sources:

1) *Model inputs*: e.g. input signals, parameters, and boundary conditions.

2) *Numerical approximations*: e.g. due to the numerical method used by the solver.

3) *Model form*: e.g. model simplifications or uncertainty in underlying equations.
These three uncertainty sources are also in line with the definitions provided by Coleman and Steele (2009). As seen above, the uncertainty sources defined by Roy and Oberkampf cover large parts of the aspects defined by Pace. However, there is not a one-to-one mapping and Pace notes that the framework by Roy and Oberkampf does not explicitly consider user effects and uncertainties in intended use.

Coleman and Steele (2009) propose the following formalization of the relations between the simulation result $S$, the experimental validation data $D$, the validation comparison discrepancy or residual $E$, and the true but always unknown value $T$. The uncertainty in the simulation result $\delta_S$ and the uncertainty in the validation data $\delta_D$ are also defined. The equation variables may be either time-series or single values, such as steady-state values:

$$ E = S - D \quad (2.1) $$

$$ \delta_S = S - T \quad (2.2) $$

$$ \delta_D = D - T \quad (2.3) $$

Hence, $E$ is the combination of all uncertainties in the simulation result and in the validation data.

$$ E = (\delta_S + T) - (\delta_D + T) = \delta_S - \delta_D \quad (2.4) $$

With the three simulation uncertainty sources described by Roy and Oberkampf (2011), the uncertainty in the simulation result can be defined as follows.

$$ \delta_S = \delta_{S\text{\_input}} + \delta_{S\text{\_num}} + \delta_{S\text{\_model}} \quad (2.5) $$

In addition to sensor measurement uncertainty (which may also include A/D conversion that implies finite resolution), the total uncertainty in validation data may depend on various characteristics of the physical test setup, e.g. uncertain boundary conditions, experimental simplifications, or placement of sensors. An example might be when comparing air stream temperatures obtained from a 1-D simulation model with experimental results. In such a case, the model typically does not take into account local effects and inhomogeneous flow patterns. Therefore, to obtain useful validation data, placement of the temperature sensor should be chosen carefully, e.g. in terms of downstream distance from a mixing point or radial positioning in a pipe. To emphasize this, the uncertainty in the experimental validation data can be defined as the combination of sensor measurement uncertainty $\delta_{D\text{\_sensor}}$ and uncertainty due to the experimental setup $\delta_{D\text{\_setup}}$: 
\[ \delta_D = \delta_{D\text{ sensor}} + \delta_{D\text{ setup}} \]  

(2.6)

In some sense, the uncertainty due to the experimental setup \( \delta_{D\text{ setup}} \) can be seen as the experimental counterpart to the model form uncertainty of the simulation, \( \delta_{S\text{ model}} \). Equations (2.4) to (2.6) can be combined to show the individual contributions to the residual \( E \):

\[
E = (\delta_{S\text{ input}} + \delta_{S\text{ num}} + \delta_{S\text{ model}}) - (\delta_{D\text{ sensor}} + \delta_{D\text{ setup}}) \]  

(2.7)

Equation (2.7) demonstrates the possibility of subtractive uncertainty cancellation resulting in \( E = 0 \). Thus, if the sum is zero, this does not necessarily mean that the simulation is correct. Rewriting Equation (2.7) to solve for \( \delta_{S\text{ model}} \) shows similarly that subtractive uncertainty cancellation may lead to misinterpretation of the model form uncertainty:

\[
\delta_{S\text{ model}} = E + (\delta_{D\text{ sensor}} + \delta_{D\text{ setup}} - \delta_{S\text{ input}} - \delta_{S\text{ num}}) \]  

(2.8)

It should be noted that \( \delta_{Sx} \) and \( \delta_{Dx} \) above are single realizations of their related uncertainty distributions (index \( x \) indicates any uncertainty source from Equations (2.5) and (2.6) respectively). The related uncertainty distributions are here denoted \( U_{Sx} \) and \( U_{Dx} \) respectively. Whereas the single uncertainty realizations are typically unknown, their related distributions are sometimes known (i.e. can be estimated individually as aleatory PDFs or epistemic intervals).

If an unknown related uncertainty distribution is to be indirectly estimated from a number of known contributing uncertainties, care should be taken regarding how to combine the contributors. If the individual contributing uncertainties \( U_{S\text{ input}} \), \( U_{S\text{ num}} \), \( U_{D\text{ sensor}} \), and \( U_{D\text{ setup}} \) are known, then Equation (2.8) can be expressed as:

\[
U_{S\text{ model}} = E \pm \Delta \]  

(2.9)

where \( \Delta \) is the combination of the contributors. One alternative for calculating \( \Delta \) is to use the square root sum of squares, in line with what is proposed in (Coleman and Steele, 2009) and (Coleman and Stern, 1997):

\[
\Delta = \sqrt{U_{S\text{ input}}^2 + U_{S\text{ num}}^2 + U_{D\text{ sensor}}^2 + U_{D\text{ setup}}^2} \]  

(2.10)

However, since the known uncertainty contributors are necessarily neither independent nor random, the square root sum of squares is not always applicable. In such situations, a
more relevant method may be to use the conservative combination of the absolute values of each individual uncertainty contributor. That is, to treat each individual uncertainty contributor as an interval-valued epistemic quantity:

\[ \Delta = |U_\text{input}| + |U_\text{num}| + |U_{\text{sensor}}| + |U_{\text{setup}}| \] (2.11)

A detailed discussion on the implications of combining uncertainties is provided by Oberkampf and Roy (2012) who give an example of how to calculate the total numerical uncertainty of a simulation result given a number of individual uncertainty contributors.

2.3.3 Propagation of Uncertainty

During propagation of input uncertainty, a separate treatment of aleatory and epistemic uncertainty may be accomplished by using a nested Monte Carlo sampling strategy with an outer loop for the epistemic uncertainties and an inner loop for the aleatory uncertainties. The result of nested sampling is one Cumulative Distribution Function (CDF) for each combination of epistemic input. The complete set of CDFs for a model output is referred to as a probability box or p-box (Roy and Oberkampf, 2011). Nested sampling is in contrast to single-loop sampling, in which all uncertainties are treated as aleatory, resulting in a single CDF for the complete uncertainty propagation. In the context of aircraft system design using computationally demanding models that may include a large number of epistemic uncertainties, nested sampling is normally not feasible. It is also noted by Roy and Oberkampf that “for more than a handful of epistemic uncertainties, the total number of samples required for convergence becomes extraordinarily large, and other approaches should be considered” (Roy and Oberkampf, 2011). Alternative approximate single-loop sampling methods which reduce the number of samples significantly are proposed by Hofer et al. (2002) and Krzykacz-Hausmann (2006).

Regardless of the sampling technique chosen, for computationally demanding system simulation models it is always advantageous to reduce the number of uncertainties for consideration as far as possible, and thereby reduce the dimensionality of the UQ problem.

2.3.4 Uncertainty Quantification Challenges in System Simulation

An appealing approach used in (Roy and Oberkampf, 2011) is the separate treatment of aleatory and epistemic uncertainty. Also, the contributions from numerical uncertainty and model form uncertainty to the total uncertainty are explicitly taken into account. However, from a practical point of view, it should be noted that input uncertainty, numerical uncertainty, and model form uncertainty are not always independent. The choice
of model structure, for example the choice of equations to describe the pressure loss in a pipe, obviously affects model form uncertainty, but since the choice of equations determines which parameters exist in a model, the input uncertainty is also affected. Furthermore, the choice of equations as well as parameter values may affect the numerical uncertainty; for example changing a parameter value may cause numerical instability. Due to these dependencies between different sources of uncertainty, the applicability of a separate treatment is in some cases questionable, especially in early development phases when the basis for a comprehensive UQ is limited. In other words, a single uncertainty estimate of a simulation result in which the contributions from input, numerical, and model form uncertainty are lumped together is better than no uncertainty estimate at all. This is also the selected approach in the uncertainty aggregation method described in paper [IV].

Characterization of uncertainties may be a demanding task in terms of the work effort required. The characterization process is typically very human-/skill-intensive and may require extensive physical testing, even for a low number of uncertainties. This is highlighted in paper [III] by providing an estimation of the work effort required to apply extensive UQ on a detailed Environmental Control System (ECS) model. That characterization of uncertainties in general is time-consuming and expensive is also recognized by Saltelli et al. (2008) and Helton et al. (2006). The challenge of limited information availability for the characterization of input uncertainties is also pointed out by Zhang et al. (2014).

The engineering work effort required for UQ mainly depends on the dimensions M&S complexity and Information availability shown in Figure 2-1, i.e. the complexity of the model to be analyzed and the availability of information for uncertainty characterization. Regarding computational cost, the critical dimensions are M&S complexity and UQ scope, where the latter describes the comprehensiveness of the UQ technique applied.
To summarize, for any non-trivial industrial M&S application, UQ often implies challenges related to computational cost but in the context of this research the greatest challenge is the engineering work effort required for UQ. From a pragmatic engineering point of view, and as for any M&S effort in general, a central challenge is to find a suitable level of detail also when performing UQ. That is, to simplify the UQ as far as possible without affecting the UQ results in such a way that they become irrelevant.

2.4 Test Worthiness

A term closely related to credibility, sometimes used in the ground, flight and simulator testing communities in the aeronautics industry, is test worthiness. Formal simulator testing for system verification purposes requires a test worthiness declaration to be performed. The test worthiness declaration may well be considered a special type of credibility assessment.

Figure 2-2 shows a simplified view of the current process for test worthiness declaration of simulator applications at Saab Aeronautics. When creating an aircraft system simulator configuration, specific versions of subsystem simulation models are selected from the model storage. In accordance with the applied process for model export and simulator integration,
each simulation model in the model storage comes with a status declaration (Andersson and Carlsson, 2012). The purpose of the status declaration is to describe the current status of the model and its related documentation. It refers to applicable documentation of the model and states, for example, model version, general information, model purpose, known limitations, and conclusions drawn from V&V activities. As for each model, a status declaration is also to be written at simulator level, providing aggregated information and referring to the status declaration of each model. The simulator status declaration, together with the aircraft configuration (which the simulator should represent) and the test plan form the main inputs to the assessment of test worthiness. In the event of an unexpected test result or due to other issues noted during testing, the assessment of test worthiness may have to be reiterated, which is indicated by the circular arrow in the figure.

![Figure 2-2: Simplified view of today’s workflow for test worthiness declaration.](image)

A sound assessment of simulator test worthiness requires a great deal of experience and knowledge of the aircraft subsystems, the corresponding simulation models, the simulator architecture, and the testing to be performed. Obviously, the current process is rather document-centric and involves some degree of subjectivity. To capture all links and details needed, the current level of documentation is considered necessary. However, there is room for methods to decrease the level of subjectivity and to present specific information about the simulation models in a more automated way. Such a method has been proposed in paper [VI], and is at the time of writing in the process of being implemented at Saab Aeronautics.

### 2.5 Philosophical Aspects on Credibility Assessment

An overview of philosophical positions related to model validation, and in a broader scope credibility assessment, is provided by Kleindorfer et al. (1998). They note that “The validation problem in simulation is an explicit recognition that simulation models are like miniature scientific theories... As such, the warrant we give for these models can be discussed in the same terms that we use in scientific theorizing in general.” The philosophical positions treated in (Kleindorfer et al., 1998) are placed in one of two main branches: foundationalism (related to objectivism and justificationism) and anti-
foundationalism (related to relativism, conventionalism and anti-justificationism). A true foundationalist believes that validation is an absolute, i.e. every detail of a theory or model shall be validated using theory-free direct experiences (empiricism) or self-evident ideas from one’s own mind (rationalism). In other words, the model and the model developer are fully separable, the model is either valid or invalid, and no human judgment or interpretation may be involved in the process of validation. Problematic questions for the foundationalist are: Can one really find a theory-free empirical foundation? and Which ideas are actually to be considered as self-evident? As pointed out by Kleindorfer et al. (1998), even if a theory-free empirical foundation were available, the problem of induction still remains, i.e. the fundamental difficulty of justifying generalization based on a set of specific observations. As opposed to foundationalism, the anti-foundationalist positions involve judgment and decision making in one way or another. As an example, it is noted that “...an extreme relativist believes that the model and the model builder are inseparable. As such, all models are equally valid or invalid and model validity is a matter of opinion.”

In many practical situations, a traditional way of validating simulation models is to compare model results with measurement data. The measurement data may be obtained from the real system of interest, from a test rig, or from lower level bench testing of equipment used in the system of interest. Validation using measurement data is related to validation techniques termed historical data validation and predictive validation in (Sargent, 2013). In practical situations, it is impossible to collect measurement data for all system operating points. If it were possible, a model of the system would not be needed. In this sense, a first complicating factor is that there is always a lack of measurement data for validation purposes. The extreme case is the common situation of developing a model of a system which does not yet exist. A second complicating factor is that measurement data from the system of interest is not the same thing as the true system characteristics, i.e. measurement data is always uncertain. A third, and probably the most severe, complicating factor is the “unk-unks” – the “unknown unknowns”, here explained using a statement of Donald Rumsfeld, made while serving as the US Secretary of Defense in 2002:

“There are known knowns; there are things we know that we know. There are known unknowns; that is to say there are things that we now know we don’t know. But there are also unknown unknowns – there are things we do not know we don’t know.”

Since validation using measurement data goes from the specific (specific sets of measurement data from the system of interest) to the general (conclusions on the validity of a model), it can be seen as an inductive process. A fourth complicating factor is thus the problem of induction, which is related to both the “unk-unks” and the lack of data for validation purposes.

As discussed above, there are several methods to support validation using measurement data in the assessment of M&S credibility, e.g. SA, UQ or face validation. In this sense,
model validation may include methods related to both objectivism and relativism. That is, the “either/or” debate (also known as “Cartesian Anxiety”) about which is the “correct” philosophical position for model validation is unproductive (Kleindorfer et al., 1998).

Due to the complicating factors mentioned, a model cannot be fully validated, but rather validated to a specific degree within a specific range of operation, with the perspective of a specific intended use. Since model validation is closely related to the problem of induction, a validation statement cannot be a guarantee or a formal proof of correctness. The main output of model validation is information on which parts of the intended use the model is likely (and unlikely) to cover. Thus, using M&S for decision making necessarily includes a leap of faith. Most likely, this is typically not what decision makers would like to hear after spending money and resources on assessing the credibility of an M&S effort. However, when looking at the alternatives, M&S accompanied by structured and transparent methods for credibility assessment is not that bad after all.

It should be noted that providing guarantees or formal proofs of correctness is difficult also when it comes to physical testing. To end to this philosophical discussion, a quote from Kleindorfer et al. (1998) is used, proposing a holistic framework for model validation which does not preclude any philosophical position: “Rather, there are situations where appeals to each of these positions would be called for. Thus, the purpose should be to lend just enough structure to provide stability and lend meaning to questions of validation, yet not so much as to diminish the importance of individual freedom and ethical behavior in model validation”.
Modeling of Aircraft Vehicle Systems

The experimental frame of reference in the conducted research is *aircraft vehicle systems*, i.e. systems found in more or less any conventional aircraft, enabling fundamental capabilities necessary for aircraft operation. Examples include electrical and lighting systems, Environmental Control Systems (ECS), landing gear, fuel systems, and hydraulic systems. For fighter aircraft, emergency escape systems are also commonly included in this group. Figure 3-1 gives an overview of Saab Gripen’s vehicle systems.

Figure 3-1: An overview of Saab Gripen’s vehicle systems.
In modern fighter aircraft, most vehicle systems may be looked upon as fairly complex, including a vast number of hardware components as well as extensive software for control, functional monitoring, redundancy management, and Built-In Test (BIT). Typically, vehicle systems are tightly integrated into the aircraft, and due to the complexity of each subsystem and the high level of interconnection between subsystems, there are significant challenges in engineering design at both subsystem and system level. Significant use of M&S is thus necessary to understand and predict system behavior and thereby achieve a sufficiently good design in a reasonable amount of time.

Compared for example to CFD or CSM, which are two domains that have come far regarding VV&UQ methodology, system simulation is a very diverse field and the application of UQ appears to be relatively uncommon. To provide the necessary system simulation context to the appended papers, the remaining part of this chapter describes concepts typically used in M&S of aircraft vehicle systems. The first section below provides a short mathematical introduction to ODEs and DAEs. This is followed by an overview of techniques commonly used in the development of system simulation models. Most of the implementation efforts of the presented work are carried out in the equation-based object-oriented language Modelica. In addition, the industrial application examples used in this research include individual Modelica models as well as closed-loop models with sub-models developed using both Modelica and Simulink. Therefore, an introduction to Modelica is provided along with a short description of the typical layout of closed-loop models. The final section of this chapter presents the industrial application examples, ranging from models representing the physical part of a single aircraft subsystem, to closed-loop models and further on to large-scale aircraft system simulators.

### 3.1 ODE and DAE Fundamentals

As aircraft vehicle system models are typically dynamic in nature and often described by ODEs or DAEs, a brief introduction to the mathematics involved may be in order. The following equations are taken from (Ascher and Petzold, 1998). An explicit ODE is defined by

\[ y' = f(t, y) \]  

(3.1)

where \( y = y(t) \) represents the system characteristics and \( f \) is in general a nonlinear function of \( t \) and \( y \). An ODE may be characterized as either an Initial Value Problem (IVP)

\[ y' = f(t, y), \quad 0 \leq t \leq b \]
\[ y(0) = c \]  

(3.2)
where $c$ is a given initial condition, or a Boundary Value Problem (BVP)

$$y' = f(t, y)$$
$$g(y(0), y(b)) = 0$$ \hspace{1cm} (3.3)

A more general form of equation (3.1) is the implicit ODE given below.

$$F(t, y, y') = 0$$ \hspace{1cm} (3.4)

If $\partial F / \partial y'$ is nonsingular it is possible to solve (3.4) for $y'$, obtaining the explicit form (3.1). However, if $\partial F / \partial y'$ is singular, this is not possible and solution $y$ has to satisfy certain algebraic constraints. In this case, the problem is referred to as a DAE. A special case of DAE is that of an ODE with constraints or semi-explicit DAE, depending on additional algebraic variables $z(t)$, and forced to satisfy the algebraic constraints given by $g$:

$$x' = f(t, x, z)$$
$$0 = g(t, x, z)$$ \hspace{1cm} (3.5)

Defining $y = \begin{pmatrix} x \\ z \end{pmatrix}$ it is possible to rearrange (3.5) into the implicit form (3.4). Note that the class of DAEs includes all ODEs. An indication of how difficult it is to simulate a DAE is given by its index, which is the number of differentiations required to obtain an explicit ODE. Solving high-index problems is in general more complicated than solving low-index problems.

### 3.2 Methods, Languages, and Tools

At Saab Aeronautics, development of aircraft vehicle systems has been supported by M&S techniques since the late 1960s (Steinkellner et al., 2010). A challenge, as central today as in the 1960s, is to find a suitable level of complexity in the modeling of physical systems. The word complexity here refers to the level of detail used in the modeling of physics in model components, as well as the structure of model components. Note that component here refers to a model of a single piece of equipment and a (sub-)model includes several components. In the development of a component for representation of a specific physical phenomenon at a specific detail level, several structural alternatives are typically available. Four examples of structural tradeoffs are:
1) **Generality between domains:** Should one single component be used for modeling in several domains, for example using different types of medium (single-phase/multiphase, incompressible/compressible, single-substance/multiple-substance)? For example, when developing a pipe component for use in a liquid cooling model (single-phase, incompressible cooling liquid), should the pipe component be general enough to also handle air cooling applications (compressible humid air)?

2) **Generality inside domains:** Should one single component be able to handle several physical phenomena or should a set of tailor-made components be used? For example, should one single pipe component handle both laminar and turbulent flow, heat exchange with ambient due to specified ambient temperature, heat exchange with ambient due to specified heat flow, etc.

3) **Level of inheritance:** Should commonalities be broken down as far as possible to achieve generic partial models for extension and reuse or should the full definition of a component be gathered in a single container?

4) **Graphical vs. textual modeling:** When possible, should components be described by graphical “drag-and-drop” modeling, by textual code, or a mixture of these two techniques?

The chosen structure affects vital characteristics such as the user’s ability to i) use the component library for building models with a specific detail level, ii) understand what is happening inside components, iii) maintain, modify, and reuse components, and iv) verify and validate components and models.

In addition to the above tradeoffs, different modeling techniques provide different abilities to model a given system. A classification of modeling techniques is given in the figure below, which is an extended version of a figure originally presented by Krus (2008).

---

**Figure 3-2: Classification of modeling approaches.**
The two main branches are signal-flow modeling and power-port modeling. In signal-flow modeling, the information flow at each connection point is unidirectional, whereas the information flow in a power-port connection is bidirectional. The industrial application examples presented in section 3.4 include both signal-flow models and lumped parameter power-port models using signal-based and equation-based connectors.

### 3.2.1 Signal-Flow Modeling

A common tool for signal-flow modeling is Simulink by MathWorks. In signal-flow modeling, the causality has to be defined, i.e. which signals are inputs and which signals are outputs. Depending on what is known and what is to be calculated, the signal-flow approach may result in several alternative model implementations for one single problem. For example, consider a simple incompressible problem with two pipes connected in series. For simplicity, only secondary pressure loss is considered, i.e. each pipe is treated as an orifice.

![Figure 3-3: Left: One pipe treated as a simple orifice. Right: Simplified problem of two pipes in series.](image)

With the simplifications and assumptions given, the pressure loss over one pipe is defined by the following equation (Miller, 2009).

$$\Delta p = \xi \frac{\dot{m}^2}{2 \rho A^2}$$  \hspace{1cm} (3.6)

Here, $\Delta p$ is pressure drop, $\dot{m}$ mass flow, $\rho$ density, $A$ cross-sectional area, and $\xi$ the pressure drop coefficient. With indexes 1 and 2 denoting the design inlet and outlet respectively, and assuming constant density $\rho$, we may introduce a constant $C$ to obtain the following equation.

$$p_1 - p_2 = C \dot{m}^2$$  \hspace{1cm} (3.7)

To simplify even further, we only consider $\dot{m} > 0$, i.e. mass flow from design inlet to design outlet. Depending on the given causality of the problem, the above equation implies the following three alternatives to describe the pressure-flow characteristics of a pipe.

---

3 [http://www.mathworks.se/products/simulink](http://www.mathworks.se/products/simulink)
\[ p_1 = p_2 + C\dot{m}^2 \]  
\[ p_2 = p_1 - C\dot{m}^2 \]  
\[ \dot{m} = \frac{|p_1 - p_2|}{C} \]

The figure below shows three alternative signal-flow models of the system consisting of two pipes connected in series. Indexes A and B denote system inlet and outlet boundary conditions in terms of pressure or mass flow, and on pipe component level equations (3.8) to (3.10) are used.

Figure 3-4: Top: Boundary conditions \( p_A \) and \( p_B \) known, \( \dot{m} \) wanted. Pipe 1 and Pipe 2 using eq. (3.10) and (3.8) respectively. Middle: Boundary conditions \( \dot{m} \) and \( p_B \) known, \( p_A \) wanted. Pipe 1 and Pipe 2 both using eq. (3.8). Bottom: Boundary conditions \( \dot{m} \) and \( p_A \) known, \( p_B \) wanted. Pipe 1 and Pipe 2 both using eq. (3.9).
3.2.2 Power-Port Modeling

Bidirectional information flow facilitates *component-based* modeling. If component equations and medium equations are fully decoupled in such a way that one single component can be used with several types of mediums (single-phase/multi-phase, incompressible/compressible, single-substance/multiple-substance), this may be termed *device-oriented* modeling (Franke et al., 2009). Since component-based modeling theoretically enables a one-to-one hierarchical mapping of system topology to model topology, power-port modeling is well suited for modeling physical systems. Nonetheless, it should be noted that signal-flow modeling is often used for modeling of physical systems, and may well be a suitable alternative for a broad field of applications. However, for modeling more complex physical systems, power-port modeling often results in less cluttered models with a topology more similar to the real system of interest. Figure 3-5 shows a power-port model of the above system with two pipes connected in series. Note that the same pipe component models are used, independently of the type of boundary conditions given in terms of $\dot{m}$ or $p$.

![Figure 3-5: Modelica power-port model of a system consisting of two pipes in series.](image)

A widely used power-port concept is lumped parameter modeling with calculation of state derivatives in the model components, in combination with a centralized solver for time integration. That is, all model equations are collected in one system of ODEs or DAEs and solved by the centralized solver. As shown in Figure 3-2, the centralized solver approach can be realized using two different concepts of defining component connectors. Here, a connector is the connection point of a component where information on physical quantities is exchanged with the neighboring component. As the industrial application examples introduced later belong to this group, this approach is described in somewhat greater detail.

**Centralized Solver and Signal-Based Connectors**

In a signal-based connector, the causality is predefined. As an example from the fluid system domain, a capacitive component (e.g. a volume) may take mass flow as input and calculate pressure as output while a resistive component (e.g. an orifice) takes pressure as input and calculates mass flow as output. Typically, an ODE system is then obtained by altering components of capacitive and resistive type, as visualized in Figure 3-6. When modeling incompressible fluid systems, this technique implies that a virtual compressibility is added in each capacitive component.
A detail which may be argued is whether the approach using signal-based connectors really should be regarded as power-port or not. However, it facilitates component-based modeling, and from a user perspective it looks and behaves like power-port modeling. One tool that uses this concept is the FORTRAN-based M&S tool Easy5\(^4\) originating from Boeing, now provided by MSC Software. At Saab Aeronautics, there are good experiences from using this technique also with Modelica, for modeling thermal-fluid systems.

It should be noted that with some implementation effort, this technique (or something close to it) may also be implemented in signal-flow based tools like Simulink. This is shown in Figure 3-7 below, where the pipes are modeled as a capacitive component (here referred to as a node) in combination with a resistive component (in the figure named flow element). Due to Simulink’s current limitation regarding positioning of subsystem ports, i.e. inputs on the left and outputs on the right, this type of modeling results in a great many crossing feedback loops.

Centralized Solver and Equation-Based Connectors

In contrast to the signal-based connector, the causality of the information exchange in an equation-based connector is not predefined. Rather, the concept of equation-based connectors involves two kinds of variables, defined as either flow or effort variables (also referred to as through or across variables respectively). At each connection point, effort variables are set equal and flow variables are summed to zero. In bond graph modeling, the product of flow and effort variables is normally power in [W] (Paynter, 1961). In the

\(^4\) http://www.mscsoftware.com/product/easy5
figure below, this is not the case since the effort variable is pressure given in [Pa] and the flow variable is mass flow in [kg/s] (and not volume flow in [m³/s]).

![Equation-based connector](image)

Figure 3-8: Equation-based connector. Note that the sign convention of the flow variables has to be clearly defined. A common approach is to define mass flow into a port as positive.

The equation-based connector approach is used in the Simscape⁵ library of Simulink, and is also typically used in Modelica models. See (Borutzky, 2002) for a detailed discussion on the relation between bond graph modeling and equation-based object-oriented modeling using Modelica.

**Distributed Modeling**

An alternative to the centralized solver approach is distributed modeling based on bilateral delay lines, also known as Transmission Line Modeling (TLM), see for example (Krus, 2005). The basic concept of TLM is to separate components by introducing a physically motivated time delay, allowing each component to solve its own equations independently of the rest of the system. To clarify, a capacitive component such as a volume in a fluid system is modeled as a transmission line element for which the physical propagation time corresponds to one time step. The time delay is related to the length of the fluid path in the component and the speed of sound in the fluid used. In this way a time delay is introduced between capacitive and resistive components. Compared to the centralized solver approach, this facilitates a higher degree of numerical isolation of components which, may serve as an enabler for parallelization of model equations for simulation on multi-core platforms (Braun, 2015), (Sjölund et al., 2010). TLM is used in the M&S tool HOPSAN⁶ developed at Linköping University.

3.2.3 Modelica

Modelica is an equation-based object-oriented language developed primarily for modeling of heterogeneous physical systems. The language is developed by the non-profit, non-governmental international organization Modelica Association⁷. The object-oriented approach facilitates reuse and variability of model components. The equation-based

---

⁵ [http://www.mathworks.se/products/simscape](http://www.mathworks.se/products/simscape)
⁷ [https://www.modelica.org](https://www.modelica.org)
approach enables declarative programming where the user can define model equations and still leave the causality open. Thus, the user does not have to consider the order of calculation. This is in contrast to conventional imperative programming languages like FORTRAN, C, C++, or Ada, where it is up to the user to define both the problem and how it should be solved. However, Modelica provides the possibility to include imperative code in one or more algorithmic sections of a component. As state equations, “algorithmic equations”, and algorithms are allowed, a Modelica model may take the form of an ODE or a DAE.

As indicated above, by the word heterogeneous, Modelica is well suited for multi-domain modeling, i.e. to combine components and sub-models from different domains, such as the hydraulic, thermodynamic, mechanical, or electrical domains. In addition to physical modeling, Modelica also provides functionality for block diagrams and state machines (Modelica Association, 2012).

The most fundamental building block when developing model components in Modelica is the definition of component interaction points, known as connectors or ports. To give some examples of what a connector might look like, Modelica code of a signal-based and an equation-based connector respectively follows (as discussed above and shown graphically in Figure 3-6 and Figure 3-8). The signal-based connector approach requires one connector type for resistive components and one for capacitive components.

```plaintext
connector pmh_resistive
  "Connector on pipe-type elements for dry air or single substance liquid"
  input SIm.Pa p "pressure";
  output SIm.kgps m "mass flow";
  input SIm.Jpkg h "spec enthalpy into pipe-type modules";
  output SIm.Jpkg hn "spec enthalpy into node/volume";
  input SIm.m X[3] "position x,y,z, from node or accumulator";
end pmh_resistive;

connector pmh_capacitive
  "Connector on node/volume for dry air or single substance liquid"
  output SIm.Pa p "pressure";
  input SIm.kgps m "mass flow";
  output SIm.Jpkg h "spec enthalpy into pipe-type modules";
  input SIm.Jpkg hn "spec enthalpy into node/volume";
  output SIm.m X[3] "position x,y,z, always given in node or accumulator";
end pmh_capacitive;
```

The equation-based connector approach only requires one connector type. In this connector, the effort variable is pressure and the flow variable is mass flow. A Modelica stream variable is used to propagate information on outgoing specific enthalpy. The concept of stream variables is a Modelica-specific solution to describe the transport of specific quantities (Franke et al., 2009). In this connector, the port position in (x,y,z) is handled as an effort variable, i.e. one component in each connection point defines the
position, which is then automatically propagated to the other component(s) in the connection point.

```modelica
connector Port
  "Interface for 1-dimensional fluid flow (incompr., single-phase, one substance)"

  LAB.SI.Pa p "Pressure in the connection point";
  flow LAB.SI.kbps m "Mass flow rate from the connection point into the component";
  stream LAB.SI.Jpkg h_out "Spec enthalpy close to the connection point if m < 0";
  LAB.SI.m pos[3] "Port position (x,y,z)";
end Port;
```

Modelica is a modeling language used to describe models and to structure models and component libraries into packages. To graphically browse, edit, and simulate a model, a Modelica simulation environment is required. Prior to simulating a Modelica model, a number of steps which are similar in all Modelica simulation environments are necessary. However, the qualities in performing each step may differ between tools. The first step in the transformation of Modelica source code into executable simulation code is to parse the Modelica code into an abstract syntax tree. The output of this step is basically a flat set of equations, functions, variables, etc., and the object-oriented structure of the Modelica source code is dissolved. In brief, the next steps are to sort equations, perform index reduction, and symbolically simplify the problem as far as possible. After this, C-code is generated and at compile-time linked with a numerical solver to obtain the executable simulation code. The following figure shows the main steps and the output of each step.

![Figure 3-9: Steps for transformation of a Modelica model into executable simulation code, adopted from (Fritzson, 2014), shown by kind permission.](image-url)
3.3 Closed-loop Simulation

In the process of finding the most suitable modeling language and acquiring an M&S tool, there are several aspects to consider – from technical to personnel related, as well as financial and business strategies. The same applies when developing or acquiring a component library.

At Saab Aeronautics, M&S of complex physical systems using the Modelica-based tool Dymola\(^8\) is put into practice with purchased as well as in-house developed component libraries. In addition, open-source alternatives such as OpenModelica\(^9\) are becoming increasingly interesting for industrial use. For modeling control systems in aircraft vehicle system applications, the main tool currently used at Saab Aeronautics is Simulink.

When a strategy of using the best domain-specific languages and tools available is applied, there is an intrinsic need to exchange and connect models from different tools. One technique for obtaining closed-loop models, including sub-models from for example Simulink and Dymola, is hosted simulation (Steinkellner et al., 2008). Using hosted simulation, a Dymola environment closed-loop model is obtained by integrating code generated from Simulink. A corresponding Simulink environment closed-loop model is obtained by integrating code generated from Dymola. Which environment to use depends on the M&S task to be performed and which tool the engineer is most comfortable using. In Figure 3-10, a typical closed-loop model of an aircraft vehicle system is sketched.

![Figure 3-10: Typical layout of a closed-loop model of a specific aircraft vehicle system. System S/W and ECU H/W denote system-specific software and hardware placed in an Electronic Control Unit (ECU), and are typically modeled in Simulink. System H/W is a model of the physical system, typically developed in Modelica. BC denotes boundary conditions in terms of for example flight case and climate profile, and the gray-dashed arrows indicate communication with other systems.](http://www.3ds.com/products-services/catia/products/dymola)

\(^8\) [http://www.3ds.com/products-services/catia/products/dymola](http://www.3ds.com/products-services/catia/products/dymola)

\(^9\) [https://openmodelica.org/](https://openmodelica.org/)
3.4 Industrial Application Examples

The work presented in this thesis involves three main types of industrial applications which are used for methodology development and evaluation; 1) Environmental Control Systems (ECS), 2) liquid cooling systems, and 3) large-scale aircraft system simulators.

The ECS is used in paper [I] to guide the reader through the workflow of developing, validating, and exporting a model to a model storage, for later (re)use in simulator applications, and in paper [III] for assessment of the work effort required for UQ of a detailed industry grade simulation model. The liquid cooling system is used in paper [IV] for development and evaluation of an approximate uncertainty aggregation method. In addition, a large simulation model consisting of two ECSs and one liquid cooling system are used in paper [V] to develop a framework for early and approximate UQ. Two large-scale aircraft system simulators are used for evaluation of the credibility assessment method described in paper [VI], one real-time Hardware-In-the-Loop simulator and one soft simulator without real-time requirements. The industrial application examples are described in somewhat more detail in the following sections.

3.4.1 Environmental Control Systems

Air supply, cabin pressurization, and avionics cooling are basic requirements for most aircraft, usually facilitated by an ECS. In the Gripen fighter, the main purpose of the ECS is to provide sufficient cooling of the avionics equipment, tempering and pressurization of the cabin, pressurization of the fuel and anti-g systems, and to provide conditioned air to the On-Board Oxygen Generating System (OBOGS), which in turn provides breathing air to the pilots. Briefly, this is achieved by means of a bootstrap configuration using engine bleed air which is decreased in pressure and temperature and dried prior to distribution. The main H/W components in the ECS are heat exchangers, compressor, turbine, water separator, pipes, and control valves. The ECS S/W, which is physically located in the General systems Electronic Control Unit (GECU), controls and monitors pressure, temperature, and flow levels in various parts of the system. The layout of the ECS is show in Figure 3-11 below.
Aligned with the real system layout, the closed-loop model of the ECS consists of three major models, viz. the ECS H/W model, the ECS S/W model, and the GECU H/W model. The ECS H/W model has been developed in Modelica using Dymola and the two other models have been developed in Simulink. Closed-loop models are obtained using hosted simulation (Steinkellner et al., 2008). To make a link with the previously discussed general template for an aircraft vehicle systems closed-loop model, the ECS H/W model is located in the rightmost box in Figure 3-10. Figure 3-12 shows a graphical overview of the ECS H/W model.
The ECS H/W model has several variants, e.g. one simple and one detailed variant. The model layout is hierarchical and the Modelica construction *replaceable* is utilized to obtain different variants applicable for model-time binding. Additional variant handling is performed by parameter selection at load-time and run-time. See Andersson (2012) for an introduction to binding concepts and binding time. The figure above shows the detailed physical model with its sub-models. This view is actually one step down from the ECS H/W model top level, in which either detailed or simplified is selected.

### 3.4.2 Liquid Cooling Systems

Liquid cooling systems are typically used for equipment with more demanding cooling requirements, such as tactical systems like radar and Electronic Warfare (EW). An application example used in this research is a liquid cooling system in the Saab Gripen Demonstrator Aircraft. The main purpose of the system is to provide sufficient cooling of the radar antenna, in this case an Active Electronically Scanned Array (AESA). More specifically, the liquid cooling system must meet a set of requirements concerning radar inlet and outlet temperature, cooling liquid mass flow, and pressure levels. The cooling liquid is in turn cooled in a heat exchanger, with air supplied by an ECS. The main
components of the system are pump, accumulator, liquid-to-air heat exchanger, piping, and a subsystem of heat loads including the radar antenna and related electronic equipment.

To support system design and specification, a model of the liquid cooling system has been developed in Modelica using Dymola. As part of an evaluation of different concepts for modeling in Modelica, the model has been implemented in three different variants using different component libraries: 1) Modelica.Fluid, which uses equation-based connectors and the Modelica.Media library and is shipped together with the Modelica Standard Library, 2) the Saab Aeronautics in-house library Modelica Fluid Light (MFL), which uses signal-based connectors and a simpler in-house media library, and 3) a prototype library that uses equation-based connectors and the same in-house media library as MFL. The intention of the third component library is to pick up the advantages of both Modelica.Fluid and MFL. The model shown in the figure below is based on components from the prototype library.

![Diagram of liquid cooling model](image)

Figure 3-13: Modelica implementation of the liquid cooling model.

The model used in paper [IV] is the MFL implementation. From a system simulation perspective, the model may appear to be fairly simple. Nonetheless, it is a component-based model of a physical system consisting of a number of components and one submodel. This 1-D dynamic simulation model is used to predict pressure, mass flow, and temperature levels at different points in the system. The components include equations describing pressure variations due to g-loads and fluid thermal expansion, internal heat exchange between equipment and fluid, external heat exchange between equipment and surrounding equipment bays, temperature dynamics in equipment and fluid, as well as fluid dynamics due to transport delays in the piping arrangement. The model contains approximately 200 equations, 100 parameters, and 50 states. The connector interface used
in the model includes information about pressure, mass flow, and specific enthalpy \((p, \dot{m}, h)\), and is shown in detail in the code example in section 3.2.3.

### 3.4.3 Large-Scale Aircraft System Simulators

At Saab Aeronautics, behavioral simulation models of aircraft vehicle systems are typically developed in Modelica, and models of control systems are developed in Simulink. Software architecture and event-driven software are typically developed using tools incorporating the Unified Modeling Language (UML) (OMG, 2013). Closed-loop models representing specific subsystems, such as the ECS described above, are obtained using code generation and hosted simulation (Steinkellner et al., 2008). These models are then used on desktop platforms for a wide range of applications such as system design, performance estimations, control design, fault simulation, and test planning. Additionally, models of each subsystem are exported for use in aircraft system simulators. The simulators are then used for development, verification, and training. Different simulator configurations are used depending on the given task. For example, in the Mission Trainer (used for pilot training) a detailed behavior of the ECS is typically not needed while in the Virtual Maintenance Trainer (used for technician training) a detailed behaviour may be essential. As indicated above, a simulator typically includes several individual simulation models of aircraft subsystems, often developed in different tools using different modeling techniques (Andersson, 2012).

In the current development of the Saab Gripen fighter aircraft, a number of different simulation environments and test rigs are used. Of these, the following two simulator applications for simulation at aircraft system level have been used for evaluation of the credibility assessment tool described in paper [VI]: 1) MySim, a purely soft simulator without real-time requirements, used on desktops for early development testing, and 2) Total system rig (T-rig), a real-time HIL simulator used for development, testing and formal verification of system requirements and flight safety.

To decrease the number of flight hours needed in test aircraft during system development and verification, and to optimize the usage of narrow resources like the T-rig, intensive use of desktop simulation environments at both subsystem level and aircraft level is essential. Figure 3-14 and Figure 3-15 show the T-rig cockpit environment and equipment rooms respectively. Further information related to the development of the simulator product line at Saab Aeronautics is provided in (Andersson, 2012).
Figure 3-14: T-rig cockpit environment room.

Figure 3-15: Picture taken of the equipment room during the development of the T-rig, showing some of the on-board equipment and interface units.
4

Contributions

Based on challenges seen in today’s aeronautical model-based system development, this work addresses methods for M&S credibility assessment in general and early model validation in particular. A workflow has been developed for development, V&V, and export of multipurpose simulation models, as well as methods for approximate UQ. As a complement to the methods intended for simulation models of individual subsystems, this work also proposes a method and an associated tool for credibility assessment of large-scale aircraft system simulators. In addition to Table 1-1 in section 1.6, which provides a very brief overview of the appended papers and how they are related to each other, the contributions of each paper are summarized below. For detailed information, see the appended papers [I] to [VI].

4.1 Paper [I]: Methodology for Development and Validation of Multipurpose Simulation Models

Until recently, a fairly common situation has been that one or a few individuals are responsible for the development, V&V, and usage of a simulation model. Current development is moving towards more all-embracing multipurpose models intended to support not a few but several different tasks. The models may be used in different phases of the product lifecycle and by different individuals from different organizational domains. This situation puts new requirements on transparency in the methodology and documentation of V&V activities. As for system development in general, model development involves a large number of activities which all require a number of balanced and informed decisions. To facilitate a common view of model development and related activities and to provide practical guidance to model developers, a methodology has been developed for development, V&V, and export of multipurpose simulation models. The
methodology, which is documented in the form of a handbook, is originally a wiki page at Saab Aeronautics but a public variant is also available (Andersson and Carlsson, 2012).

The proposed methodology enables reuse of models in different applications with different purposes and the scope of the handbook is simulation models representing physical environment, physical aircraft subsystems, avionics equipment, and electronic hardware. The word \textit{export} here refers to the process of exporting a model for integration in a simulator application and \textit{multipurpose} refers to the fact that models are not only used by engineers specialized in M&S but also by several other user groups, and integrated in several kinds of applications with varying intended uses. Some examples are:

- Model of a single subsystem (S/W or H/W) used in a specific M&S tool
- Closed-loop model (S/W and H/W) used in a specific M&S tool
- Desktop simulator that includes models of several subsystems
- Large-scale simulator that includes models representing a complete aircraft
- Simulator for pilot training
- Simulator for technician training
- HIL simulator, e.g. test-rig with some subsystems replaced with simulation models

The workflow starts with the definition of intended use of a simulation model and ends with the issuing of a simulation model status declaration. The top level view of the methodology is shown below. In paper [I], the ECS model from section 3.4.1 is used as an industrial application example to guide the reader through the workflow.

![Figure 4-1: Workflow in Saab Aeronautics Handbook for Development of Simulation Models.](image)
The arrow-shaped blocks correspond to activities and the stars are output items in the form of 1) a simulation model running in a specific M&S tool, 2) an interface description, 3) test cases, 4) generated model code, and 5) simulator integration layers. The symbols named “RS” and “SD” represent documentation such as Requirement Specification and Status Declaration. The figure above provides a chronological view of the workflow; it must, however, be stressed that the duration of each activity may vary significantly depending on the character and intended use of the simulation model of interest. Another important aspect is that activities normally overlap and that there are iterative loops between activities, i.e. the workflow is not sequential.

Developing a useful workflow for simulation model development has proven to be a challenging task. To suite the wide range of models included in the scope, a generalized methodology is required. However, to be a useful engineering aid in model development, a sufficient level of detail has to be provided. Nonetheless, the real challenge is not to develop and maintain a description of a methodology, but to get people to know about it, to train engineers to use it and apply it in their daily work. The methodology has been developed by a small interdisciplinary team, with experience from M&S of vehicle systems as well as simulator development. This increases the probability of defining a relevant methodology. Special care has been taken to ensure usability, mainly by means of 1) a user friendly format, easy to overview and update, 2) keeping the amount of text down, and 3) providing relevant examples, templates, and checklists.

The methodology is implemented and has been in broad use at Saab Aeronautics for a few years now. The following list contains experiences and technical recommendations gained during the development and use of the methodology:

- It is difficult to predict all future uses of a simulation model. A model may be developed with a specified intended use but at a later stage when the model is in operation, new areas of use may be discovered. It is nonetheless important to obtain as complete a picture of the intended use as possible prior to model development. This certainly applies to multipurpose simulation models to be integrated in several applications, since requirements from different stakeholders may conflict and some conflicts may need to be resolved at an early design stage.

- The amount of useful and value-adding documentation depends on the application, but a strong recommendation is to always provide a status declaration for each model in the model store.

- Ideally, all model functionality shall be placed in the core model, i.e. the additional integration layers should only concern signal interface, with possible changes in signal names, data types, and, if necessary, unit transformations. If functionality is spread out in the different integration layers, it implies an increased risk of functional errors. It may also make it difficult to obtain an overview of the model.
functionality, which complicates model updates and increases the long-term maintenance costs.

- Validating a simulation model often demands a substantial amount of work. A model cannot be fully validated in all operating points, but validated to a specified degree in a certain range of operation. To make the validation of multipurpose simulation models feasible, an efficient method is to perform as much of the validation effort as possible in the environment best suited for this activity, and to use common test cases to ensure the consistency of applicable integration layers.

Finally, providing engineers with descriptions of a logical workflow supported by templates and checklists gives them a framework to act within and a feeling of better control of the model management activities.

4.2 Paper [II]: Study of Industrially Applied Methods for VV&UQ of Simulator Models

To find suitable VV&UQ methods already in use in industry, an exploratory case study has been carried out at Saab Aeronautics and the Swedish National Road and Transport Research Institute (VTI). The study also investigates whether the NASA Credibility Assessment Scale (CAS) (NASA, 2008) is useful in the context of aircraft system simulators and driving simulators. The main contribution is the increased insight into how V&V is conducted in two organizations with extensive experience in simulator development and usage. Both technical and organizational needs are pointed out. The study is based on a questionnaire and on interviews with engineers working with model development and validation, and the results are of both qualitative and quantitative nature.

The workflow from initiating the study to drawing conclusions is shown in Figure 4-2, which is inspired by Runeson and Höst (2009). The choice of population is carried out by self-selected sampling of models developed at Saab Aeronautics and VTI. The study contains a total number of 17 simulation models (14 models represented by 9 individual model developers at Saab Aeronautics, and 3 models represented by 3 individual model developers at VTI). For the analysis of open-ended questions, an adaptation of the Affinity Diagram Technique (Kasunic, 2005) has been applied. Basically, the method has been used as a tool for transforming qualitative textual answers into quantitative results, via grouping of answers followed by coding into nominal numbers.

The evaluation of the CAS is carried out by the use of semi-structured interviews. A description of the CAS was provided to the respondents prior to the interviews and detailed information about the CAS is provided in situ during the interviews. A subset of the 17 models has been included in the CAS evaluation; 7 models at Saab Aeronautics and one model at VTI.
The majority of the models describe physical systems (from domains such as thermal-fluid or mechanics) whereas a smaller number describe electronic hardware, such as the hardware part of an Electronic Control Unit (ECU). Included in the study are models of flight mechanics, ECS, hydraulics, fuel systems, braking systems, landing gears, ground vehicle dynamics (passenger cars), and ECUs.

In total, 31 V&V techniques are identified in the study, 27 of which can be approximately mapped to Balci’s list of techniques mentioned in section 2.2. That is, four additional techniques have been identified in the case study. An interesting method identified is the use of a credibility measure embedded in a model, i.e. an output signal describing one or more aspects of model credibility. This is an essential contribution which is generalized and further developed in paper [VI].

The results also include correlations between answers to different questions and – for cases considered interesting – the lack thereof. Data is presented in bubble charts showing the number of models located in each XY-combination. The circle area is proportional to this number, which is also printed in the center of each circle. As an example, an interesting result is that more than 50% of the models do not have an explicitly stated intended use.

Figure 4-2: Workflow of case study.
This is contrary to textbook modeling processes, but well in line with the authors’ experiences from practical model development. As can be seen from Figure 4-3, most model developers assess their own model to cover its intended use, but in many of these cases the intended use is not explicitly defined in the model documentation. This contradictory result can be partially explained by the fact that the assessment is carried out by the model developers with respect to their own understanding of the intended use. Nevertheless, the lack of well-defined intended uses is indeed problematic in a V&V context.

![Figure 4-3: Definition of intended use vs. coverage of intended use.](image)

Historically, model development has often been initiated with an implicitly understood intended use and as the model becomes available to users, the scope of and expectations from the model often increases with time. However, an explicitly stated intended use is desirable and the results from the case study show a weak trend of improvement.

To summarize the results from the case study, it is observed that roles and processes are applied to enable some degree of independence in V&V, especially regarding integration testing. However, none of the studied organizations have a dedicated group or department for V&V of simulation models. This implies a great responsibility on the part of individual model developers to decide how far to go with building confidence in a specific model in terms of both planning and execution. Thus, a transition towards more
independence in V&V is desirable, but what is a suitable degree of independence in a specific context is open to discussion.

Furthermore, the results show that SA is applied only on a few of the studied models, and UQ is not applied on any model included in the study. This implies a large gap between academia and this part of industry and indicates that simplified and approximate methods are needed for SA and especially UQ to be widely used in aircraft system simulation. To take a step towards a more continuous process for model V&V, a higher degree of automation is also needed. What should be automated is repetitive work with limited necessity for human judgment. Obstacles to a higher degree of automation are the lack of well-defined intended uses, validation metrics, and acceptance criteria.

Regarding the applicability of the CAS, it is considered to be a usable tool for achieving a uniform level of V&V for all models included in a simulator, although its implementation at the studied organizations requires careful tailoring and coordination.

4.3 Paper [III]: Enabling Uncertainty Quantification of Large Aircraft System Simulation Models

Even though UQ is commonly seen as a value-adding activity increasing M&S credibility, estimations of UQ workload and analysis of industrial applicability for use in aeronautical system development are hard to find. In this paper, the engineering work effort as well as the computational cost for carrying out a structured and comprehensive UQ of a large system simulation model have been estimated. The theoretical frame of reference for the estimation is the established six-step UQ framework proposed by Roy and Oberkampf (2011), and the estimation uses the ECS model from section 3.4.1 as an industrial application example of a large system simulation model. It should be noted that this study considers the physical part of the ECS closed-loop model, corresponding to the rightmost box in Figure 3-10.

The number and types of uncertainties in the ECS model are analyzed by model inspection, and the engineering work effort and computational cost related to each of the six steps are estimated in a collaborative manner. Two out of the six UQ steps were carried out in practice for the ECS model. For the remaining steps, more approximate, experience-based estimations are given. Fully representative estimates cannot be obtained without actually completing each step, which is beyond the scope of this paper. The estimates are however believed to give a reasonably good indication of the total work effort and computational cost required.

In the initial step of the estimation, parameters and input signals are identified and placed into one of the following categories A to F. The list includes the number of parameters or input signals representative of the ECS model.
A) Geometry parameters considered certain (e.g. pipe diameter): 92  
B) Geometry parameters considered uncertain (e.g. pipe surface roughness): 76  
C) Physical parameters considered certain (e.g. pressure loss coefficient of simple pipe bend): 31  
D) Physical parameters considered uncertain (e.g. participating thermal mass or pressure loss coefficient of complex geometry): 196  
E) Input signal considered certain (e.g. one-seater or two-seater aircraft): 7  
F) Input signal considered uncertain (e.g. engine bleed pressure): 4

Although some simplifications are introduced already during the estimation, the results in Table 4-1 indicate that a comprehensive but not excessive UQ of a large aircraft system simulation model may be extremely time-consuming.

<table>
<thead>
<tr>
<th>Step</th>
<th>Engineering work effort [hours]</th>
<th>Computational cost [hours]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Identify all sources of uncertainty</td>
<td>16</td>
<td>N/A</td>
</tr>
<tr>
<td>2) Characterize uncertainties</td>
<td>3396</td>
<td>N/A</td>
</tr>
<tr>
<td>3) Estimate uncertainty due to numerical approximations</td>
<td>4</td>
<td>N/A</td>
</tr>
<tr>
<td>4) Propagate input uncertainties</td>
<td>40</td>
<td>990</td>
</tr>
<tr>
<td>5) Estimate model form uncertainty</td>
<td>430</td>
<td>N/A</td>
</tr>
<tr>
<td>6) Determine total uncertainty</td>
<td>40</td>
<td>N/A</td>
</tr>
<tr>
<td>Total:</td>
<td>3926</td>
<td>990</td>
</tr>
</tbody>
</table>

Most certainly, there is an uncertainty in the estimation but the work effort is not believed to be overestimated. The total work effort is primarily dominated by the step characterize uncertainties, followed by estimate model form uncertainty. As a comparison, the estimated engineering work effort required for UQ is about twice the total time used for model development, verification, and steady-state validation of the ECS model.

The uncertainty propagation is demanding when it comes to computational resources. The time required can however be shortened by distributing a number of batch simulations over several cores. That is, the 990 hours of execution time for uncertainty propagation can be significantly reduced. However, it should be noted that the estimation considers only single-loop sampling, which implies a vast simplification. Nested sampling, as proposed in (Roy and Oberkampf, 2011), is far too demanding in terms of computational cost and is considered to be out of scope for the application studied.

To enable UQ of large system simulation models, approximations, compromises, and simplifications to ease the UQ workload are necessary. The most significant measures
Contributions

proposed concern reducing the number of parameters treated as uncertain, simplifying the
characterization of uncertainties, and simplifying the uncertainty propagation. The
proposed simplifications are discussed in more detail in papers [IV] and [V].

4.4 Paper [IV]: Definition and Implementation of a
Method for Uncertainty Aggregation in Component-Based System Simulation Models

To enable dimensionality reduction of a UQ problem and to thereby make UQ more
feasible for industry-grade system simulation models, the paper describes a pragmatic and
approximate method for uncertainty aggregation termed component output uncertainty.
Briefly described, each component in a model, e.g. a pump, a pipe, or a heat exchanger,
is typically validated against some available data. The result from such validation efforts
is information about component level uncertainty. The central idea is to integrate the
information obtained during component level validation directly into the component
equations and to utilize this information in model level UQ. That is, uncertainty is
expressed at component level rather than at parameter level. In this way the proposed
method enables one form of early model validation even though system level
measurement data are unavailable or very limited.

In the context of processes for system simulation model development and V&V, the
proposed uncertainty aggregation method adds two major steps. As seen in Figure 4-4,
data from component level validation are used as input to the step “Characterize and
Quantify Component Output Uncertainty”. The component output uncertainties are then
used in the generic step “Model Uncertainty Quantification”. The foundation of the process
shown below stems from the workflow described in paper [I].
To enable the integration of uncertainty information directly into the component equations, an extension of the original component is necessary. The uncertainties may be expressed in absolute terms or relative to some characteristic of the original component. With the aim of achieving an intuitive uncertainty description suitable for aircraft vehicle system models within the thermal-fluid domain, it has been chosen to define the component output uncertainties in terms of pressure and temperature. Figure 4-5 shows the graphical layout of an original component extended with an output uncertainty description component $UC$. 
Figure 4-5: A new uncertain component (red-dashed) based on an original component (in this case a pipe component) connected with an uncertainty description component (here named UC). The original component’s nominal pressure drop $\Delta p_{\text{component}}$ is shown, as well as the added uncertainty in the pressure drop $\Delta p_{\text{UC}}$.

The equations required to describe the component output uncertainty are placed in the extension UC, i.e. there are no internal modifications of the original component. The uncertainty description component has been implemented in the MFL library introduced in section 3.4.2. See paper [IV] for mathematical definitions and Modelica implementation aspects.

In addition to the definition and implementation of the uncertainty description component, the paper also provides a guiding example of how to make practical use of equation (2.7) for characterization and quantification of a component’s output uncertainty based on component-level bench test measurement data. Furthermore, the industrial applicability of the uncertainty aggregation method is demonstrated in an approximate UQ of the liquid cooling system simulation model from section 3.4.2. The analysis includes a basic comparison of a brute force Monte Carlo sampling versus a Latin Hypercube Sampling (LHS), both in combination with aggregated uncertainties. Figure 4-6 shows the PDF and CDF of the pressure at the heat load inlet port for Monte Carlo sampling with $1.5 \cdot 10^5$ samples and LHS with 250 intervals. It can be noted that the shape of the distributions are preserved fairly well even for this low number of LHS intervals.

Figure 4-6: Resulting PDF and CDF for the heat load inlet pressure using Monte Carlo sampling and LHS respectively.
In the industrial application example used, the proposed method reduces the number of uncertain parameters by more than 50%. To clarify, all original component parameters are kept with nominal values, and the model is extended with a new but significantly smaller set of uncertain parameters. In combination with the more intuitive characterization of these uncertain parameters, this implies a substantial improvement in the conditions for conducting a model level UQ.

For the industrial application example used, the system characteristics obtained with a computationally relatively cheap LHS scheme (250 intervals) are in fairly good agreement with the system characteristics obtained with a high number of Monte Carlo samples ($1.5 \cdot 10^5$ samples). Thus, the reduced number of uncertainties in combination with efficient sampling techniques actually makes approximate probabilistic UQ feasible for this type of application.

It should be noted that the selected approach implies a simplified lumped treatment of epistemic and aleatory uncertainty, which makes the UQ less precise. However, this is a deliberate choice motivated by the common lack of information and resources for detailed uncertainty characterization.

To summarize, for cases when comprehensive UQ is not feasible, the demonstrated approximate UQ using aggregated uncertainties is considered to be a pragmatic alternative suitable for industrial M&S practitioners.

4.5 Paper [V]: A Framework for Early and Approximate UQ of Large System Simulation Models

The paper addresses the challenging task of UQ of a large system simulation model that includes a set of connected sub-models representing different aircraft subsystems. That is, the proposed UQ framework is directed towards simulation models of significantly higher complexity than the single subsystem application example used in paper [IV]. In paper [V], the industrial application example is constituted of two ECS models, one liquid cooling system model, and an engine model. The challenge is addressed by exploring methods for i) simplifying the uncertainty characterization process and ii) reducing the dimensionality of the UQ problem. If appropriate approximations in these two areas can be found, this can mitigate the problems identified in paper [III], regarding both computational cost and engineering workload. The proposed framework uses the approximate uncertainty aggregation method described in paper [IV] as one of the means for dimensionality reduction. However, due to the size of the total simulation model considered in paper [V], further simplifications are proposed.

In order to get rid of the inner sampling loop discussed in section 2.3.3, the uncertainty characterization may be simplified by treating all uncertainties as epistemic or by assigning the epistemic uncertainties uniform PDFs. The approach selected in this paper is to
simplify as far as possible by characterizing all uncertainties as epistemic and apply an interval-based minimum/maximum analysis of the total simulation model. If it is possible to efficiently find the uncertainty settings which give the widest interval for the output of interest, one can drastically reduce the number of simulations needed of the total simulation model.

The total simulation model consists of several connected sub-models, and due to the high computational expense, thorough sampling of the total simulation model is not an option. However, the sub-models can be run much faster when simulated as standalones. To utilize this, a solution alternative investigated in the paper is to use linear response surfaces to approximate the output dependency of the input uncertainties for each sub-model. In this process, each sub-model is treated as a standalone. This implies that thorough sampling is carried out on individual sub-models only, and that the number of simulations of the total simulation model can be kept relatively low.

In the industrial application example used, the output of interest with respect to the total simulation models is the overall cooling performance, which can be analyzed both in steady-state and in transient conditions. Therefore, it should be investigated how each sub-model and its inherent uncertainties affect the overall cooling performance – which is also the objective of the sub-models’ response surfaces. That is, the response surfaces are used to find the settings of uncertainties in each specific sub-model expected to give the minimum and maximum overall cooling performance respectively. These uncertainty parameter settings are here referred to as the minimum/maximum output parameter setting. For example, regarding the liquid cooling system, a high cooling liquid temperature generally implies a low overall cooling performance, and vice versa. Therefore, the output of the response surface for the liquid cooling system simulation model is the uncertainty of the cooling liquid temperature. The minimum/maximum output parameter settings for each sub-model are then combined to obtain the resulting uncertainty in the overall cooling performance.

Treating all uncertainties as epistemic reduces the characterization workload, but the uncertainties still have to be assigned well-founded intervals. A reasonable approach is to find the most important uncertainties, which affect the results the most, and devote the most resources to characterize these uncertainties with the most accurate intervals possible. The less important uncertainties with respect to the output of interest can be roughly assigned with “big enough intervals” or where justified ignored. To reduce the dimensionality of the UQ problem, the following three means are considered: i) experience, ii) screening by approximate characterization combined with SA, and iii) uncertainty aggregation according to paper [IV].

Based on the comprehensive framework proposed by Roy and Oberkampf (2011), in combination with a number of methodological simplifications briefly described above, a framework for early and approximate UQ of large system simulation models (including several sub-models representing different subsystems) is proposed. The framework consists
of six steps and is as illustrated in Figure 4-7. A prerequisite for carrying out a study according to the proposed workflow is that a main output of interest is defined and that a qualitative analysis is made of how each sub-model affects this output.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
</table>
| 1    | Estimate numerical uncertainty  
Varying solvers and tolerances, analysis of mass and energy conservation in nodes and branches |
| 2    | Identify input and model form uncertainties  
SME and analysis of simulation model |
| 3    | Approximate characterization and screening SA  
SME and SA with truncation threshold |
| 4    | Detailed characterization of input and model form uncertainties  
Measurement data, high-fidelity models, literature, and SME |
| 5    | Find sub-model min/max output parameter settings  
Response surfaces of sub-models |
| 6    | Uncertainty propagation at model top-level  
Application of min/max output parameter settings to total model |

Figure 4-7: Proposed workflow of an approximate UQ for use in early system development phases, “SME” is an abbreviation of Subject Matter Expert. The number of uncertainties for consideration is reduced in three steps.

The proposed framework has been applied on the industrial application example for approximate UQ of cooling performance estimations. The stationary cooling performance is assessed in terms of endurance time in minutes for different operating points in the flight envelope, which is spanned by aircraft altitude and speed. In the conducted analysis, the number of evaluated operating points in the flight envelope was set to 59. Three simulations (minimum, nominal, and maximum stationary cooling performance) at each evaluated operating point give a total of 177 simulations.
Figure 4-8 (left) provides a stationary cooling performance plot in terms of nominal endurance time in minutes without any uncertainty information. Figure 4-8 (right) shows the same nominal data extended with the uncertainties resulting from the approximate UQ.

![Figure 4-8](image)

As can be seen from Figure 4-8 (right), the uncertainty of the endurance time is greatest in the limited regions. The reason for this is that in the other parts of the flight envelope shown in the figure, the cooling performance of the system is either i) so high that the control system manages to keep the temperature at the set point value regardless of the uncertainties, or ii) so low that the endurance time is zero minutes regardless of the uncertainties. In this aspect, the proposed interval-based epistemic UQ approach is beneficial, since a more comprehensive UQ of such operating points would not be value-adding.

Regarding uncertainty identification and characterization as well as response surface generation, it should be noted that this study is limited to the uncertainties in the liquid cooling system sub-model. Adding uncertainties from other sub-models would not increase the number of simulations needed of the total simulation model.

If a more accurate but still computationally relatively cheap UQ is required to explicitly handle significant uncertainties characterized in detail as aleatory, the proposed framework can be extended; the proposed UQ framework can be used as described to find minimum/maximum output parameter settings for the epistemic uncertainties, and combined with sampling of the aleatory uncertainties.

Finally, it should be noted that if strong nonlinearities exist, the combination of the minimum/maximum output parameter settings for the individual sub-models does not necessarily result in minimum and maximum response in the total simulation model. Such characteristics could possibly be detected by means of additional random samples at the model top level.
4.6 Paper [VI]: Development and Evaluation of a Concept for Credibility Assessment of Aircraft System Simulators

As a response to some of the needs identified in paper [II], a concept and a related tool for credibility assessment of simulator applications is developed and evaluated. The concept aims to support simulator users such as test engineers or system engineers in credibility assessment both prior to, during, and after a simulator run. In addition, the concept is seen as a complement to the traditional document-centric approach to simulator test worthiness declaration described in section 2.4.

Several disciplines were represented in the development and evaluation of the proposed concept; engineers specialized in M&S of aircraft vehicle systems, simulator software developers, test engineers, and methodologists specialized in flight testing, simulator testing, V&V, and configuration management. The main steps in the development and evaluation of the proposed concept are shown in Figure 4-9.
As seen in Figure 4-9, three groups of aircraft vehicle systems are included in the development and evaluation of the concept: 1) ECS with four simulation models, 2) Fuel System with two simulation models, and 3) Secondary Power System with one simulation model.

First, a number of static and dynamic credibility measures were defined. These were then evaluated and a subset of the credibility measures was chosen for implementation in the aircraft vehicle system simulation models. In parallel with this, a Graphical User Interface (GUI) for use during simulator testing was designed and implemented. The tool was tested in the Saab Gripen fighter aircraft system simulators MySim and T-rig, described in section 3.4.3. To enable evaluation in realistic scenarios, the testing was carried out in parallel with already planned simulator testing. To provide structure for the user evaluation, a questionnaire was devised made up of both preparatory questions and questions to be answered during simulator testing.

In addition to the evaluation carried out by simulator users, workshops were arranged to collect feedback from methodologists and decision-makers with influence on a possible large-scale implementation of the concept.

The central idea of the proposed concept is that each model in a simulator is extended with a meta-model describing different aspects of credibility. This credibility information is then presented to the simulator users directly during simulation. The meta-model includes a number of static credibility measures and a dynamic measure that may vary during simulation.

Based on quantitative rating by test engineers, methodologists, and decision makers, the following three static measures, with a number of sub-levels each, are considered relevant: i) System type (1..5), ii) Fidelity\(^{10}\) (1..3), and iii) Signal propagation (1..3). The first measure, System type, has very limited influence in the assessment of simulator credibility in practice; however, it is kept as it quickly gives the user a general understanding of the composition of a large-scale simulator. Fidelity is here treated as a measure of the level of detail used in the modeling process. Clearly visualized modeling detail levels facilitate an enhanced understanding of what can and cannot be expected from a simulator. This may also point out weaknesses in the simulator configuration, such as unconsidered connections between models of dramatically different detail level. For example, in some cases it makes no sense to let a simple model provide input to a detailed model, i.e. “garbage in, garbage out”.

The static measure Signal propagation describes to what extent signals are propagated through the model and is useful to highlight flaws in functional chains, in particular when using simulator testing for fault simulation to verify software functionality for fault detection, fault isolation, and pilot warning messages.

\(^{10}\) The term fidelity may be interpreted in various ways. With the aim of reducing confusion, after the submission of paper [VI] it was decided to change the name of this credibility measure to Detail level. However, for consistency with paper [VI], the name Fidelity is kept here.
The dynamic credibility measure is a generalization of a measure identified in paper [II]. The proposed dynamic measure is called \textit{Credibility\_Level} and has three sub-levels:

\textit{Credibility\_Level} (1..3):

1) No credibility: e.g. outside the operational limits of the model or unrealistic input data to the model.

2) Degraded credibility: e.g. near the operational limits of the model or input data to the model is considered a special case.

3) Normal credibility: Model output is considered to meet expectations for a model of the current \textit{Fidelity} level.

It must be noted that the third credibility level, i.e. “Normal credibility”, is connected to the static measure of \textit{Fidelity} for the specific model, and it is essential that the users understand that “Normal credibility” could have completely different meanings for models of different \textit{Fidelity}. The concept of \textit{Credibility\_Level} and how it may vary throughout a simulation are shown in Figure 4-10.

In addition to presenting credibility information of individual models, the proposed tool also provides an overview of model dependencies. Two visualization techniques have been considered: dependency graph and Design Structure Matrix (DSM) as shown in Figure 4-11.
Figure 4-11: Left: Dependency graph with four models, M1 to M4. The directions of the arrows indicate dependencies, and the color of the ellipses may represent a selected credibility measure. Right: A DSM representation with identical information content to the dependency graph. Model input indicated by “X” in the columns, model outputs indicated by “X” in the rows.

For small-scale simulators that contain only a few separate models, the dependency graph may be an intuitive visualization but for a more complex system like a large-scale simulator, the DSM is more feasible and is therefore the selected visualization technique.

According to the evaluation, the tool facilitates an intuitive overview of model dependencies as well as credibility information for individual models and for a simulator as a whole. This implies a support for detecting test plan deficiencies, or that a simulator configuration is not a suitable platform for the execution of a particular test. Furthermore, model developers and end-users are encouraged to reflect upon central credibility aspects like intended use, model fidelity, and test worthiness in their daily work. In this way, the tool creates a link between model developers and end-users, supporting continuous model improvement. Finally, as input to future improvements of the tool, it was found that registering and automatic post-processing of credibility information is needed; during simulator testing, test engineers can at best handle only aggregated high-level credibility information. Further development and industrial implementation are currently ongoing at Saab Aeronautics.
Discussion and Conclusions

The overall impression from the feedback obtained during conferences, from paper reviews, and from the numerous discussions with contacts in both industry and academia is that there is significant interest in and appreciation of M&S credibility assessment in general, and model validation in particular. However, in practice model validation is too often seen as a final check rather than a continuous activity integrated in the development and use of simulation models. As M&S projects are typically difficult to estimate, this may result in available resources being spent before any structured validation activities are initiated. Stakeholders may also be eager to obtain simulation results, which may lead to models being used prior to proper validation activities have been carried out. Here lies the paradox of model validation; the best or, as argued by some the only, way to conduct model validation, and in a broader scope credibility assessment, is to actually use the model. Obviously, and especially for safety-critical applications, this is not good enough. Thus, a reasonable conclusion is that a balanced effort of assessing and improving M&S credibility has to take place prior to model usage and during model usage. With limitations normally present in system development projects related to budgets, time schedules, and availability of skilled personnel, this is easier said than done. As concluded by Pace (2004):

“Administrative guidance, not even when stated in standards or DoD directives and other formal documents, seems unable to change modeling and simulation behaviors in ways that are needed to improve the credibility of modeling and simulation results.”

To change behaviors, methods need to be simple, efficient, and “good enough”, and there has to be an endeavor for assessing and improving M&S credibility, both from M&S
practitioners and from management. In short, credibility assessment and its related activities such as model validation have to be more fun; as an M&S practitioner, the author is well aware that developing a model, making it run, and occasionally also using the model are commonly seen as enjoyable parts of M&S. Validating the model, which means questioning the model and thereby questioning oneself, is not always considered equally rewarding. In addition, when lacking sufficient organizational and methodological support, model validation tends to become a tedious and thereby rare event.

The following subsections provide the main conclusions from the research, which are believed to support a change in behaviors in ways that are needed for balanced assessment and improvement of M&S credibility to take place in practice. The final subsection provides condensed answers to the research questions defined in section 1.3.

5.1 Organization and Independence

Part of the solution to the implications described above may be to increase the level of independence by handing over specific V&V activities to someone for whom V&V is the main interest. A parallel can be drawn with the software community, where software testing alone includes a set of established professions such as test leader, test designer, and tester. Another parallel can be drawn with product development and the domain of physical testing, with its common distinction between design engineers and test engineers. Thus, a transition towards more independence in V&V is desirable, but what is a suitable degree of independence in a specific context may be discussed.

For the organizations studied in paper [II], it is seen that roles and processes are applied to enable some degree of independence in V&V, especially regarding integration testing. At Saab Aeronautics, this has been enabled by the introduction of the methodology presented in paper [I]. However, since none of the studied organizations have a dedicated group or department for model V&V, individual model developers still have a great responsibility to decide how far to go with building confidence in a specific model in terms of both planning and execution.

At the time of writing, an independent organizational body for performing V&V and judging the credibility of M&S is not considered feasible for the studied organizations. In current situation, the following three measures to improve V&V are seen as feasible for implementation; 1) increased collaboration between model developers, 2) increased involvement of end users, and 3) introduction of a special group to support model developers with V&V expertise. Most likely, the implementation of the first two measures would be easier with the third measure already in place. Regardless of the level of ambition, any V&V organization should be carefully designed to avoid unnecessary overhead, and an appropriate timing for introduction of new V&V routines is likely to be a crucial factor for a successful implementation.
5.2 Automation

As stated above, efforts to assess and improve M&S credibility need to be balanced and take place both prior to model usage and during model usage. As M&S is used throughout system development, this implies that also activities related to assessing and improving M&S credibility should be conducted coherently throughout system development. That is, rather than being isolated rare events, activities like V&V should be carried out in a continuous process. For this to happen, the work effort required for thorough V&V has to be decreased. As shown in paper [II], a key enabler for this is a higher degree of automation.

What should be automated is repetitive work with limited necessity for human judgment. Areas typically suitable for a high degree of automation are for example import of measurement data, pre-processing, batch simulations, post-processing of simulation results, result plot generation, residual analysis, and documentation. Common obstacles to an increased level of automation are the lack of well-defined intended uses, validation metrics, and acceptance criteria.

As the uncertainty aggregation method proposed in paper [IV] raises the level of abstraction from uncertainty of component input parameters to uncertainty of component output characteristics, it is well suited for continuous model validation supported by UQ. That is, with fewer uncertainties, UQ is easier to apply and iterate as new information becomes available. If a model’s components include an uncertainty description, this is a good starting point for a higher degree of automation in model validation and UQ.

Furthermore, paper [IX] provides a framework for semi-automatic validation against system-level measurement data, using an ECS model as application example. The proposed automation efforts simplifies continuous model validation in both steady-state operation and in transient conditions, however with a focus on the later system development phases when system-level measurement data are available.

5.3 Approximate Uncertainty Quantification

In early system development phases, SA and UQ are seen as vital complements to traditional model validation using system level measurement data. However, regarding industrial application of SA and especially UQ, the research results indicate a large gap between academia and the organizations studied in paper [II]. Even though system simulation models are used throughout all system development phases for design and verification of both physical systems and control software, the application of structured methods for UQ is rarely seen. Three out of several reasons for this are that i) system simulation models often have a large number of parameters and inputs which may be uncertain, ii) the information required for characterization of uncertainties is often missing or difficult or expensive to extract, and iii) UQ is often of interest not only for steady-
state simulation results but also for time-series data representing for example full-length flight missions including transient conditions for multiple output.

Even though the challenge of computational cost is great, in an industrial context where information and resources are limited, the challenge of uncertainty characterization is even greater. This is especially true in early development phases. As the dimensionality of a UQ problem has a great impact on both the engineering workload and the computational cost required for UQ, it is vital to keep the number of uncertainties necessary for consideration down to a minimum and to offer approximate methods to UQ practitioners – otherwise UQ will not be widely used outside high-consequence areas with extensive resources for UQ.

In paper [IV], it is shown that – compared to using a model’s original component parameters – the proposed uncertainty aggregation method may result in a significant reduction of the number of uncertain parameters that require consideration in UQ. With only limited system level measurement data available for model validation purposes, the proposed uncertainty aggregation method provides means to use knowledge of component level uncertainty for assessment of model top level uncertainty. This approach mitigates the common situation of lacking data for characterization of individual parameter uncertainties.

The approximate UQ using aggregated uncertainties, demonstrated in paper [IV], is positioned somewhere in-between the concept of a comprehensive UQ resulting in probability boxes and no UQ at all, and is considered to be a pragmatic alternative suitable for the common M&S practitioner within the area of system simulation.

5.4 End-User Support for M&S Credibility Assessment

Traditionally, the only objective means for a model or simulator end-user to assess the credibility of simulation results is thorough model documentation – at best. Model documentation is indeed important, but far too often the documents are not read by end-users, implying that M&S credibility is assessed with a high level of subjectivity (if assessed at all). As a consequence, models and simulators may unintentionally be used outside of their operational domains and for testing not covered by their intended uses. In some cases this may be fine but in other cases this may lead to simulator run-time errors, invalid simulation results, invalid interpretation of simulation results, or that design errors are detected at a later stage. The more complex the simulator, the likelier the user is to become overconfident in the results and – unfortunately – the harder it is to make an objective assessment of simulation result credibility.

The proposed method for simulator credibility assessment is a response to traditional document-centric approaches, and is considered to be well-suited for use in large-scale simulators. The method encourages both model developers and simulator users to reflect upon central V&V aspects like intended use, operational domain, fidelity, etc.
Furthermore, simulator users are notified if the credibility of a specific model is degraded directly during testing, encouraging further investigation and an increased use of existing model documentation.

In most development testing carried out in simulators, the unit under test is software for control or functional monitoring. The proposed concept helps the user to consider the models surrounding the unit under test, such as models of physical systems. In addition, the concept of credibility assessment will become increasingly relevant as the level of detail in the models also tends to increase for real-time simulators. However, it should be noted that the proposed concept is deliberately pragmatic rather than mathematically stringent, and does not provide any guarantee of credibility at either model or simulator level.

The implementation of the proposed methodology, which is currently ongoing at Saab Aeronautics, affects several engineering disciplines: i) model developers need to develop, implement, and maintain information on model credibility; ii) simulator developers need to reintegrate, test, and maintain the updated models; iii) it will affect test engineers’ way of working regarding test planning, execution, and analysis of test results; and iv) it will affect test leaders’ and decision makers’ way of working regarding test worthiness declaration. To keep the cost of implementation and maintenance down, the methodology is implemented only for a carefully selected subset of simulator models.

5.5 Answers to Research Questions

This section provides summarizing answers to the research questions from section 1.3.

RQ1 How can model validity be assessed in the absence of system level measurement data?

The uncertainty aggregation method proposed in paper [IV] (also described in papers [VII] and [VIII]) addresses a challenge pointed out by Pace (2004), namely the one of limitations in items required for effective V&V, such as measurement data for model validation purposes or data for characterization of uncertainties. With only limited system level measurement data available for model validation purposes, the proposed uncertainty aggregation method provides means to use knowledge of component level uncertainty for the assessment of model top level uncertainty. Thus, the uncertainty aggregation method provides an answer to RQ1 regarding how model validity may be assessed in the absence of system level measurement data.

RQ2 How can uncertainty information obtained on component or sub-model level be used to assess uncertainty on model top level?

Paper [IV] describes the definition and implementation of a component output uncertainty description. Furthermore, it is shown how component level uncertainty may
be characterized and how such information may be used in an approximate model level UQ. To summarize, the proposed uncertainty aggregation method addresses some of the needs identified in papers [II] and [III], and has been found useful for the industrial applications tested in papers [IV] and [V]. Thus, the proposed uncertainty aggregation method also provides an answer to RQ2 regarding how component level uncertainty information may be used to assess uncertainty on model top level.

RQ3 Which model validation techniques are suitable for use in an industrial applicable process for development and integration of aircraft vehicle system models?

As a short answer to RQ3, a number of established techniques fit straight into the picture and are currently used at Saab Aeronautics. Examples are bottom-up testing, desk checking, sensitivity analysis, historical data validation, and visualization/animation. A large number of validation methods already in industrial use are identified in paper [II]. More interesting is probably who validates (the model developer or someone else?), how validation is executed (degree of automation in validation activities?), and when validation is executed (validation seen as a final check or a continuous activity?). These aspects are related to the methodology for development, V&V, and export of multipurpose simulation models presented in paper [I]. It has been found that providing engineers with descriptions of a logical workflow supported by templates and checklists gives them a framework to act within and a feeling of better control of the model management activities. It was found that with the described methodology in place engineers were more willing to take full responsibility for new models, but also for “legacy” models when those are updated and re-engineered.

An interesting area related to RQ3 is also the emergence of new roles and organizational changes in the M&S domain. New organizational roles able to adopt a holistic approach for development, V&V, and usage of simulation models would provide significant support in the industry’s endeavor to expand the use of M&S towards virtual testing and virtual certification.

RQ4 How can credibility information for individual models and for a simulator as a whole be presented to simulator users to support the detection of deficiencies in model representativeness – prior to, during, and after a simulator run?

The methodology and tool for simulator credibility assessment presented in paper [VI] provide one possible answer to RQ4 regarding how to support simulator users in the detection of deficiencies in model representativeness. The proposed methodology defines a number of credibility measures which are presented to the user in a GUI directly during simulation. In addition, model dependencies are clearly visualized by means of an
automatically generated DSM. Model dependencies and part of the credibility information can also be presented to the user and decision makers already prior to simulation, and thus provide decision support to the simulator test worthiness declaration.

**RQ5** Which credibility measures and which level of detail in the credibility assessment is industrially applicable for large-scale aircraft system simulators?

In the realization of a credibility assessment tool running in large-scale simulators, there are numerous design decisions to be taken. RQ5 regards such a design decision, namely which level of detail in the credibility assessment is industrially applicable for aircraft system simulators – from one single static quality stamp on each individual model to a dynamic run-time uncertainty estimation of each signal in each individual model interface in the complete simulator? As an answer to RQ5, it was found that signal level uncertainty estimation is out of scope from an industrial point of view. Available resources do not allow the development and maintenance of such a detailed approach. As described in paper [VI], rather a set of static credibility measures in combination with a dynamic credibility measure on individual model level has been found industrially applicable, i.e. simple enough for industry-scale implementation and good enough to actually provide user support.
To utilize the full potential of simulation models, including an expanded use of M&S towards virtual testing and virtual certification, the development of efficient methods for credibility assessment is a key challenge. This work demonstrates methods for early model validation by means of approximate UQ, as well as a method facilitating less document-centric credibility assessment of large-scale system simulators. Even though the methods are promising and some of them already in use at Saab Aeronautics, there is clearly room for improvement and generalization, which requires continued research and development.

However, new methods and tools are not all that is required. From an industry perspective, a considerable challenge for the near future is to ensure that good examples of existing methods are disseminated and applied on a larger proportion of the models used in system development. This requires an organizational setting that enables thorough knowledge exchange between different M&S teams from different organizational domains and joint efforts in further development and application of methods and tools for credibility assessment and V&V. To obtain the long-term focus needed for this to happen in practice, the introduction of a dedicated organizational entity to support model developers with V&V expertise and coordination is probably required. To enable fruitful collaboration, such an organizational entity should not be used for judging the credibility of M&S efforts. As pointed out by Rider (2014), M&S professions have to end up owning the quality of their work, and V&V should engage as a resource to improve quality.

In addition to the organizational aspects described above, an interesting area which may have an impact on future methods for UQ in system simulation is the evolution of the Modelica language. The Modelica language is at present deterministically oriented. To clarify, uncertainties in parameters and inputs need to be handled separately from the Modelica model. Bouskela et al. (2011) propose a set of requirements concerning the Modelica language that enable parameter uncertainties to be defined in the form of PDFs. Future developments of the Modelica language and related tools towards better support
for handling uncertainty and performing efficient stochastic simulations will be followed with great interest. Equally interesting will be to follow and hopefully also to influence the development of the FMI standard. This is related to alternative UQ methods, but is also of great interest since the standard has the potential to make model integration in general more efficient. Fortunately, industrially applicable methods for model integration, co-simulation, and V&V taking advantage of the FMI standard will be investigated and further developed in the recently started EUREKA/ITEA3 project OPENCPS (2016).

Regarding the proposed approximate uncertainty aggregation method, an area for future work is to conduct a detailed analysis of the degree of uncertainty introduced by the proposed aggregation method. Comparisons with established methods require an application example within component-based system simulation, sufficiently limited and with extensive measurement data available to allow thorough and stringent UQ resulting in probability boxes.

Finally, it will be very interesting to follow the continued development, implementation and industrial use of the simulator credibility assessment tool. At the time of writing, the main focus is on implementing what has already been scientifically evaluated during this research project. However, an interesting but equally challenging area for future research is the definition of a fidelity measure, not solely related to the level of detail used in the modeling process. A recommendation is to investigate whether the fidelity levels defined in (ICAO, 2009), can be extended to be applicable also for the wide range of aircraft subsystem models used in today’s development simulators.


Appended Papers

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-127380