Linköping University | Department of Management and Engineering
Master’s thesis, 30 credits| Industrial engineering and Management
Spring 2020| LIU-IEI-TEK-A--20/03889--SE

A Conceptual Study on Model-Based Systems Engineering and Data Driven Methods in the Context of Complex Products and Systems.

Appu Balachandran
Dennis Karlsson Tunhult

Supervisors: Nicolette Lakemond & Gunnar Holmberg
Examiner: Dag Swartling

Linköping University
SE-581 83 Linköping, Sweden
+46 013 28 10 00, www.liu.se

Abstract

Increased use of data is influencing the existing practices in the engineering domain, including that of systems engineering. Complex products and systems (CoPS), along with its predominant methodology of development, Model-based systems engineering (MBSE), is no exception to this. This thesis explores the possible integration of the emerging data driven methods and the established model-based methods in the context of CoPS development. It also explores what the implications of such an integration could be for the organizations building such systems, the system integrators. To analyse the current state of the art in CoPS development and model-based methods as well as the emerging trends in data driven methods, this research employs an integrative literature review method. The literature search concluded in 71 selected articles to be reviewed. These articles where divided over three main categories, CoPS, Model-based systems engineering (MBSE) and data driven methods. The results of the analysis suggest that data driven methods and the model-based methods complement rather than compete throughout the innovation life cycle of CoPS. The findings indicate that an integration of the methods is beneficial to the architectural, systemic, and component level innovation in CoPS. MBSE and data driven methods could however have different levels of influence in these three types of innovation. The findings indicate that MBSE could have more influence in architectural innovations, while data driven methods could be more influential in systemic and component innovation. The continuous innovation in the use phase of system is also seen to be improved by this integration. The system integrators benefit from the improved project to project learning resulting from the integration which enhances their economy of repetition. An integrated method could also increase the speed of which decisions can be made while still maintaining reliability in the system. The results indicate that the number of iterations could increase due to the increased feedback of data and the learnings gained from it, which could pose some challenge to the existing project management methods. Further research is needed to find out what are the full benefits of an integrated method and identify other potential conflicts.

Keywords: Complex product systems, Model-based systems engineering, Literature review, Innovation, Project management.

Acknowledgements

After having spent almost seven months working with our thesis, we would like to thank everyone who guided and supported us throughout this project. This master thesis project has been a great learning experience for both of us.

First, we would like to thank our supervisors, Nicolette Lakemond and Gunnar Holmberg for introducing us to this topic and guiding us throughout the project. Thank you for your insightful suggestions, thought provoking questions, encouragement, and support.

We would like to thank Dag Swartling, our examiner, who provided us with thoughts and advises to improve the overall quality of our report. Thank you for your support in this thesis project.

Finally, we would like to thank our opponents Jacob Hjalmarsson and Markus Drugge, who provided us with important suggestions and feedback that helped to improve the quality of our report. Thank you for your efforts in providing a good opposition to this thesis.

This thesis, we hope, will be an informative read to those who are interested in understanding complex products and systems, and how model-based systems engineering, and data-driven methods play an important role in their development.

Table of Contents

List of Figures

List of tables

1. Introduction

This section presents the background of the thesis as well as presenting of the general theoretical areas that the thesis will cover. The background and problem formulation then lead into the purpose statement and two specific research questions that will be answered. The section closes with an outline for the rest of the thesis.

1.1. Background

Complex product systems (CoPS) are highly customised technical systems, with a hierarchical structure consisting of several interconnected subsystems (Hobday, 1998). According to Hobday (1998), it requires a large amount of knowledge and skill to develop them due to this high degree of interconnectedness. Telecommunication networks, control systems, capital goods, aircrafts are some examples of CoPS. When compared to the mass-produced goods, the development of CoPS is considerably different. It involves multiple stakeholders such as system integrators, suppliers, regulators, and users, spanning the organizational boundaries (Hobday & Rush, 1999). As CoPS are business critical for the users and mostly tailormade, the degree of user involvement in the development process is considerably high (Bonaccorsi & Giuri, 2000) (Hobday, et al., 2000). The dynamics of innovation in CoPS is also unique when compared to that of mass-produced goods (Hobday, 1998). These special characteristics make the development of CoPS difficult, requiring special capabilities to manage various aspects of the system. According to Rhodes and Ross (2010), the dynamic nature of a CoPS poses a challenge to modelling, testing, validation, and evaluation of such systems. Understandably, CoPS development requires a holistic and multidisciplinary approach.

According to Ramos et al. (2012), systems engineering (SE) is a methodology that is ideal for developing complex systems where there is a need to deal with different competencies, multiple stakeholders, interconnection between subsystems, etc., making it a preferred method. According to the International Council on Systems Engineering (INCOSE), origins of systems engineering (SE) practices can be traced back to the defence programs that were initiated in the US and the UK and later it emerged as a preferred methodology owing to its ability to handle complexity and manage the associated changes (INCOSE, 2015). Systems engineering can be defined as an interdisciplinary approach aimed at developing systems successfully by capturing the customer requirements and the functionality needed, early in the system development phase, documenting it, and subsequently performing activities such as design synthesis, verification and validation in relation to the whole system life cycle (INCOSE, 2015). Even though SE started as a document based method, as the complexity increased in system building, the use of models for conceptualization became a common practice and it started replacing the document driven method (Madni & Sievers, 2018).

Ramos et al. (2012) define Model-Based SE (MBSE), as the process of formally applying the principles, tools and methods associated with modelling to the development of complex sociotechnical systems, that are interdisciplinary in nature, throughout its life cycle. In MBSE, in contrast to the document focused approach to systems engineering, the model is the true source of knowledge and the system model the prime artefact in the development of a system (INCOSE, 2015). According to Holland (2015), the development of computing power has enabled MBSE to capture the system characteristics using models which can subsequently be used for verification and validation. In the development of CoPS, MBSE maintains an 'information model' - visible to those involved in the development - from the identification of system requirement to the subsequent activities such as decomposition of requirements to components, system integration and verification and validation (Freidenthal, et al., 2014) (Holland, 2015). In this way MBSE covers the whole development cycle of the system. The benefits of using MBSE are the enhanced communication between stakeholders, collaboration among specialists, knowledge capture, standardization, reduced risk, improved quality, traceability of changes to name a few (Freidenthal, et al., 2014) (Holland, 2015).

Though MBSE is shown to have many advantages over the document-based systems engineering, it is not without shortcomings. The customers who are used to the document-based systems need a cultural transformation to adopt MBSE in their system development (Bonnet, et al., 2015). According to Madni and Sievers (2018), integration of the models could be difficult as the models are heterogenous as they originate from the different disciplines. Establishing and reviewing a baseline satisfying the interest of the multiple domains involved, could be challenging. Apart from solving such interoperability issues, MBSE also need to develop an approach to bridge the gap between the stakeholders, who have a non-technical/non-expert background, and the system engineers (Madni & Sievers, 2018).

The emergence of new technologies based on data have started to affect the established engineering practices, including systems engineering. Artificial intelligence (AI), big data analytics and internet of things (IoT) are some of the important technologies that have the potential to impact the current system development methodologies. The abilities of organizations to gain insights and take effective decisions have improved due to the access to large amount of data that can be analysed using specific techniques to gain insights (Marjani, et al., 2017). According to

Gao et al. (2013), data driven methods, that based on the big data analytics and signal processing, do not necessarily require prior process knowledge. Though the output characteristics are different, data driven methods can process data and give output faster when compared to the model-based approaches (Geffner, 2018). However, the quality and accuracy of data becomes very important as the results would be affected greatly if these factors are overlooked (How, et al., 2019).

According to Hybertson et al. (2018), the predominantly model based nature of systems engineering needs a new perspective which include emerging data driven methods in it. In the last few years there can be seen some attempts to create conceptual models that incorporate data driven methods with the established model-based approach. For example, the triple V model by Li et al., (2019), and framework on Evidence-based systems engineering by Hybertson et al. (2018). However, further studies are needed to realize all benefits that both model-based and data-driven methods brings in the development of complex systems and understand its impact on established management practices.

1.2. Purpose

The purpose of this thesis is to explore what implications the increase of data driven approaches may have on established MBSE methodologies in the context of complex product and systems and identify how these two areas could be integrated on a conceptual level. This includes what changes might be necessary in the technical processes of developing complex products and systems using an MBSE methodology, as well as its implications for the system integrators.

The aim of this thesis is then to contribute to an integration of data-driven approaches and model-based approaches, where both types of methods together drive the progress. The thesis should also contribute towards the development of a conceptual model that utilizes the benefits of both methodologies.

This purpose leads to two specific research questions for the thesis:

i) How could data driven methods be integrated with MBSEs in the development of complex products and systems?
ii) What could the implications of such an integration mean for the system integrators?

1.3. Structure of the thesis

Chapter 2 outlines the methodology of the thesis, how the data was collected, analysed and why this methodology is the appropriate one. Chapter 3, theoretical framework, aims to give the reader a necessary understanding of Complex products and systems capabilities, innovation life cycle, organizational structure, and key capabilities. Chapter 3 also presents the basics for model-based systems engineering and data driven methods. Chapter 4 gives a descriptive analysis of the chosen articles followed by Chapter 5 which presents an integrative analysis of complex products and systems, model-based systems engineering, and data driven methods. Chapter 6 presents the conclusions of the thesis and chapter 7 discusses the recommendations for future research.

2. Methodology

In this section, the methodology used to perform the thesis is presented. First, a general research design is defined followed by a more detailed strategy. The strategy aims to first put the thesis in a broader context, making it more generalisable, to then in detail explain how the research was carried out, the data collected and analysed.

2.1. Research Design

As the purpose of this thesis is to analyse and synthesize how two different approaches to develop complex systems might be integrated, and what effects this could have on system integrators, this study will mainly take an exploratory approach. An exploratory approach fits well when the aim of the research is to get a better understanding of an area or phenomenon and is especially suited to answer questions stated in a "how or "what" fashion (Saunders, et al., 2019). Along with the exploratory approach, this thesis will mainly use a qualitative research design. In a qualitative methodology, in contrast to quantitative, the data collected focuses on words, meanings, concepts and relationships rather than numbers and quantifiable results (Bell, et al., 2017). A qualitative research approach is also connected with the researchers interpreting the data as they need to make sense of different meanings in the studied subject (Saunders, et al., 2019). As this thesis aims to develop a conceptual framework, where a conceptual framework can be considered a synthesis of relationships of concepts (Jabareen, 2009) which requires creativity (Torraco, 2016), a qualitative approach was deemed the preferred choice of research design.

2.2. Research Strategy

The research strategy was developed to put the thesis and its result in a broader context by placing it in an established theory building framework. The strategy was also made to incorporate a methodology that would support the understanding of both emerging, as well as more established fields. The methodology was also chosen to allow for the combination of different concepts in those fields.

2.2.1. Theory building framework

Even though this thesis is only meant to establish a conceptual model, and can therefore not claim to be building theory, it is the aim that by putting the conceptual model in the context of a theory building framework, this can help establish a broader structure and support with more concrete definition of terms. This is done especially since the conceptual development is treated as a separate and necessary phase in several theory building frameworks (Storberg-Walker, 2003).

For this thesis, the choice was made to use Lynham's General method of theory building (Lynham, 2002). The reason for choosing Lynham's method over other theory-building frameworks was for its general usage, that it is not restrictive to any specific philosophical view, research design, or approach to the research (inductive/deductive) (Storberg-Walker, 2003). The framework was also specifically developed for applied disciplines (Lynham, 2002).

The framework consists of a total of five stages. These phases are:

i) Conceptual development – The purpose it to develop a conceptual model where the final output should be a conceptual framework that often is represented through a model or metaphor.
ii) Operationalization - where the purpose is to establish a connection between the conceptual model and practice.
iii) Confirmation or Disconfirmation - in this phase a research agenda or studies should be planned, implemented and then either confirm or disconfirm the theoretical framework for the specific area in which it applies.
iv) Application - the confirmed theory then needs to be applied to the real world to address the issue or phenomenon which it was developed for.
v) Ongoing refinement and development - the final stage of theory development process ensures that the theory is kept up to date with the latest findings in the area. It ensures that it is always reliable and when it is no longer accurate to its application, it is updated, changed or discarded as false.

The phase of Lynham's framework that this thesis aims to contribute towards is the conceptual development phase. Clarification on the definition of a conceptual framework and its necessary properties will be further defined in section 2.2.2.

2.2.2. Conceptual development of a framework/model

A conceptual framework or model can generally be described as a collection of interlinked concepts which together contributes to the understanding of the issue or phenomenon (Jabareen, 2009). A concept can then further be described as consisting of several distinct, but non separable components and can be understood as the accumulation of these components (Jabareen, 2009).

With these definitions in mind, the conceptual development process can then be described as the process of gaining deeper understanding about a subject to depict the current and best practice of the area of study, with the purpose of developing a conceptual model or framework (Lynham, 2002). There is no specific method that should be used to develop a conceptual model, rather, the method should be chosen in accordance to the purpose of developing the model. The method used in this thesis is the integrative literature review, which will be explained further in section 2.2.3. and 2.2.4. However, whichever method is chosen to develop the conceptual model, there are some characteristics that needs to be defined and developed during the conceptual development. During the process, the key elements of the theory should be identified, the relationships between these elements needs to be mapped and the scope under which the model can be expected to function should be defined (Lynham, 2002). This process should then culminate to an informed conceptual model or framework which is not simply a collection of concepts or elements, rather an interpretation of both their relationships to each other and an interpretive understanding of the real world (Jabareen, 2009). This thesis will focus on identifying the key elements and mapping their relationships, which can contribute to further development of a conceptual model.

2.2.3. Literature review

This thesis focuses on the advancements in MBSE, data driven and CoPS. Hence it is important to look at the state of the art in these fields. Critical review of existing literature and synthesis of new knowledge, according to Torraco (2016), is one of the main purposes of conducting a literature review. According to Snyder (2019), to build new conceptual models or theories, it is important to know the gaps in research, which a literature review method can reveal. In case of emerging topics, it is likely that there are contradicting viewpoints which none of the individual literature discusses about (Torraco, 2016). A literature review gives an opportunity to investigate different aspects and bring about a clear understanding of underlying issues. Keeping these factors in mind, a literature review method was found to be a suitable method for answering the research questions of this thesis.

According to Snyder (2019), the literature review method can be broadly categorized into three type namely systematic literature review, semi systematic literature review and integrative literature review.

i) Systematic literature review: it is aimed at collecting empirical evidence that is selected based on a pre specified inclusion criteria, often using statistical methods (e.g. meta-analysis) to identify different patterns and relationships that emerge. This method reduces bias and provides reliable results.
ii) Semi systematic literature review: it is aimed at studying topics that several groups have conceptualized differently involving diverse disciplines, which hinder the use of a

systematic review. It looks at how topics have evolved overtime across different research traditions, providing a historic overview.

iii) Integrative literature review: this method is suitable for evaluating, critiquing, and synthesizing existing literature to aid the development of new theoretical frameworks and identification of emerging perspectives.

For this thesis, an integrative literature review is found to be suitable as it aligns to the purpose of the research and seems appropriate to aid in answering the research questions. The details of which are discussed in the next section.

2.2.4. Integrative literature review

When developing a conceptual framework, it is important to look at the issues from multiple perspectives. This requires the researchers to be creative in collecting the data from diverse sources to get a holistic view of the topic and the integrative literature review, can be a suitable method for this type of research (Whittemore & Knafl, 2005) (Snyder, 2019). The integrative literature review is suitable for both mature as well as emerging topics. In mature topics, it can result in an upgrade of the existing concepts and in emerging topics, it can result in the creation of new concepts (Torraco, 2016). Model driven methods and complex systems both fall under a more mature topics and data driven methods is still an emergent field of research. Though the integrative literature review is a suitable method, there are some key aspects that need to be taken into consideration to make it rigorous. According to Snyder (2019), the literature review should have a step by step approach and it is important to select the articles in a transparent manner to ensure the quality and reliability. Since this method allows creative ways of collecting data, such as combining experimental and non-experimental data, such diversities arising out of the breadth, could result in higher complexity (Whittemore & Knafl, 2005). According to Whittemore and Knafl (2005), the strategies for extraction of primary data as well as the strategies for data analysis are of prime importance. Developing a right strategy is important in enhancing the rigour of the integrative review (Whittemore & Knafl, 2005).

Whittemore and Knafl (2005) propose certain strategies which can act as counter measures to strengthen the scientific rigour of the integrative literature review.

i) In the problem formulation stage, framing a purpose that is well defined is important. This can specify the variable which help in identifying if the information gathered is relevant or not.

ii) In the data collection stage (literature search), having the right keywords is essential since inconsistent search can result in a loss of about 50 % of eligible literature. The search method should be explicit with keywords, databases, secondary search methods, criteria for selection and rejection of literature. The use of a decision tree is highly recommended.

iii) In the data appraisal stage, the defined quality criteria for the data is considered. Employing more than one criteria for the primary sources, can increase the quality of data.

iv) In the data analysis stage, the focus is on interpretation and synthesis of the collected data using a methodical approach. In integrative reviews, methods such as categorization, grouping and coding are done before data reduction, comparison, conclusion, and verification. It is possible to use diverse methodologies to handle varied data in the integrative method.

v) In the conclusion stage, the emergent patterns are subjected to interpretation. One challenge could be handling the conflicting evidences, if such a situation emerges. This can also be an indication of the need for future research.

A thematic structure that Toracco (2016) suggests was used for this thesis to organize the sub-themes as it helps in building clarity on how the different topics are linked together as well as brining coherence to the different ideas. Table 3 in section 2.4.1 shows the emergent sub-themes of the literature review. To ensure the quality and reliability of the research, recommendations from Whittemore and Knafl (2005), Torraco (2016) and Snyder (2019) were incorporated. These recommendations call for a transparent step by step approach supported by a sound strategy to conduct an integrative literature review, which were used as a guide. Section 2.4.1 describes in detail, how these recommendations were implemented in this thesis.

2.3. Pre-study

At the start of the thesis, a pre study was performed to gain further understanding about the subject. There was specific emphasis put on gaining understanding within the methodology of systems engineering as this was considered the outer boundary of the content of the thesis. The pre study then continued with a focus on what characterizes a complex system. Finally, focus was put on understanding the two main concepts of this thesis, model-based systems engineering, and data driven methods for systems engineering. To gain deeper understanding of these two concepts, different literature sources where studied, as well as working through a short practical case example. The short case was developed using a model-based system engineering methodology, and once done, reflected upon on how the system would be impacted if data driven methods would have been utilized.

The various phases of the model-based system engineering activities were explored by using the example of a simple door access system that allows selective access based on identification such as a key card or tag. This case study helped the authors to understand how MBSE provides a robust way to decompose requirements to design elements. The business and stakeholder requirements were converted into functional requirements. The system architecture was designed to fulfil these requirements. The architecture was mapped to the design definition and to the physical components. Verification and validation criteria were created and were mapped back to the system design. Since all the aspects ranging from requirements to verification and validation are linked together, traceability and impact analysis was made easy with the use of MBSE. It was reflected upon by the authors that the process of capturing the requirements from stakeholders and decomposing them to functions and components is in dependent on previous experience and knowledge. It was also reflected upon at how the incorporation of data driven methods could instead influence the system at various levels.

The knowledge gained from this initial pre study contributed to the definition of the purpose for this thesis, the specific research questions, the themes of the thesis and the different keywords in the themes. In this sense, the pre-study can be viewed as the first of Whittemore & Knafl (2005) strategies for a strong integrative review.

2.4. Data Collection

From the pre-study, the main themes of this research were identified as complex products and systems, model-based systems engineering, and data driven methods, with systems engineering as the background. The articles from the pre-study phase were used to identify the keywords and key phrases to be used for the literature search.

Complex products and systems	Model-based systems engineering	Data driven methods
"Complex products and systems	"Model based systems engineering"	"Data driven"
"Complex product systems"	"MBSE"	"Data analytics"
	"Complexity"	"Data based"
	"Complex"	"systems engineering"
		"Complexity"
		"Complex"

Table 1. Keywords and key phrases for the literature search

The keyword search where structured in a way so that each article would be in the context of complex systems. The keyword/key phrase search was then divided into the three themes to identify the most relevant articles. The three different searches, labelled S1, S2, and S3, and the structure of the different keyword searches that were carried out can be viewed in table 2.

S1	*"Complex products and systems" OR "Complex product systems"*
S2	*("Model based systems engineering" OR "MBSE") AND ("Complexity OR "Complex")*
S3	*("Data analytics" OR "Data driven" OR "data based") AND ("Complexity" OR "Complex") AND ("Systems engineering")*

Table 2. Strings for the literature searches S1, S3 and S3

A comparison study of databases by Falagas et al. (2008) explored the capabilities of Scopus, Web of science and Google scholar. The study suggests that Scopus had several advantages when compared to Web of science and Google scholar owing to its wider journal and subject range. A trial search of databases (Scopus, Web of Science, ScienceDirect and Google Scholar) performed by the authors of the thesis showed a similar result. Google scholar returned the most results, but with a large portion being not relevant to the subject of the thesis. Scopus returned a good number of relevant articles while the return in ScienceDirect and Web of Science was low and returned several duplicates with Scopus. For these reasons, Scopus and Google scholar where chosen as the databases for the literature search. By finding and using keywords which will provide a reliable results, along with a search strategy involving the use of databases that has the potential to give the most relevant articles, part of the second strategy from Whittemore & Knafl (2005) is achieved.

2.4.1. Article selection process

The initial process for article selection started with the definition of criteria for the selecting an article in the sample. According the Snyder (2019), the design of the inclusion and exclusion criteria is critical to ensure the quality of the review. The initial criteria for the first stage of selection in the review where; The first 200 articles in the search, sorted by reference. If the search gave a result that was reasonably close to 200, all articles in the search where selected. Only articles that where published in a peer reviewed journal or conference proceedings where included. There was one exception done for this criterion where a survey of MBSE methodologies published by INCOSE was included. This was done since the survey had a significant impact in the MBSE field. Further, only articles written in English where included. Finally, only articles published between 1993-2019 where included. The reason for having 1993 as the limit is because the early literatures describing CoPS were published that year. After the search with the criteria applied, 1023 articles where included in the initial selection. In the second stage of the selection, the abstracts of the 1023 articles where read and reviewed by the authors. At this stage focus was placed on the abstract's connection to one of the three main themes as well as the overall purpose. After the abstract review stage, 94 articles remained in the selection. The final stage of the article selection process consisted of a full read of the 94 articles. The emphasis here was on the match between the article and the purpose of the thesis. After this final read, 71 articles where chosen to be included in the literature review with articles published between 1997-2019 with a distribution according to figure 1. With clearly established inclusion and exclusion criteria which the authors thoroughly considered throughout the process, the third strategy by Whittemore and Knafl (2005) is addressed.

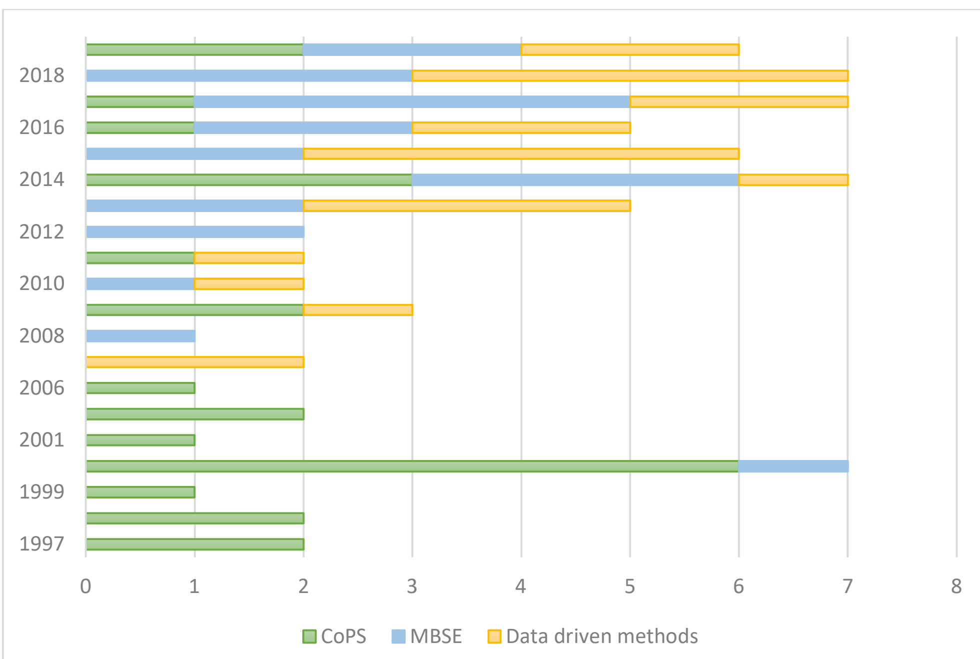

Figure 1. Distribution of articles per year

After the full read of the articles, subthemes where identified from each of the three major themes of the thesis. This was done to get a better structure of the review, as well as give a better understanding of the area based on the selected literature. The subthemes are presented in table 3.

CoPS	MBSE	Data driven methods
Characteristics of Complex products and systems	Traditional Model-based systems engineering	Data driven control and optimization
Innovation life cycle of Complex products and systems	Emerging cultural and technical challenges in Model-based systems engineering	Data driven modelling
Organizational structure in Complex products and systems suppliers	Case studies on Model-based systems engineering	Data driven monitoring and fault diagnostics
Key capabilities of developing Complex products and systems	Model-based systems engineering and data	

Table 3. Subthemes of the literature search

A representation of the article selection process can be viewed in figure 2 in the form of a process tree.

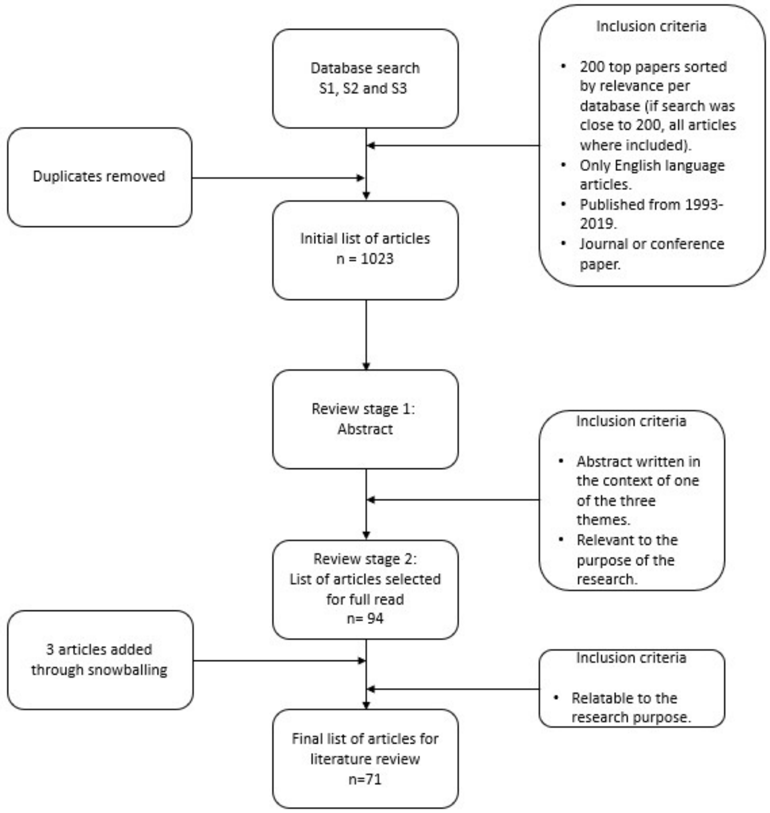

Figure 2. Process tree for the literature search

After the articles had been selected for the literature review, a descriptive analysis was done for each article. The aim of the descriptive analysis was for the authors to identify and describe important concepts and perspectives from each article relative to the thesis purpose. Finally, the findings from the descriptive analysis where subjected to a critical review. In the critical review, the different concepts and perspectives where critiqued and compared to achieve an integrated analysis where complementary and opposite views where analysed to build the conceptual framework. In the analysis, the authors both critically review and synthesise the different concepts from the literature, therein addressing the fourth strategy for a strong integrative literature review by Whittemore & Knafl (2005). The fifth and final strategy by Whittemore & Knafl (2005) is achieved in the chapter 6. Conclusions, and in chapter 7. Future research areas, where the findings are interpreted, and gaps for further development are identified.

2.5. Limitations

While the integrated literature review method allows for more creative ways of identifying and combining different concepts, it may not give a full view of the different fields. As the selection was done with focus on the purpose of this thesis, the selection is not meant to give an overview of either MBSE, data driven methods or CoPS. It is therefore possible that with the lack of a systemic review of the areas, concepts that could have been useful and influential to the research may have been missed in the selection. As the articles where sorted by relevance in the databases, it is also possible that more impactful articles within the field may not have been included in the selection.

3. Theoretical framework

The theoretical framework chapter aims at giving a broad overview to the three different areas of analysis (Complex products and systems, model-based systems engineering, and data driven development). The main purpose of this chapter is to provide the necessary knowledge about the main categories to be further built upon in the more in-depth analysis of chapter 4 and 5.

3.1. Characteristics of Complex Products and Systems

The concept of Complex product and systems (CoPS) was first investigated by Miller et al. (1995) (Ranjbar, et al., 2018). Although at that point simply referred to as complex systems. Miller et al (1995), while investigating the development of the flight simulator industry, laid the foundation for CoPS by defining certain systems of products that are large scale, have a high degree of connectivity, high customization and often show emergent behaviour, as a specific group of product systems that do not follow the general innovation model of mass produced goods. Hobday (1998) continued the development of CoPS, establishing it as a specific research area, defining characteristics as well as providing examples of products and systems that can be classified as CoPS. The definition provided for CoPS as "high cost, engineering-intensive products, systems, networks and constructs" (Hobday, (1998, pg. 690)). Since then, this definition has been commonly used by other scholars (Ranjbar, et al., 2018).

CoPS are high cost and are either produced in single unit, or small batches, and usually done through projects (Hobday, 1998). This high interoperability between stakeholders does however lead to an inherent issue in coordinating information to a higher degree than in mass produced goods (Hansen & Rush, 1998).

As implied by the word "complex", CoPS contain a high number of customized components and developing these components require a high amount of knowledge and skill (Hobday, 1998). The component architecture of these complex products and systems often become very large, difficult to manage, consisting of many interconnected, customized elements (Miller, et al., 1995). CoPS often also contain a high degree of technological novelty in its development (Ren & Yeo, 2006). Due to the architectural complexity and customized elements, technical uncertainty during development is a recurring issue in CoPS projects (Hansen & Rush, 1998).

Even though most complex products and systems exhibit mainly the same characteristics, these characteristics are not represented to the same degree. The different characteristics run on a scale of complexity and it is therefore difficult to make generalisable conclusions on different types of complex products and systems. (Hobday, 1998)

3.2. Innovation life cycle of Complex Products and Systems

According to Anderson and Tushman (1990), the industry life cycle of mass-produced goods is marked with the emergence of a technological discontinuity that causes a period of ferment in an industry, followed by the appearance of a dominant design that is accepted as the industry standard. After the selection of a dominant design, the industry enters a phase of incremental change till the appearance of the next technological discontinuity (Anderson & Tushman, 1990). Utterback and Suarez (1993) observed that the emergence of a dominant design results in an industrial shake out and shifts the focus from product innovation to process innovation.

As CoPS differ from mass produced goods in various aspects, there are some striking differences in the innovation life cycle of CoPS as compared to the conventional model. Even when technological discontinuities emerge, there is stability in the industry (Miller, et al., 1995). There are considerable entry barriers for newcomers in CoPS, and the mass entry and mass exit of firms are not observed (Hobday, et al., 2000). The industry shake-out following the emergence of a dominant design, a key aspect of the standard life cycle model, does not seem to hold good. A study of 'turboprop industry' by Bonacorsi and Guiuri (2000), found that even when there was a high concentration of competitors, the shake out did not occur. Instead, there was a stability resulting out of the coexistence of the competing firms. The technological changes affects the suppliers mostly as there can be entries and exists in the supply chain as a result (Miller, et al., 1995).

As CoPS are tailor made to specific requirements, the economies of scale do not apply and hence the shift of innovation from product to process is not seen (Peltoniemi, 2011). There can be multiple feedback loops throughout the development of CoPS and the innovation often continues even after the product is delivered to the customers, in different forms such as upgrades to sub-systems, performance enhancements etc. (Hobday, 1998).The innovation process is affected not only by the product characteristics but also by the organisational structure in CoPS (Nightingale, 2000).

3.3. Organizational structure in Complex Products and Systems

Hansen and Rush (1998) highlight organization and project structure as one of the 'hotspots' in CoPS. The high degree of innovation in CoPS warrants an organization structure that creates a conducive environment. Since such system development entail a close collaboration of multiple actors, the coordination requirements are much higher (Hobday, 1998). According to Hobday (1998), when a greater number of firms get involved in the different phases, the complexity of coordination increases. For the customers, CoPS are business critical units and hence they have a deeper engagement in the various development phases of CoPS (Hobday, 2000) (Davies & Brady, 2000). CoPS customers are often few and are very demanding which is an important aspect (Davies, et al., 2011). As the customers of CoPS are of prime importance and source of input for the development and innovation process, the customer focus required is much greater as compared to mass produced goods (Hobday, 1998). These aspects highlight the need of a structure that is flexible, facilitates coordination and has a strong focus on the customer requirements. An organic structure rather than a mechanistic one is suited for CoPS development (Hobday, 1998) (Hobday, 2000).

Clark and Wheelwright (1992) in their study, highlighted the advantage of 'heavyweight project teams' over the other type of development project teams in terms of specialization as well as integration in a new product development (NPD). A project is a temporary organizational form which strongly focuses on customer value while maintaining the close contact with the organizational members (Tonnquist, 2012). A single firm may not have all the capabilities and domain knowledge to develop CoPS, hence a network style of functioning is often adopted to enable collaboration (Hardstone, 2004). The project is the main co-ordination mechanism that enables stakeholders to interact, agree upon and realize the system, while maintaining the effectiveness in resource/skill mobilization and deployment (Hobday, 1998) (Hobday, 2000). The key issues in coordination are organizational structure, communication, technological competence development and customer interaction (Hansen & Rush, 1998).

The emergent properties in CoPS increase the degree of complexity and uncertainty, making the project management tools and techniques that are used by functional and matrix organizations ineffective when applied on CoPS (Davies, et al., 2011). Hobday (2000) emphasises on a 'project-based' organization structure which has dominant project lines, as opposed to the weaker ones in functional organization, for CoPS development. It is suited when there is need for a concurrent model of project management to promote innovation, ability to deal with uncertainties and ability to cope with emergent properties. It is also useful for resource sharing across firms when needed. The weak areas of this structure are interproject learning, coordination of resource across the different projects and reduced ability to exercise senior management control over the project (Hobday, 2000). While Hobday (2000) recommends a project led organization structure which is a balance between the project based organisation and matrix organization for CoPS, Davies and Brady (2000) argue that companies can use different organizational structures at different stages of the project as per the requirement. Davies and Brady (2000) observe that when it comes to the organizational structure, such firms adopt a project based form in the early phases such as proposal stage, a matrix form in later phases such as implementation stage and a functional form during operational support.

3.4. Key capabilities in developing Complex Products and Systems

As CoPS are mainly developed in projects, project management capabilities, such as risk management, scheduling and resource allocation are repeatably mentioned as key to succeed in the development process (Hobday, 1998) (Davies & Brady, 2000) (Nightingale, 2000) (Hardstone, 2004). Since issues discovered late in the process can cause feedback loops to earlier development stages and other parts of the project due to the complex and emergent aspects of CoPS, causing significant delays, costs and re-work to the project, good project management practices are especially important in CoPS development (Nightingale, 2000). Davies & Brady (2000) argue that "project capabilities" should alongside strategic and functional capabilities be treated as essential to organisations supplying CoPS. Project capabilities are especially important in the bidding phase, to successfully win contracts, as well as in executing the project after a successful bid. During the bid phase some necessary project capabilities are to gather requirements from customers, cost estimation, project scheduling and risk management. During the execution organisations also need to allocate resources, integrate different organisational functions and team management (Davies & Brady, 2000).

Being able to manage the uncertainty that CoPS entails is a capability that any organisation developing and supplying CoPS need to develop. Nightingale (2000) outlines six different factors of uncertainty in CoPS that affect the project and cause costly feedback loops, and ways of mitigating these uncertainties.

i) Technological traditions established: If the organization has established design processes that they are able to re-use, they are better prepared to handle uncertainty.
ii) Intrinsic uncertainty of the technology: By fully understanding the technology and what effects changes will have, as well as using well established technology, an organization can decrease the uncertainty.
iii) Complexity of the product: The complexity increases with the number of subcomponents included, as well as increasing the likelihood for emergent properties, putting pressure on the organization to have a good system for managing their "work in progress".

iv) Systemic relationships between subsystems: The amount of interdependencies of different subsystems will increase the complexity and uncertainty. Changes to one part of the project can therefore cause feedback loops throughout the system. Being able to perform analysis on the system throughout the development and simulations can decrease the uncertainty.

v) Fixed and unfixed problem: Changes to the problem caused by emergent properties, regulatory changes or customer specifications all increase the uncertainty and adds costs. Having good project management practices in place, contingency planning and risk management can all help in decreasing the uncertainty.

vi) Organizational rigidities: Using inappropriate organizational structures for the project, culture and physical distance can increase the uncertainty of the development. Having clear communication channels both horizontal, and especially vertically is essential to decrease the uncertainty.

Since developing CoPS require knowledge from several different domains, it is important that organizations develop their capabilities in building alliances with other organizations (Hardstone, 2004). Due to the amount of domains as well as stakeholders involved in supplying CoPS, two frequently mentioned capabilities that are core to any organization developing CoPS are systems engineering and systems integration where organizations are able to combine the knowledge and subcomponents/subsystems into the final CoPS (Hobday, 1998) (Hardstone, 2004) (Nightingale, 2000).

3.5. Model-based systems engineering (MBSE) methodology

Model-based systems engineering is based on the principle of using a common project model, or system model, throughout the development of the system (Ogren, 2000) (Ramos, et al., 2012). The system model is constructed by connecting different sub models. These sub models should contain all the relevant information about the systems and be an accurate representation of the requirements, functions and structure of the system (Ramos, et al., 2012).

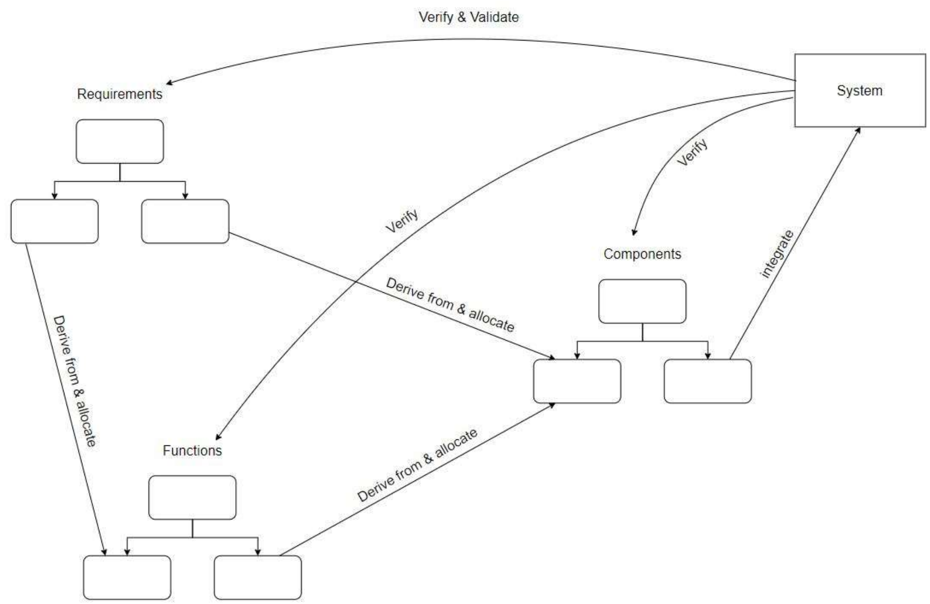

Figure 3. Conceptual representation of the MBSE development process

When moving from a document-based approach, towards an MBSE approach, there are four phases, or levels of advancement that an organization can choose to incorporate MBSE methods (Brown. B in (Holland, 2015)).

i) The first level is when no system models are used, and documentations are free form and purpose created for each instance.
ii) The second level, system models are used, and diagrams are drawn from them to support information from documents.
iii) The third level, system models are the primary source of information and substantial information is drawn from the model to create documentation.
iv) The fourth level, documentation is created automatically using information generated from the system model. Barely any editing is done.

3.5.1. The model

A model can be described as an abstraction or representation of an element of the physical world. The element can represent for example a system, a process, a product, or a phenomenon. The model is often used to describe certain aspects of these elements such as a function or geometry (INCOSE, 2015). Models should be developed for a specific purpose and to meet one or several established stakeholders needs and/or requirements. However, no one model can satisfy all the questions posed by the different stakeholders (Madni & Sievers, 2018). Inherently, no model can represent the physical world with complete accuracy. There is always some degree of uncertainty in the models used (Holland, 2015) (Madni & Sievers, 2018).

When working in a model-based approach it is necessary to establish the scope of a model so that it fits with the models purpose in addressing the relevant stakeholder needs and requirements (INCOSE, 2015). When a model can accurately address the purpose and the stakeholder questions imposed on it, it is said that the model is "fit for purpose" (Madni & Sievers, 2018).

There are no specific rules on what type of model or set of models to be used to answer stakeholder questions (Sargent, 2015). Rather, the choice of type of models depend on the purpose of the use of the models, the characteristics of the system of interest and on what level of accuracy is needed (INCOSE, 2015), as well as, the resources that are available (Sargent, 2015). While there exists many different definitions and ways of sorting types of models, INCOSE (2015), presents a well-structured and relatively comprehensive taxonomy of model types.

i) Physical model: A simplified model of the physical system or part of the system such as a wind tunnel or a prototype of the system.
ii) Abstract model: An abstract model can be expresses in many various ways consisting of different informal and formal models. An abstract model acts as a representation of the system of interest or system element and can vary in degrees of how concrete the model is.
iii) Informal model: An informal model can simply be a representation using simple drawings or be in text form. Although this can be useful, it must have a high degree of relevance so that it may be useful for the abstract model.
iv) Geometric model: A geometrical model is used to show the geometric properties of the system of element and/or the connections in the system.
v) Quantitative model (mathematical model): A quantitative model is based on mathematics to represent the system or parts of it to acquire a numeric result.
vi) Logical model (conceptual model): A logical model, or conceptual model, is used to represent the relationships and interconnections between different parts in the system. The representation could be of for example, function, processes, or activities. The model often consists of diagrams, tables, graphs, etc.

3.5.2. Model-based systems engineering and life-cycle stages

Using a set MBSE methodology of processes, methods and tools can greatly reduce the risks in the system development project and increase the likelihood that the that the developed system fulfils all the different stakeholder requirements (INCOSE, 2015). While many organizations develop their own MBSE methodology and life cycle approach to develop systems, most of them are based in one for three life cycle models, the Waterfall model, the "Vee" model or the Spiral model (Estefan, 2008). While the different parts in the methodologies vary in the sequence and amount of iterations each step is done, they mostly consist of the same stages (INCOSE, 2015).

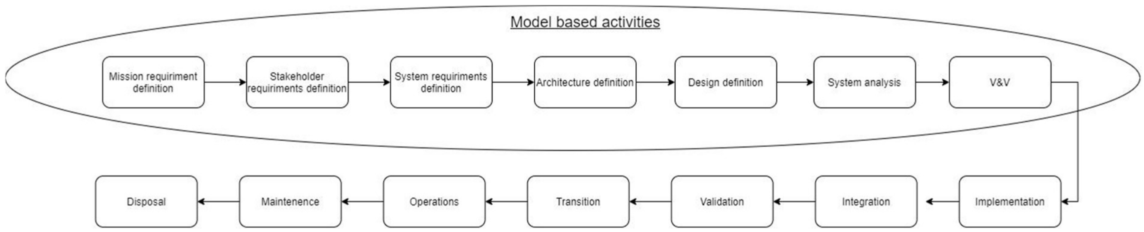

Figure 4. Life cycle stages of MBSE methodologies

The international standard ISO 15288 outlines some generic stages in the life cycle development of a system (concept, development, production, utilization/support and retirement) (INCOSE, 2015). While the focus from an MBSE perspective is on the concept and development stage, the developed models need to consider the other stages as well, such as how the system will operate or how the end of life process will look like.

In the concept stage, a preliminary concept is first developed from the system requirements that are derived from the business needs and mission requirements and the stakeholder requirements (INCOSE, 2015). Models helps to synthesize, evaluate alternate concepts from the preliminary concept and aid in a clear definition of the system requirements (Freidenthal, et al., 2014). The system attributes can be linked to the objects in the model which helps in efficient management of requirements (Ogren, 2000). The parameters that are critical to the system can effectively be communicated through the models (INCOSE, 2015). Models help in the validation of the system requirements against the stakeholder needs, acting as a checkpoint before proceeding to the next set of life cycle activities (Freidenthal, et al., 2014).

The design development stage uses the outputs of the concept stage to create the system architecture and design (INCOSE, 2015). The system architecture is vital as it is a formal representation from which the logical, behavioural, structural and other related representations are derived (Madni & Sievers, 2018). In the design stage models aid in converting the system requirements to functional and then the component level (Freidenthal, et al., 2014). Both top down and bottom up approaches can be applied in the design development phase, to distribute requirements to the objects and to find reusable objects for the requirements respectively (Ogren, 2000). A variety of models can be used to represent different aspects of the system design based on the need to cover both functional (e.g. Interface, functionality, performance and physical requirement) as well as non-functional requirements (e.g. Reliability, maintainability, safety and security) (Freidenthal, et al., 2014).

Verification and validation (V&V) are important steps in MBSE. Model verification and validation help to eliminate the flaws in the system and ensures that the system meets the external requirements and is close to reality (Madni & Sievers, 2018). In the system integration and verification stage, models support the hardware and software integration and in the test phase, models can help to define various test cases (INCOSE, 2015) (Ogren, 2000). Lower level components of both these categories are integrated to the higher-level system design which in turn aids the verification (Freidenthal, et al., 2014). Validation ensures that the system modelled is dependable (Madni & Sievers, 2018).

According to Madni & Sievers (2018), there are various approaches to model-based V&V as mentioned below.

i) Model appraisal: The domain experts from various disciplines evaluate the model. This method improves the quality of the design but is expensive.
ii) Guided modelling: This aids the designers to effectively model the system. It uses pattern based method, which uses previously validated patterns as hints to design; template based methods, which use pre-verified information as a starting point; feedback enabled approaches, which use verified and validated standardized models and lessons learned from previous projects as a tool.
iii) Simulation: It executes the model in a cost-effective way against the operating conditions, to understand the behaviour and take corrective measures. This is used especially if the other means of testing is hazardous or inaccessible for humans making the tests expensive.
iv) Formal proof: This uses mathematical / logical methods to verify the system against the specifications. Model checking and theorem proving are two formal methods of system verification.
v) Digital twin and digital thread: They are the digital equivalent of the system that can be used for verification, where digital twin is an accurate representation of the system that can be used throughout its life cycle and digital thread is a framework for sharing information among multiple stakeholders in the development activity.

The factors affecting the choice of MBSE based V&V method are the domain, the design and development preferences and availability of tools. V&V method is indispensable as it ensures that the different stakeholder requirements are met, and that the system fulfils its purpose. (Madni & Sievers, 2018)

The other life cycle activities such as training, maintenance / diagnostics, interactive simulations are also assisted by models depending on the requirement. By playing a crucial role in almost all the life cycle stages, models help in the system evolution by capturing, applying, and reusing knowledge. This helps the organization in knowledge management and enhancement of its competitiveness in a changing environment. (Freidenthal, et al., 2014)

3.6. Data driven methods

Data driven methods make use of empirical models to derive relationships between system variables from a large set of data (Mosallam, et al., 2015) (Villarejo, et al., 2016). The relationships are modelled by applying methods such as machine learning and 'computationally intelligent algorithms' to the complex datasets (Mount, et al., 2016). These methods have limited dependency on the domain specific background knowledge and hence can be useful at times when the hypothetical knowledge is limited due to increased system complexity (How, et al., 2019) (Mount, et al., 2016).

Drawing from software engineering, Geffner (2018) classifies programs in either Learners or Solvers. In aggregated terms, Learners are model-free and utilizes data or experience to achieve the output. This data driven method is characterized by a slow training period but are fast after the learning period. Solvers are model-based and through a model automatically achieve an output. This model-based method is more general and can solve any problem if it fits the model. As Solvers use models, they require no preparation but are slower than Learners in achieving the output.

Data driven methods in complex systems can roughly be categorized in either data driven modelling, data driven monitoring and fault diagnostics, or data driven control and optimization. One big benefit of using data driven methods is that it requires no previous information about the process. Data driven methods are instead based on signal processing, and data analytics. (Gao, et al., 2013)

3.6.1. Big data analytics

The term big data was initially made to capture the emergence of the vast amount of data being created from an incredibly diverse set of sources making it hard to handle for existing structures, with the amount of data approximately doubling every two years (Hu, et al., 2014). The definition of big data used in this thesis is adopted from Mauro et al., (2016, pg. 131) as: "Big data is the information asset characterised by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value". The area of big data can be categorized into one of four main themes; Information, technology, methods or impact (Hu, et al., 2014).

i) Information: information is structured data and is what drives big data. For organizations, information can be turned into knowledge and used to create value as per the data-information-knowledge-wisdom hierarchy.
ii) Technology: It is a necessary enabler to be able to gather and analyze big data. Big data puts a lot of pressure on technical systems due to the speeds in which it needs to be processes, and the amount of data that needs to be stored.
iii) Method: The usual statistical methods used to process data is not enough to handle the amounts of data that big data entails, rather more complex methods with examples like, neural networks, regression models and cluster analysis are needed. The method change will also require a cultural change in the organization with a focus on proper data management and implementing data in the decision making.
iv) Impact: big data is already having a large impact on the society and is continuously being adopted in an increasing number of domains. Big data is also impacting organizations internally where they must question their processes so that they are able to utilize the data in the most impactful way.

It should be noted that big data and system specific data are different. System specific data can be data generated from models, simulation, tests and operational conditions that are related to the system, often used for taking decisions related to alternatives and to verify if the system model meets the user requirements (INCOSE, 2015) (Madni & Sievers, 2018). Big data is much larger in comparison and need specific methods such as data analytics to gather system specific insights from it.

3.6.2. Internet of things (IoT)

Internet of things refers to the communication network between objects in an environment enabled by the information technology, aimed at harnessing the information/data from these objects for various purposes such as process enhancement, productivity improvement, decision making, trend prediction, pattern finding etc. (Marjani, et al., 2017) (Gubbi, et al., 2013) (Lee, et al., 2015). The development of sensors in the recent years along with the rise of technologies like digital technology, advanced telecommunication devices, wireless sensor networks have enabled the monitoring of a wide variety of applications (Gubbi, et al., 2013). The information collected in this manner is voluminous and can be analysed to gain insights which aid data driven decisions and can also be used for creating a common pool of information to trigger new applications (Marjani, et al., 2017) (Gubbi, et al., 2013).The applications of IoT cuts across industries such as home, transport, healthcare, defence, agriculture, enterprise, mobile to name a few (Gubbi, et al., 2013).

The transmission and storage of data is a critical part of IoT since the amount of data generated is huge. Intelligent storage and retrieval of data in a centralized manner enabled by cloud storage technologies will be prevalent in the industry (Gubbi, et al., 2013). As the sensors gather all different types of data, the nature of IoT data is different as compared to the other types of bigdata and hence suitable processes are needed to handle IoT data to eliminate some of the issues associated with them (Marjani, et al., 2017).

Big data analytics is crucial for supporting IoT as the structured, unstructured, and semi-structured data obtained from the different sources need to be transformed into homogenized data that can be analysed and interpreted. Analytical tools employ different algorithms and methods to achieve this, such as classification, clustering, association rule mining and prediction categories. (Marjani, et al., 2017)

Classification, a supervised method, uses existing knowledge to train the system to handle the data and categorise them into groups (Marjani, et al., 2017). Deep learning (DL) is another supervised method where by minimizing the error function linked to the training data set, the system gains knowledge about the paraments and utilizes it when a bigger data set is given (Geffner, 2018). Clustering is an unsupervised method which categorizes data based on their distinct features (Marjani, et al., 2017). This is comparable to deep reinforcement learning (DRL), where the method learns on its own, by minimizing the error function which is not based on a training data set but based on the successive data sets given to it (Geffner, 2018). Association rule mining works by build meaningful relationships among different data types for predicting trends, behaviour and demand whereas prediction categories use historical data as a training data set to find out patterns and trends (Marjani, et al., 2017).

3.6.3. Digital Twin

Madni et al. (2019) discusses the benefits of digital twin supported by internet of things (IoT) for the system developers in the model-based system engineering context. Digital twins aim to integrate the physical and virtual systems to aid real time monitoring of systems, collection of data for various developmental and maintenance purposes, reduce downtime by preventive interventions explore new business opportunities and future system upgrades. Digital twins can bring down the cost of system verification and testing by using the information collected from the physical twin. With the advancements in IoT, the cost of implementing digital twin has come down, making it a viable option. With the support of IoT, a digital thread, which is an information chain, connects digital twin to its physical twin, throughout its life cycle, capturing all the necessary data. The models in digital twin gets updated accordingly. The relationship between the twins continue even after the product sales, throughout the service life of the product. According to Madni et al. (2019), a digital twin is considerably different from a CAD model, as it represents a specific instance of the system, reflecting the performance of the physical twin. It also maintains the traceability of the physical twin during the life cycle phases and its age, through the operational data. The digital twin can be used for many different purposes including validation of the system model, predicting system changes, provide decision support, and discover the new possibilities in application. (Madni, et al., 2019)

4. Descriptive analysis

The descriptive analysis aims at giving an overview of the selection of articles including the authors, distribution over time, where the articles where published and the citation count for the entire category, as well as for sub-categories. The section will also give a description of the main findings of the literature, giving a more through view of the different areas.

4.1. Complex product and systems

For the CoPS part of the review, 25 different articles where included. In this selection of articles, 44 unique authors contributed with 5 authors being included in more than one article. The articles where published in 15 unique journals or conference proceedings with 5 journals contributing with more than one publication. The selection was distributed between 1997 and 2019 according to figure 5. In total, the selection had a citation count of 3604 citations over all articles.

The distribution of articles in the subthemes of Complex products and Systems are presented in table 4 with the number of articles in each section and number of total citations.

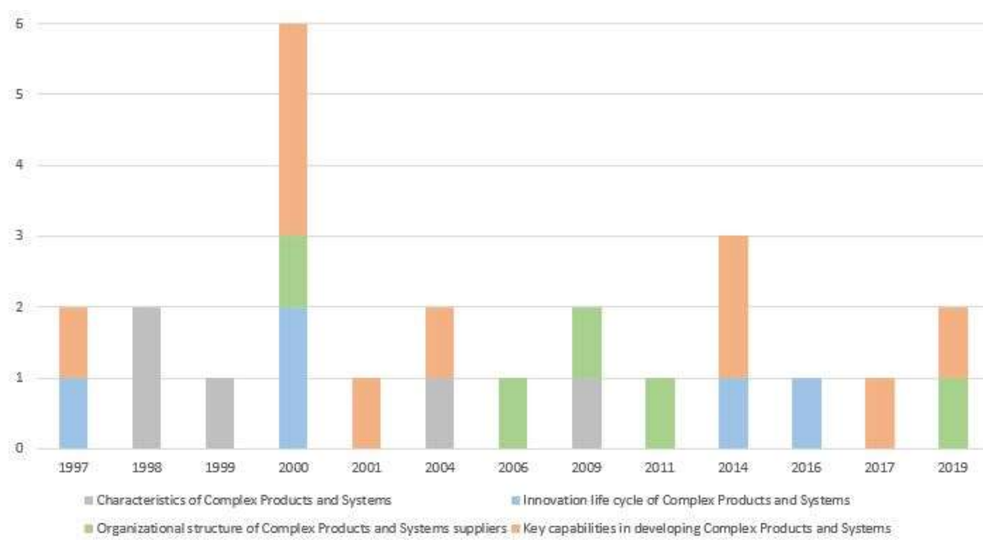

Figure 5. Distribution of articles in subthemes of CoPS.

Subtheme	Number of publications	Number of citations
Characteristics of Complex products and systems	5 articles	959 citations
Innovation life cycle of Complex products and systems	5 articles	613 citations
Organizational structure in Complex products and systems suppliers	5 articles	792 citations
Key capabilities of developing Complex products and systems	10 articles	1240 citations

Table 4. Article and citation count in subthemes of CoPS.

4.1.1. Characteristics of Complex Products and Systems

Hobday (1998) in his conceptual paper outlined many of the characteristics of CoPS which still holds true in industries today. Hobday (1998) found that a key contributor to the complexity of CoPS is its hierarchical nature and product architecture. Due to the size and complexity of CoPS, there tend to be a very large amount of design alternatives of the system architecture. The difficulty of these decisions is only made harder with increasing amounts of custom-made components and sub-systems in the system, which is a common characteristic of CoPS. In addition, feedback loops from later to earlier stages requiring architectural changes is a reoccurring issue in CoPS development and what could be viewed as an incremental change to the system at first, could later on show to have significant impact on the components functions (Hobday, 1998). Hobday (1998) also argues that due to that CoPS tend to be very high cost and only developed in unit or batch size, performing tests and experiments on these systems may be extremely costly, and in some cases impossible. Not being able to test the designs only makes it harder to identify issues and potential feedback loops in the system. For this reason, step-by-step continuous learning throughout the development process is an important aspect in CoPS industries.

Owing to the size, complexity and importance of many CoPS, there is often a high number of stakeholders involved in the decisions made in the development. These stakeholders commonly involve major users, suppliers and regulatory bodies which all can affect the innovation path of the CoPS. Having efficient coordination between stakeholders is often key in establishing design paths for the architecture and during development since it is uncommon that one firm internally have the necessary span of control of the project, or the breadth of knowledge required. (Hobday, 1998)

Hansen & Rush (1998) in their multiple case studies found four general problem areas, or "hotspots", where many of the problems seem to originate from the characteristics in technical uncertainty and coordination established by Hobday (1998). The first area was dependency on the suppliers and difficulty in the procurement systems, which proved to be prevalent in most projects. This could be due to poor performance in the reporting system, or lack of control mechanisms. Overdependence on key suppliers could also lead to delays in integration and delays from suppliers could in some cases cause delays for the entire projects. The second problem area was in technical uncertainty/difficulties. Some causes for this could be that possible technology reuse from previous generations is not communicated or that short cuts were made to satisfy short term projects constraints but adding uncertainty and risk in the long term. The third area is organisational and project structure, where the structure needs to facilitate knowledge transfer within the projects, as well as, with customers and between other projects, and facilitate learning. The final problem area is management of requirements capture. Requirements capture is often initially done under significant time pressure. Along with "knock on effects" from new technology, and changes in client's needs, the management of requirements can often lead to difficulties for the project.

Hobday & Rush (1999) suggest that an underlying characteristic to some of these issues may be the fast pace of technological change. This can cause capability gaps both within teams, as well as between firms. The complexity of CoPS also makes it hard to transfer knowledge from one sector to another, making it difficult to have set best practice and lead to very low learning across sectors. The fact that CoPS tend to be developed in projects also causes problems in learning between projects. This lack of project-to-project learning is a key reason for the variety in performance of CoPS projects (Hobday & Rush, 1999).

The way that development and innovation in CoPS industries differs from traditional mass-produced industries where development happen in a more linear way. In CoPS actors work together in a "web of innovation" including users, buyers, suppliers, regulators, among others, who collaborates to develop and improve the systems. Central to this innovation network is the systems integrator (Hobday & Rush, 1999). Rutten et al. (2009) define the term Systems integrator as an organization that through systems integration adds value in project-based industries. There are, according to Rutten et al. (2009), two main tasks to be performed by the systems integrator. First, to establish the network of organizations that will be involved in the development of the CoPS. Second, to coordinate and delegate the work of these actors. Through these interorganizational networks, the system integrator collaboratively achieves innovation with the actors by developing new components and/or new ways of linking components.

How large of an impact of CoPS industries have on a nation is very difficult to measure and cannot be done through conventional methods since CoPS span through multiple sectors and due to the project network structures involved in CoPS innovation, measuring economic contributions is difficult (Hobday & Rush, 1999). However, Acha et al. (2004) did attempt to quantify the impact of CoPS industries in manufacturing and construction in the UK. The research found that, at that time, CoPS accounted for 19% of overall production and total gross value of the UK economy as well as 21% of employment in manufacturing and construction. Acha et al. (2004) also found that CoPS tend to spend more on IT than other organisations, while they will spend less on advertisement.

The main findings of this section are summarised in the table below.

Key findings
• CoPS have a hierarchical system architecture. • CoPS face technical uncertainties and emergent behavior. • Feedback loops from later to earlier stages can impact system development. • Capturing requirements is challenging. • High degree of coordination with stakeholders.

Table 5. Key findings from characteristics CoPS

4.1.2. Innovation life cycle of Complex Products and Systems

Bonaccorsi & Giuri (2000) presents that there are two classes of industries that does not conform to the traditional industry life cycle theory and therefore will not show the pattern of shakeouts. Together they represent all previously known cases on non-shakeout. Class I industries, where shakeouts do not occur because of the vertical separation between process or product research and development, and manufacturing. In these industries incumbents cannot build up large barriers of entry due to R&D investments, and large investments in process technologies is risky since they do not have control of the product-process connection. Class II Industries instead suffer from a lack of economy of scale, combined with customers with very specific needs. In these industries suppliers need a high amount of knowledge, which often tends to be tacit, to serve their customers. For such organisations it is often preferable to acquire specialised knowledge and focus on serving one marketing segment. However, Bonaccorsi & Giuri (2000) showed that organisations operation in class II which are in "violations of increasing returns", may serve several segments by using a multi brand, multi divisional structure and if the supplier's market is highly independent, for example by subcontracting or purchasing strategies. In their case of the turboprop industry, there was a stable coexistence between a market leading generalist serving several segments, and specialised organisations. Bonaccorsi & Giuri (2000) argues that systemic industries, such as single or batch production complex products and system fits well in the class II discontinuity which suffers from a lack of increasing returns.

Similar to that systemic industries such as CoPS does not show a pattern of shakeout, disruptive innovations do not cause the same changes to the industry as in mass producing industries. According to Dedehayir et al. (2014), the disruptive innovation in the context of CoPS have unique characteristics, which is also confirmed through a case study. Early adopters, policy makers and regulators play an important role in paving way for disruptive changes. CoPS are produced in limited volume based on specific customer need. Hence, the chances of a firm missing the performance mark is lower in CoPS industries. It is also difficult to interchange technology of sub systems within a CoPS network. Since it is difficult to find niche market for the nurturing of technology, the new technologies often must compete with the incumbent technologies directly. This means that the new technologies need to be superior to survive in CoPS marketplaces. As the disruptive technology tends to be superior, it increases the complexity of the further, making it costlier. However, in the absence of a clear superiority, the disruptive technology tends to co-exist with the incumbent technology. (Dedehayir, et al., 2014)

Hobday et al. (2000) in their editorial paper emphasises the length of the innovation life cycle of CoPS. According to Hobday et al. (2000) it is common for the product life cycle to last for decades and even the decision from stakeholders to invest may take several months to years. A characteristic of CoPS, as compared to the conventional model, is that innovation to the system is on-going, even after implementation of the system with continued updates and upgrades throughout the life cycle (Hobday, et al., 2000).

Davies (1997) argues through his case study on the telecommunications industry, that CoPS industries, which are engineering intensive, low volume products and system that serve individual business, go through two life cycle phases, the architectural phase and the new product generation phase. The architectural phase is characterized by a high amount of architectural innovation. Architectural innovation concerns the function of the components and subsystems, and their interconnections. The architectural phase concludes when the core components are selected and there is agreement on technical standards. The new product generation phase then starts once the new product system architecture is commercialised. The new product generation phase is instead characterised by a high amount of component and systemic innovation. Component innovations are technical changes to components in the system where the change only affects that component and no modification to other parts are necessary. Systemic innovations are technical changes to a component or its functions that also affects other parts of the system, requiring changes to other parts of the system as well. According to Davies (1997) one of the largest challenges for suppliers of CoPS is managing the life cycles of individual components. While the system may have a lifespan of several decades, technology in individual components may have be changed, or become obsolete within years. Davies (1997) also suggest that one limitation of this framework is that it may only apply to infrastructure networks such as telecommunication or transportation systems and not to standalone products such as aircrafts or smart buildings. However, Huenteler et al. (2016) in their study on technological change in the life cycle of solar photovoltaics and wind power, found that the life cycle of wind turbine technology, which is arguably a more standalone product than a network, more closely followed Davies (1997) model than the traditional mass production industry life cycle. Huenteler et al. (2016) does however argue that products may lie on a scale of industrial life cycles where the traditional mass production industry life cycle, and the CoPS life cycle by Davies (1997) are two extremes on the scale which is based on complexity of the product architecture and scale of production.

The main findings of this section are summarised in the table below.

Key findings
• Innovation can be architectural / systemic / component in nature. • Innovation continues even after system becomes operational through updates and upgrades. • Industrial phases consist of architectural and new product development phases. • Innovation is a collaborative effort with stakeholder involvement. • Incumbent technologies can coexist with disruptive technologies.

Table 6. Key findings from innovation life cycle of CoPS

4.1.3. Organisational structure of Complex Products and Systems suppliers

Hobday (2000) compared two companies developing CoPS, one company using a pure project-based structure (PBO) and one company using a matrix structure, to look at the process differences, advantages, and disadvantages of the structures in the development of CoPS. Hobday (2000) contrasts the PBO with more established structures used in mass produced markets, such as the matrix structure, and claims that the PBO structure is effective in dealing with emergent issues, respond to changes in requirements, respond to uncertainty, etc. The reason being PBO having a concurrent approach that is external focused as compared to the inward focused approach of functional/matrix structure. This makes it effective form in developing CoPS and fits best in large, high risk projects where all the resources of the firm revolve around that single project.

However, Hobday (2000) identifies that the PBO structure is less effective in areas such as coordination of resources and capabilities across projects, routine production or engineering tasks and project to project learning. The challenge in choice of structure in the developing CoPS is to balance the need for the current project, and to develop the organisations capabilities for future projects and markets opportunities. To overcome this, the company case showed how they took a step back from the PBO structure to a project led organisation, where the project manager is still in power but there is now task coordination in functional units to capture the learnings and gain technical expertise.

According to Hobday (2000), PBO uses project as a vehicle for integrating and coordinating key business functions covering the broad spectrum - R&D, new product development, engineering, production, marketing, and finance – of activities spanning the organizational boundaries in CoPS. They are also used for exploring new strategic opportunities in terms of technology development, market expansion as well as for the revival of the organization (Davies, et al., 2011).

According to Davies et al. (2011), the ability to provide project-based customer centric solutions requires organizations to develop capabilities in areas such as system integration, operational services, finance, and consulting. This can be achieved through changing the organizational structure, having new performance parameters, and engaging in longer life cycle of projects. A long-term strategic engagement with the network players is required, to solve customers' problems. Right organization structures are vital for organizations to ensure horizontal collaboration across the network of CoPS. Concepts like co-located organizations for enhancing cooperation can be rewarding.

Roehrich et al. (2019), focuses on the impact of organizational structure on the innovation in CoPS, in the context of the inter organizational functioning and explores the effectiveness of the concept of integrated project teams (IPT), a form of management innovation. IPT is characterized by cohesiveness, cross functional integration, relationship focus and strategic project related activities. The increased product service integration demands management innovations to support the related activities. IPT is a suitable form of collaboration between organizations bringing the knowledge and expertise on the same table resulting in integration, trust building and improved relationships. By combining complementary skills from both the organizations, integrated solutions could be developed in an efficient manner.

Moddy and Dodgson (2006) analyses a case of technology transfer in a small satellite project and gives insights on the issues faced in the project. The authors found two major areas of issues in the project. Varied and changing goals and objectives, where a large number of stakeholders had high ambitions, but low engagement, and, external relationships, where the communication and trust with the physically distant supplier started breaking down as well as issues. Issues also emerged between other actors due to a difference in management approaches, leading to a breakdown in collaboration.

According to Moddy and Dogson (2006), one of the main contributions to the eventual success of the project was how the organisational structure evolved as the project proceeded. The structured moved from an initially more relaxed and organic structure to a more hierarchical one. This changing structure provided the flexibility to deal with new and changing conditions while the risk of feedback loops was high, as well as, providing safety and division of responsibility towards the critical period of project delivery. By changing the organizational structure from organic to hierarchical, the project succeeded in benefitting from both flexibility and discipline as per the requirement.

Geraldi (2009) explored how multi project firms deal with chaos and order because of the environmental change, through a study carried out in a large CoPS developing organization. Organizations often need to adapt to order and chaos because of the demand from multi project environment and associated heterogeneity. Even though order and chaos are diametrically opposite, the same organization often faces them due to the changing environment. Geraldi (2009) identifies two strategies that organizations adopt: i) organizations respond with flexible structures when faced with inflexible customer demands ii) they respond with internal rigidities to compensate the external chaos. This is how the balancing between chaos and order in response to external changes take shape. A conceptual model developed by Geraldi (2009) which consists of two dimensions, complexity of project portfolio and flexibility of organization structure, suggests the ideal path for organizational transformation.

The main findings of this section are summarised in the table below.

Key findings
• Project led organizations are suitable for CoPS. • Integrated / co-located project teams can be beneficial. • Organizational structure evolves throughout the project.

Table 7. Key findings from organizational structure of CoPS

4.1.4. Key capabilities in developing Complex Products and Systems

Park & Kim (2014) through their analysis of CoPS literature to identify all the main capabilities of CoPS development. The study found that there are three comprehensive capabilities in CoPS development needed. Close networking among actors, broad, deep and integrated skills and, the ability to leverage institutions and policies. Park & Kim (2014) validate their findings through a case study of Korean e-government, showing that Korea utilized these capabilities to successfully implement their e-government system.

Davies & Brady (2000) in their multiple case study found that in addition to strategic and functional capabilities, organizations that operate in CoPS industries also require project capabilities. Project capabilities are essential in both product development and implementation, which within CoPS are both project-based activities. Project capabilities are especially important in the preparation for a bid as well as in the execution of the project. Some examples of project capabilities include, requirements gathering, risk management, defining conceptual components, resource allocation integrating organisational functions etc. The way that organizations in CoPS industries develop their capabilities is through what Davies & Brady (2000) call economics of repetition, where organizations focus on their project effectiveness to be able to, more effectively execute, a greater number of similar projects. When learning from repetition organizations go through four phases. In the first phase, new knowledge and routines are created and the organization must acquire new functional capabilities and project capabilities. In the second phase organizations transfer this new knowledge and experience to other current or upcoming bids where more knowledge is captured and again applied to succeeding bids. In the third phase the organization implements the lessons learned from the projects into the functional organization, transforming their process and embedding the learning into the routines. In the final stage organizations may create completely new business units dedicated to specific areas such as project management and system integration which utilizes the learnings from the previous phases.

Gann and Salter (2000) observed through their study of the firms producing complex systems in the construction industry, that project groups often operate at the boundaries of the firm, working with other network players, creating value and generating profits. Integrating the continuous business process and projects activities becomes important in this context. Business activities result in creation of routines that aid standardization, process improvements and economies of scale while projects usually deal with non-routine activities, limiting the scope of benefits of routinization.

According to Gann and Salter (2000), effective integration of business process and projects is essential for the firm's ability to manage project portfolio and increase the competitive advantage. The learnings from the project need to be incorporated into the business process through feedback loops, to achieve synergy between the two. The increased demand for services related to the various aspects of the product is an opportunity for organizations to enhance the value proposition. These can include financial structuring, consultancy, customer support, training, and facilities management. The services tend to be offered between supply networks and project-based firms as well as between project-based firms and customers. Service enhanced capabilities are the result of successful linkage of project learnings and the continuous business process.

Nightingale (2000) identifies six main areas of uncertainty in CoPS development, established technological traditions, using uncertain technologies, the complexity of the product, the systemic relationship between subsystems, changes in requirements and regulations, and organisational rigidities. The reason why these uncertainties are more prevalent in CoPS is due to the high potential of feedback loops from later to earlier phases of development which can cause significant delays. To decrease the uncertainties in the projects, organizations need to be able to match the design to the requirements, and make sure that the specifications for the design are correct. Organizations that can be flexible in their resource allocation should also be able to reduce uncertainty. Nightingale (2000) also gives some examples of practices that can help reduce uncertainty such as reuse of technology, using already established technologies, contingency planning and analysis throughout the design process.

According to Liu & Su (2014), during the R&D process organizations developing CoPS need a mix of technology and market orientation to eventually succeed in the project. A market-based approach will allow organizations to identify customers and their requirements. However, with too much focus on market-based activities, an organization may lose its ability to develop new products with novel and complex technology. To be able to stay innovative, organizations should, according to Liu & Su

(2014), adopt a market orientation with strong technological values. Balancing market pull with technology push to identify and satisfy customer needs, while still being able to innovate in new complex technologies.

Rush (1997), identified requirements gathering as a hotspot. Requirements are gathered under time pressure by a team which may not consist of senior technical members, communication between technical staff and customer may be limited. This result in a lack of clarity in capturing the needs. There could be requirements arising as the projects progress, incorporation of which could cause ripple effects on the project. It is important to maintain a continuity in the flow of information from the bid phase to the implementation phase. According to Rush (1997), Isolated efforts by firms cannot make the projects successful. Rather, an optimization of the whole network is needed to deliver CoPS projects where the network players are well integrated.

According to Naghizadeh et al. (2017), the case study of IR-150 aircraft development, a first of passenger aircrafts designed by Iranian firms, showed that CoPS projects face challenges in integration which are often context dependant. For example, experience of developed countries with CoPS may not be sufficient to ensure success of another CoPS development in a developing country. The structure of CoPS integration itself can be imagined as an integrated system where every part of it need to work effectively and efficiently to make the project a success. Integration structure changes throughout the development phase due to the inclusion of more network players and hence it can be observed that integration structure is dependent on project life cycle.

Zhang and Igel (2001) studied the evolution of stored program control (SPC) switch manufacturing industry in China, as it is an emerging CoPS industry. It gave insight into the current state of strategies, industry structure, product development and innovation. The SPC switches and the industry did not fully demonstrate CoPS characteristics and the organizations did not adapt many strategies that are commonly adopted by CoPS developing firms. Zhang and Igel (2001) foresee that as the market gets more complex, SPC switch manufacturers will need to adapt more of CoPS strategies such as development of potential suppliers, integration of inhouse research with external parties, competence building for external networked players, horizontal management style with employee empowerment, improving innovativeness of users etc. According to Zhang and Igel (2001), as the external environment becomes complex and challenging, the emerging CoPS developing firms will have to adopt new capabilities.

Hardstone (2004) found that, along with traditional systems capabilities such as project management and systems integration, alliance building, and collaboration was especially important in systemic industries. When new technology emerged, market capabilities seemed to become even more important for organisations. With the variety and complexity in CoPS, the development projects seem to support a diverse set of capabilities, structures and strategies. According to Hardstone (2004), core capabilities of CoPS organisations do not act as rigidities and does not hinder organizations from adopting new technology as it may in commodity markets. Due to already established market connections and built up capabilities, it is more likely that incumbent firms, rather than new entrants will lead the development of the new technology (Hardstone, 2004).

Lehtinen et al. (2009) found that when working with many external stakeholders, engagement can be crucial, however, stakeholder disengagement might be just as important. Organizations should therefore not follow a strategy of being fully transparent or closed off, rather a strategy of calculated engagement and disengagement. Disengagement towards external stakeholders showed to be more effective in earlier phases where uncertainty in the project is higher and increase engagement as parts of the projects are fixed. Lehtinen et al. (2019) also find that when internal stakeholders consider engaging with external stakeholders, they should do so with a systemic view of what is optimal for the outcome of the CoPS rather than what is optimal for the individual organization.

The main findings of this section are summarised in the table below.

Key findings
• System integration capability and stakeholder management are critical. • Economics of repetition is an important capability. • Service enhanced capabilities can be developed by linking project learnings and business process. • Capabilities to balance between market pull and technological push is desirable.

Table 8. Key findings from capabilities in developing CoPS

4.2. Model-based systems engineering

For the MBSE part of the review, 23 different articles where included. In this selection of articles, 79 unique authors contributed with 11 authors being included in more than one article. The articles where published in 18 unique journals or conference proceedings with three journals with more than one publication. The selection was distributed between 2000 and 2019 according to figure 6. In total, the selection had a citation count of 1110 citations over all articles.

The distribution of articles in the subthemes of Model-based systems engineering are presented in table 5 with the number of articles in each section and number of total citations.

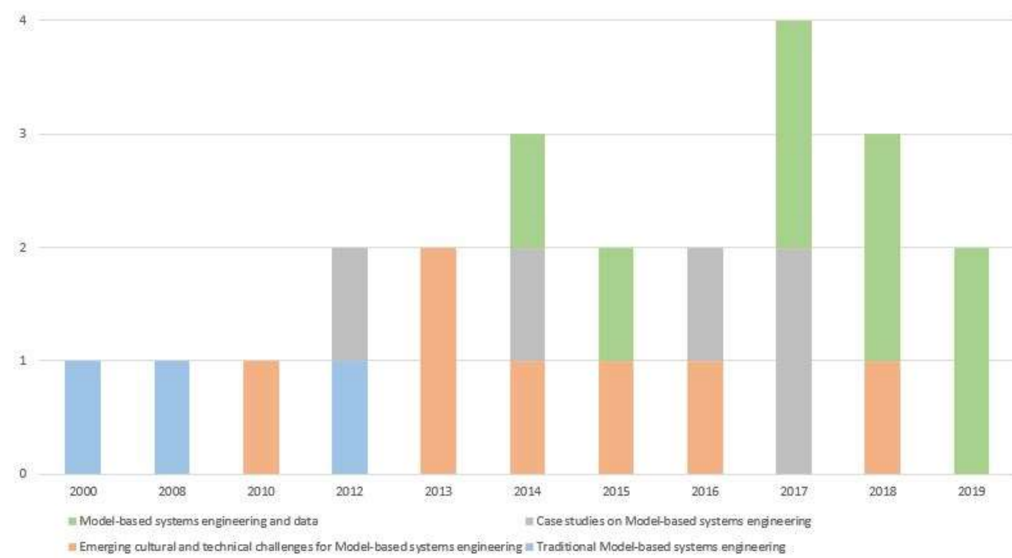

Figure 6. Distribution of articles in subthemes of MBSE.

Subtheme	Number of publications	Number of citations
Traditional Model-based systems engineering	3 articles	898 citations
Emerging cultural and technical challenges in Model-based systems engineering	7 articles	86 citations
Case studies on Model-based systems engineering	5 articles	20 citations
Model-based systems engineering and data	8 articles	106 citations

Table 9. Article and citation count in subthemes of MBSE.

4.2.1. Traditional Model-based systems engineering

Ramos et al. (2012) presents an overview of the current state of MBSE which is emerging as the suitable method for systems engineering. According to Ramos et al. (2012) MBSE is becoming a standard practise among system building industries and the trend shows that the system engineering is evolving towards a more unified practice in the future, in which MBSE plays a major role. According to Ramos et al. (2012), SE is an activity where the holistic, multidisciplinary, and complex visualisation is required. The interrelationship between the systems is very critical when it comes to SE, even if the individual silos are well developed.

According to Ogren (2000), "waterfall" or "big bang" methodologies have previously been common ways of developing system where stages and activities are planned out beforehand and sequentially executed. However, with complex systems it is not possible to plan and specify the system completely before the project, many activities are more concurrent rather than sequential as well it is very difficult to know the real and all the requirements in advance, sometimes even impossible. Ogren (2000) instead suggests that a more incremental approach should be used where stages are performed concurrently, especially "requirements management", "development" and "verification with test". This should according to Ogren (2000) be done through a central model based on design objects where requirements and test cases are connected to the objects. When the central system model's subsystems are integrated and functions in such a way that allows for a common understanding of the system, it can then be classified as a "common project model (Ogren, 2000). A common project model will according to Ogren (2000) increase the quality as well as the possibility for project success. The "common project model should evolve as the project goes on and could be seen as the projects backbone (Ogren, 2000).

Estefan (2008) claims that three of the most commonly used life cycle development models for large scale systems are the "waterfall" model, "spiral" model and the "vee" model where the "spiral" model is often used in software intensive projects and the "vee" model is commonly used in systems engineering. These life cycle development models provide a framework for organisations to build their methodology of processes, methods and tool is their project and domain specific environment.

Estefan (2008) goes on to emphasise that a central notion of model-based engineering is to elevate the models of the system to a governing role where the system model grows and get more detailed as the projects proceeds.

Estefan (2008) promotes the use of a "information model" in the development of systems through model-based engineering. The "information model" is according to Estefan (2008), a very important part in the development and allows stakeholders to view the information that is to be used in the development of the system and their relationship to each other. The "information model" should show the requirements of the project, where requirements may be decomposed into new requirements. The requirements should specify components in the system where components may be decomposed into other components. The model should show design alternatives which represent the components and satisfy the requirements and finally the models should represent the components and be able to execute the design alternatives (Estefan, 2008).

According to Ramos et al. (2012), the development of standards is critical for advancement of this field as it will establish benchmark practices in the different areas within the MBSE domain. There are different standards by groups such as standard committee of INCOSE, seventh committee of international organization for standardization (ISO), international electro-mechanical commission (IEC), institute of electronics and electrical engineers (IEEE) and Object management group (OMG). Apart from the core standards, there are other standards such as AF group that focuses on support systems architecture standards. The methodologies are also an emergent category where object-oriented SE methods (OOSEM), Harmony SE, Rational unified process for SE (RUPSE) and Object process methodology (OPM) represent some of the informal emerging methods. The informal and formal standards together represent the core set of norms that help the SE evolve. The modelling languages are an important subset of the modelling which help in moving across the different level of abstraction in the system. When it comes to languages, the SySML by OMG and Object process diagrams (OPD) and Object process language (OPL) by the OPM represent the present state of the art. The MBSE methodology offers many benefits owing to its integrated nature and is poised to expand its boundaries. But it has to overcome the cultural and technical challenges to be successfully deployed. According to Ramos et al. (2012), the future areas that need to be addressed will be more agile based MBSE methodologies, as well as the effective use of graphical modelling language to aid the collaboration of stakeholders. MBSE will have to evolve further to match the increased expectations from the stakeholders.

The main findings of this section are summarised in the table below.

Key findings
• MBSE is suitable for developing CoPS. • System development is more concurrent than sequential. • In MBSE, the system model has a governing role.

Table 10. Key findings from traditional MBSE

4.2.2. Emerging cultural and technical challenges for MBSE

Bonnet et al. (2015) capture the learnings from the MBSE implementation by Thales, an organization that focuses on engineering solutions for aerospace, defence, transportation, space, and security markets. Thales were developing Arcadia® and Capella®, which is an MBSE tool and its corresponding workbench for its clients. During a workshop conducted, it found that the obstacles and enablers to the implementation of MBSE are closely linked. According to Bonnet et al. (2015), cultural change is very important when it comes to adopting a new methodology like MBSE. Top management commitment is essential as the lack of it could create a risk for the other stakeholders. The existing tools, techniques and IT policies could slow down the implementation, if they are incompatible. The difficulty in measuring the return on investment (ROI) also could be an obstacle. Bonnet et al. (2015) found that the enablers in the MBSE deployment are the expected benefits arising out of the consistency of data, a good deployment strategy, coaching, and improved communication between the stakeholders. Bonnet et al. (2015) observe that the deployment of MBSE has its complexities and associated costs and must be carefully tackled to succeed in the transformation to MBSE.

As the systems get more complex and MBSE becomes the method of choice, managing the models and model life cycles in the context of the multidisciplinary environment becomes important. Fisher et al. (2014) argue that the convergence of engineering disciplines to create complex and smart cyber physical systems, decreased time to market, increased regulation, higher product quality requirements, etc., require a holistic system design methodology. This drives the system developing firms to build more multidisciplinary modelling and analysis techniques. According to Fisher et al. (2014), the integration of different modelling tools, advancement of standards and collaboration across the value chain are some of the important challenges that the industries are facing. As a result, the system data is spread across different platforms. Bajaj et al. (2016), claim that the main challenge for MBSE implementations is that they must deal with heterogenous models and the need for a model that serves as 'single source of truth'. In the development of intelligent systems, as the functionalities increase, the interactions become unmanageable due to complexities which makes it important to have effective cooperation and communication of inter disciplinary activities (Gausemeier, et al., 2013).

Bajaj et al. (2016) focuses on the needs that drive the next generation model-based systems engineering and tried to address them by developing an approach named 'MBSE++' which proposes an integrated model called Total System Model (TSM) that takes the role of a blueprint throughout the system life cycle. TSM uses SysML language as the supporting tool to bring together the heterogenous models. The TSM must be able to handle this heterogenous mix of tools and repositories and should be able to spot inconsistencies across them and fix them. Model transformation capability is needed to achieve this, where the version control and model-based connection between sources and target model elements are well developed, ensuring integrity of the TSM. TSM should be able to handle inter model connections (connection within a model) and intra model connections (connections between the models). Traceability is a key feature of TSM where user can review impacts, prioritise connections for exploration etc. According to Bajaj et al. (2016), an integrated model which is transparent despite the underlying heterogeneity is required and the system model needs to be connected to the project management. A concept like TSM shall facilitate functionalities like timeline views of the system, workflow management, verifying requirement against simulation, impact analysis, report publishing etc. Bajaj et al. (2016) recognizes some of the potential tensions associated with TSM. The need to have diverse tools to support the engineering breadth while being able to have a transparent and unified model is challenging. Similar is the case where the TSM need to have different type of connections for different disciplines while being able to create and visualize them through a unified framework. Ensuring traceability at a unified system level while still being able to prioritise impacts at individual connection level is also another challenge. (Bajaj, et al., 2016)

According to Fisher et al. (2014), model life cycle management (MLM) is set to be an important aspect in the further development of MBSE. MLM is about integrating and synchronising the information's related to the models that constitute the system design, where it has to handle multi-dimensional issues such as different tools, users in different geographies, tool revisions, maintenance of system consistency, validation of system design information etc. It also must handle the variants, product families, commonalities, and unique features to satisfy the various requirements. MLM face a big challenge due to a lack of robust APIs and standard/custom metamodels that can handle the different modelling and simulation tools and repositories in the market (Fisher, et al., 2014).

Gausemeier et al. (2013) developed a discipline spanning specification technique called CONSENS®, to manage the system engineering, in line with MBSE methodologies. The conceptual design specification technique consists of the following aspects: environment, application scenarios, requirements, functions, active structure, shape, and behaviour. CONSENS® employs a software support tool called mechatronic modeller. It uses a metamodel that defines the relevant models linked to the principal solution as well as their relationships. It can also handle complex dependencies in the system and can perform operations like tracing requirements, checking consistency, etc. In CONSENS®, the method by Gausemeier et al. (2013), the information used for the system model can also be used for the management of the development process. (Gausemeier, et al., 2013)

Ramos et al. (2013) has the criticism that the current MBSE are too complex and lack focus on the human-system integration. Ramos et al. (2013) also argues that the current standard for systems engineering (IEC 15288) is too flexible and lack solidity and that the systems engineering process should be intuitive, easy to use, easy to tailor, universal and logical. According to Ziegler et al (2018), even though MBSE has seen success in many disciplines that are based on traditional systems engineering practices, the methodology is not yet developed to handle issues in complex systems engineering such as emergent behaviour, especially in the case of complex adaptive systems of systems.

The Agile Systems Modelling Engineering (LITHE) methodology is a MBSE methodology developed by Ramos et al. (2013) emphasising more agile principles such as continuous communication, feedback and short iterations, while still following traditional systems engineering processes, using common system languages and using a systems model as main artefact. To achieve this the LITHE model, follow the SIMPLE (State the problem, investigate alternatives, Model the system, Integrate, Launch the system, assess performance, and Re-evaluate) process model. These stages include process steps such as; Characterize the operational domain, Identify/evaluate alternative design, develop software/hardware units and install the system. This is done with specific focus on human-systems integration on each step follow a systematic approach and are performed in an iterative and integrative fashion as per the agile methodology and similar to a spiral methodology. The functions, assess performance, re-evaluate and model the system, are performed transversal along the systems engineering process continuously supporting the other functions.

Zeigler et al. (2018) argues that with the increase in big data and IoT that incorporates multiple domains, experimentation with models is necessary to understand functions of the system, as well as engineer the system itself. Zeigler et al. (2018) claim that the existing MBSE methodology lack the simulation capabilities of analysing such system, especially emergent behaviour which is an inherent characteristic of any complex system. Zeigler et al. (2018) suggest that to overcome these restrictions, MBSE must evolve to handle human machine interaction analysis and resilient system design. Zeigler et al. (2018) suggests that discrete event system specification (DEVS), together with MBSE could bridge this gap in developing complex systems with MBSE. A workflow that according to Zeigler et al. (2018) could incorporate the use of simulation and MBSE would start with the development of a systems entity structure (SES). The SES forms the structure of families of simulation models, defining what is needed and how they interact. Due to the number of components in complex systems, there can be a large number of potential configurations of the SES. The SES then needs to be ”pruned” into a Pruned Entity Structure (PES) which is a specific configuration of the SES. Different PES is then simulated in parallel using DEVS simulation protocol, where the results is analysed using AI, which under human supervision improves the PES based on the simulation results.

Model-based systems engineering has according to Rhodes & Ross (2010) developed into a mature practice and was made and used in projects to enhance requirements, design practice and connect structural aspects to structural aspects of the systems. However, Rhodes & Ross (2010) argues that when it comes to more complex systems, to fully represent the system it is not enough to connect the structural aspects and behavioural aspects of the system. Due to the complexity, uncertainty, dynamic nature, and accelerating pace of change in these systems Rhodes & Ross (2010) suggests that along with the structural and behavioural aspects, the engineering method needs to address contextual, temporal, and perceptual aspects as well. The contextual level concerns the external environmental factors of the system. Rhodes & Ross (2010) argues that these factors are generally not fully considered but influences decisions in the system, especially over a longer period where the context of the system may change. The temporal aspects of the system are needed to characterise changes to the system over time and asses its adaptability to different contextual changes (Rhodes & Ross, 2010). The final aspect, perceptual, is according to Rhodes & Ross (2010) needed to identify and represent different perspectives of the system to different stakeholders. It is also needed to identify how the different stakeholder preferences could change over time as the context changes.

The main findings of this section are summarised in the table below.

Key findings
• Management commitment is important in transformation from a document-based system to MBSE. • Integration of heterogenous models are challenging. • Agile methodology can be used in the MBSE process. • Dealing with emerging behaviour is challenging for MBSE.

Table 11. Key findings from emerging cultural and technical challenges for MBSE

4.2.3. Case Studies on MBSE

Kaslow et al. (2017), based on the existing MBSE methodologies, tries to frame a generalizable step by step approach to develop CubeSats to improve the mission success rate. Cubesats are small satellites built by students that are inexpensive and faster to develop (Kaslow, et al., 2017). Kaslow et al. (2017) captures the developmental activities in an eight-step approach, which is generalizable and not tool dependent. The starting point is capturing the CubeSat mission requirement. In the first step, requirements were used to identify the use cases in the enterprise level. The second step is to build the relationship between use cases and requirements. In the third step, the use cases are captured along with the participating stakeholders. In the fourth step, the use case descriptions are further developed to include aspects like primary actors, supporting actors, preconditions, triggers, and post conditions. In the fifth step, the use case descriptions are captured into the models. In the sixth step, the use case scenarios are built, which capture the functionality that the system need to fulfil. In the seventh step, all the identified activities are linked to the use cases. In the step eight, the decomposition of activities is completed for all the subsystem level to cover the full functional decomposition. According to Kaslow et al. (2017), this approach is helpful in maintaining traceability and is recursive.

Aleina et al. (2016) and Fusaro et al. (2017), in their cases studies, followed a structure of stakeholder analysis, mission and objective definition from which an initial requirements list is derived and relationships between them are established through the MBSE approach. The process was then followed by the identification of necessary functions and the identification of all possible products that would be able to perform these functions.

In the case study by Aleina et al. (2016), on the development of a space tug which is a re-usable space vehicle for moving objects between orbits, explore the benefits of using systems engineering processes with MBSE, especially in the initial conceptual stages. Aleina et al. (2016) especially emphasises the selection of tools to be used in the development of the system. According to Aleina et al. (2016), these steps should all be performed iteratively. Aleina et al. (2016) put special focus on the selection of software to be used throughout these stages and it was decided in the case to use Doors® for requirement management and Raphsody® for functional and structural design and analysis. The reasoning behind the choice of the toolchain was based on the possible level of interaction between the tools and potential for traceability. In the space tug case, Aleina et al. (2016) ensured proper traceability and verifiability by choosing software with well-established interactive capabilities, allowing for allocation and verification across the different platforms.

The case study on a hypersonic and suborbital transportation system by Fusaro et al. (2017) aims at establishing a baseline for the system through MBSE methods with the implementation of Quality Function Deployment (QFD). To achieve the best combination of alternative solutions, Fusaro et al. (2017) in this case applied the use of a QFD. Fusaro et al. (2017) argue that by using a QFD with weighed criteria derived from the stakeholder analysis, a baseline that will fulfil the requirements and perform the maximum amount of mission scenarios can be identified. After the most relevant high-level system structure has been identified through the QFD, the individual elements of the system were then improved upon in further detail, both in functional and structural aspects. Fusaro et al. (2017) argues that the use of an MBSE methodology, with the integration of a QFD, together supported the development of different mission alternatives and allowed the project to select the most optimal system baseline to perform the mission alternatives.

Claver et al. (2014) describes a successful case study of the Large Synoptic Survey Telescope (LSST) and the methodology developed in the process. The project group choose to develop a methodology based on MBSE principles using the systems engineering language SysMl. One of the main reasons for using MBSE in the project was due to the members being geographically separate, and MBSE and the common system model allowed them to work, document and share information on the same interface (Claver, et al., 2014). The development model is based on a triangulation between the systems requirements, the system structure and the system behaviour where the structure and behaviour should satisfy the requirements and the behaviour is allocated to the structure of the system (Claver, et al., 2014). According to Claver et al (2014) the development the model also consists of three views which is the requirements, the logical view where high level logical structure are and functions are defined to increasing levels of details until they can be allocated to the final view, the physical view. The project group call the view structure the LSST System Architecture (SysArch) model. By ensuring both traceability within the views, and through the triangulation of the requirements, structure and functions, the SysArch model keeps the system consistent throughout development and was according to Claver et al (2014) a very effective method of developing the complex system LSST.

Do & Cook (2014) through their case study on a ground-based air missile system, which used a methodology of the object-oriented systems engineering methodology (OOSEM) with SysMl, identifies a set of research challenges for the development of complex systems using MBSE. Do & Cook (2014) found three key areas where further development is needed when developing complex system with MBSE and especially with SysMl. The three areas are, model-based requirements engineering, model-based systems engineering design and analysis, and model integration and integrated tool environment development. According to Do & Cook (2014), modelling requirements in SysML lack the necessary traceability required for large projects and inconsistent requirements will eventually arise. According to Do & Cook (2014), MBSE practice with SysML mostly catalogues the different requirements and more research is needed to improve modelling and mapping of requirements. Do & Cook (2014) argue that MBSE with SysML is capable of "capturing system design", it mainly does so, based on designs that have been developed outside the SysML infrastructure and then built in SysML. According to Do & Cook (2014) the capabilities of MBSE with SysML to design and analyse systems needs further development to better support engineers needs as well as to build executable models that where trade-off analysis can be performed, and designs can be better evaluated. Finally, Do & Cook (2014) argue that further research on integration between tools is necessary. In MBSE tools mainly focuses on one aspect of systems engineering and integrating them can lead to more executable models and better incorporation of "artificial reasoning".

The main findings of this section are summarised in the table below.

Key findings
• Limitations are seen in tool / model integration due heterogeneity. • SysML language needs to evolve to better support MBSE. • Traceability, and identification of inconsistencies is challenging.

Table 12. Key findings from MBSE case studies

4.2.4. Model-based systems engineering and data

According to Zhan et al. (2015), even though MBSE has been growing in use to develop complex systems, the major MBSE methodologies put too much focus on the representation of the models themselves and not enough focus is placed on the data elements of which the models are build, verified and analysed. The systems engineering process can be structured according to three core phases, requirements analysis, functional analysis and architectural design, and while the models may visualise these phases, it is the data elements which are the drivers for the process (Zhan, et al., 2015). Zhan et al. (2015) suggests a data centric approach to MBSE. This approach is based on the notion that to fulfil all the system requirements, the system with its subsystem needs to fulfil specific capabilities. This can according to Zhan et al. (2015) be done by building a high-level data meta model which is based on answering questions that can be grouped according to 5W1H (What, When, Where, Who, Why and How). By making connections between the 5W1H, answering questions such as who does what when, and mapping the subsystems to what function needs to be performed and how it will be performed to fulfil what capabilities, therein fulfilling the requirements, the data elements are driving the systems engineering process and the models act as a visualisation of the data (Zhan, et al., 2015).

Lindblad et al. (2018) suggests that a reason for MBSE not yet being a widespread common method, at least in the aerospace industry, is because the method itself becomes too complex as well as being too inflexible. In the aerospace industry, MBSE is mainly used in the initial project phases and not iterated upon throughout the life cycle. This is according to Lindblad et al. (2018) because when underlying engineering data is changed, if that change is not connected throughout the system, in MBSE the effect on the system as a whole may not be clearly recognized, leading to inconsistencies in the underlying assumptions of the system. Lindblad et al. (2018) suggests that by basing the development on a structure and consistent database, where the engineering data are connected will allow for better traceability, make it easier to spot inconsistencies early and allow for automation and optimization of the system structure. By using a structured database with clear relationships between the data, optimizer tools can be used to automatically come up with optimal solutions and thereby according to Lindblad et al. (2018) MBSE can become "more than a visualisation tool" and become a tool to support engineers in their decision making.

Li et al. (2019) presents an iteration to the established "V" life cycle model for systems engineering based on the inclusion of a big data driven section to the model. Li et al. (2019) argue that the "V" and the "double V" model, which is the "V" model with the addition of a MBSE branch, still lack in its performance in several aspects when developing complex systems. According to Li et al. (2019), MBSE still faces challenges in verification, where verifying the consistency of the model against the physical system and verifying the simulation accuracy continues to be an issue, leading to the continued need of physical tests. There are also challenges for new fields, or fields with low accumulated knowledge. In these areas it is difficult to claim how relevant the model actually is. There is also no way of integrating domain models in a holistic way, causing domain models to be analysed separately, dividing the relationships between them. This is because the development of models take place in their respective domains and they inherit the characteristics of those perspectives, which limits the scope of a collaborative analysis. Li et al. (2019) propose adding another "V" to the model, consisting of a big data life cycle, making it the "triple V" model. With the "triple V" model the product, its model and the data are all integrated. The data in the in the "triple V" provides the connection between verification and validation with the model and the product, supporting each other throughout iterations in the process. By adding the data branch in the model, data generated by simulations as well as the physical product can be used to learn about the model and the product, as well as improve upon them. In the model by Li et al. (2019), data will be at the centre of the interaction between the model and the product.

Herzig et al. (2014) claims that complexity in technical systems is often managed by understanding the system from different viewpoints that consists of factors such as concerns of interest, different levels of abstraction, variety of factors and context. In a complex system, there may be several inter-relations due to the different viewpoints and this can result in inconsistencies in the system given that there are heterogenous models present in the system. According to Herzig et al. (2014), identifying and resolving inconsistencies is important for the verification and validation. Identifying the inconsistencies is important as early detection can save the developmental cost. Inconsistencies arise in many different forms such as mismatch between model and test data, violation of well-formedness rules (rules related to modelling language), not following guidelines etc. Herzig et al. (2014) proposes a way to identify the inconsistencies through a graph method, where all models are represented in a graph. Pattern matching method is used to identify the inconsistencies, which is similar to identifying inconsistencies through deductive reasoning.

Herzig et al. (2014) builds on the previous theories on handling inconsistencies and proposes that a triplet can represent a relationship – a subject, predicate, and object – at the basic level, in the form of a graph. Herzig et al. (2014) argues that by searching for a sub-graph, it is possible to identify the presence of an inconsistency. Herzig et al. (2014) proposes an architecture concept which represents the key elements of the inconsistency management apparatus. With the aid of machine learning and heuristics, it is possible to improve the performance of pattern matching. The automated identification of inconsistencies can aid the verification and validation phase. Some limitation of the method proposed by Herzig et al. (2014) could be that pattern matching may not work in certain cases where relationships are missed. Instead of deductive reasoning, abductive reasoning could be used to overcome this.

Heber & Groll (2017) reasons that traceability in the development of a system is a crucial aspect, especially with an increase in complexity. By connecting Product Data Management (PDM) and Product Lifecycle Management (PLM) with MBSE, and integrating digital twin and blockchain technology, Heber & Groll (2017) suggests that the traceability throughout the life cycle can be solidified. By using a blockchain, where a new block is created whenever a new element is created or updated, and having the information in the block connected to MBSE artefacts, this chain can according to Heber & Groll (2017) act as the "backbone" of the PLM system and enable traceability of elements and their relationships throughout the life cycle. The digital twin adds usability and manageability to the system by being an easy to interact with representation of the system at each stage in time of the development from the point of sale. The digital twin is also connected to a blockchain containing all necessary information needed by the engineers about the system and the parts such as geometry or source code.

Madni et al. (2019) categorizes four levels of maturity for 'digital twins. Level 1 is pre digital twin, which is equivalent to a virtual prototype. It only helps in validating key aspects of the system. Level 2 is a 'Digital twin' that captures the performance, health, and maintenance data from the physical twin, through batch updates. These updates support the conceptual design and development for the product. At this level, the aim is to use the digital twin to find the 'what-if' scenarios, so that it can aid in supporting the physical twin in the right way. Level 3 is the 'adaptive digital twin', where the system incorporates the user behaviour in addition helping to understand various contexts. In this level, supervised machine learning algorithms can be of help to understand and predict the patterns. Real time update of digital twin is desired in this level. Level 4 is the 'intelligent digital twin', which uses unsupervised machine learning to understand and predict patterns from the operational data. High autonomy of the digital twin is the highlight here, where it can analyse various type of real-world data. Digital thread can help in the knowledge transfer across the value stream of MBSE, making the digital twin a sole source of truth. The systems engineering using MBSE can be transformed greatly by incorporating the digital twin as the data and analytics improves the system knowledge. (Madni, et al., 2019)

Schluse et al. (2017) identifies that there is a gap between the systems engineering and the simulation technology. The simulations lack a framework that can combine the different simulations of various domains, due to the complexities of the models. MBSE methodology, already consist of model preparation and simulation, but lacks an integrated approach when it comes to simulation. According to Schluse et al. (2017), experimentable digital twins (EDT), which has its beginning in the eRobotics methodology, could be the possible way to integrate the simulation methods using the digital twin concept supported by internet of things (IoT). As the models get more complex, classical methods of simulation could become less effective. EDT can intelligently improve the model generation time and transition between modelling and simulation using an 'understandable structuring element'. For the system engineers, EDT takes away the burden of running the simulation for the whole system as it intelligently chooses the needed simulations from the EDT networks and runs the required algorithms to produce the results. EDTs could act as an effective link between the various aspects of MBSE and simulation (Schluse, et al., 2017).

According to Schluse et al. (2017), interoperability issues between simulations prevent the engineers from having a holistic view, which is a blockade in achieving the synergy of the simulation methods. An EDT essentially combines a simulated data processing system (DPS) and simulated human machine interface (HMI). EDT and virtual test bed (VTB) make it possible to handle the modelling, simulation and verification of a dynamic system at a micro level including the possible interactions. Simulation based optimization can benefit from EDT by evaluating the parameters between the simulation runs through a cost function. EDT can help MBSE to achieve its full potential where models support the full life cycle by integrating simulation and modelling, leading to better designs and cost-effective solutions. As EDTs and VTB link simulation with reality, they can become vital in the development of intelligent systems.

According to Di Maio et al. (2018), MBSE models are usually too abstract, leading to difficulties in complex analysis. MBSE models also commonly cannot easily integrate domain models (for example CAD models) making it more difficult to perform analysis on the systems configuration or advanced “in the loop” simulation, leading to lower confidence in the decision making. Di Maio et al. (2018) argues that the methodology Closed-Loop Systems Engineering (CLOSE) can help overcome these weaknesses. CLOSE is based on a previously developed methodology, the Model Driven Engineering Process (MDEP). The MDEP model is based on separating the development in three areas, the client & environment, product realisation and system design where product realisation and system design is coupled through a functional matrix. By separating the system design from product realisation, the process can according to Di Maio et al. (2018) decrease the amount of feedback loops in the development. CLOSE adds to this by incorporating EDTs which are a virtual one-to-one representation of the system, at the centre of these three areas of MDEP. With the EDTs interacting and connecting the areas, this could allow for analysis and verification throughout development with more and more detail as the domain models developed. The EDTs can according to Di Maio et al. (2018) provide a holistic executable model of the entire system and not simply aspects of it. The CLOSE model with the incorporated EDTs should also decrease cycle time in development, allow for a high degree of re-use of model elements and with its decoupling of the system design and domain specific product realisation decrease the amount of feedback loops in the development (Di Maio, et al., 2018).

The main findings of this section are summarised in the table below.

Key findings
• Linking data to models using a structured database can help optimization. • Data can help in verification and validation and can improve simulation accuracy. • Concepts like digital twin can link MBSE and data and help in simulation optimization.

Table 13. Key findings from MBSE and data

4.3. Data driven methods

For the data driven methods part of the review, 23 different articles where included. In this selection of articles, 80 unique authors contributed with 8 authors being included in more than one article. The articles where published in 23 unique journals or conference proceedings. The selection was distributed between 2007 and 2019 according to figure 7. In total, the selection had a citation count of 1380 citations over all articles.

The distribution of articles in the subthemes of Data driven development are presented in table 6 with the number of articles in each section and number of total citations.

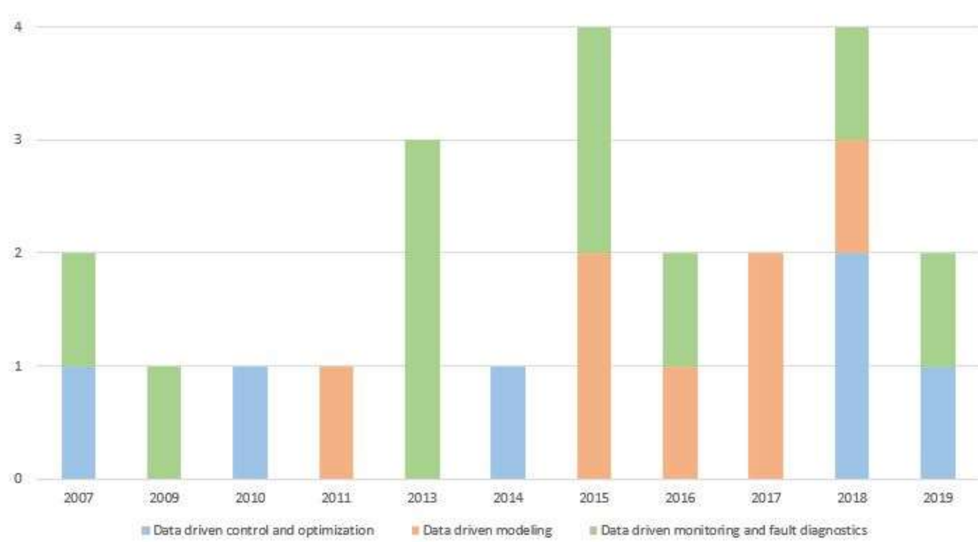

Figure 7. Distribution of articles in subthemes of data driven methods.

Subtheme	Number of publications	Number of citations
Data driven control and optimization	6 articles	827 citations
Data driven modelling	7 articles	69 citations
Data driven monitoring and fault diagnostics	10 articles	465 citations

Table 14. Article and citation count in subthemes of data driven methods.

4.3.1. Data driven control and optimization

According to Uraikul et al. (2007), There are three general approaches, analytical, data driven, and knowledge based, to support intelligent decisions in monitoring, control, and diagnostics. The analytical approach uses theories and mathematical models. The models generate outputs in the form of state estimations, parameter estimation and residuals. The data driven approach bases their results on process data. The knowledge-based approach on the other hand uses heuristics and reasoning based on artificial intelligence technologies to achieve the results. In a survey of what according to Uraikul et al. (2007) where four of the most comprehensive frameworks dealing with intelligence controls and diagnostics, Uraikul et al. (2007) found several general characteristics that were common between the frameworks, however, they varied in their execution. In all four frameworks the three different approaches were integrated, although with different priority. All four frameworks also use different methods of integrating the solutions. The frameworks did however also show some weaknesses. For example, one framework showed that scaling up the system could be an issue while another showed issues in coordination between tasks and a third framework was shown to be too rigid and could have issues in dealing with changes. From the survey, Uraikul et al. (2007) identify six desirable general attributes of intelligent frameworks. The framework should be able to coordinate between different control tasks. It should be capable of integrating the three different approaches (analytical, data driven, knowledge based). The framework should be able to coordinate between different representative views like models or cases. It should be able to support a global database and management of process knowledge. It should have a hierarchical structure for data models at different levels of abstraction. Finally, the framework should have capabilities to deal with change according to its environment.

Ren et al. (2019), discusses the effect of vibration factor in the life of downhole drilling tool and explores how field reliability big data (FRBD) can aid in the analytics, tool life prediction, prognostic health monitoring, condition-based monitoring, maintenance planning and optimization. FRDB can be a huge set of data collected from all over the world, with sizes amounting to Terabytes (TB) and Petabytes (PB), containing covariate and time varying information as well as reliability related information. This data is obtained from sensors that measure, monitor, and record the downhole - axial, lateral, and torsional – vibration data. Traditionally, the reliability data is obtained from population data from real world testing experiments and the analytics uses empirical, probabilistic, and statistical data. In comparison, FRBD can help to minimize in service failure and life cycle cost by provide more insights. It can also aid decision making in the areas of design, testing, operation, maintenance, and warranty. FRBD can be analysed to measure the actual usage information of drilling tool and prolong its life by modifying the chosen parameters as per the recommendation. It can be used for prognostics for short- and long-term predictions of the remaining useful life (RUL) of the tool. FRBD should be integrated including information for various source such as product design, testing, manufacturing, quality, and field to enhance the breadth of reliability analytics. (Ren, et al., 2019)

Liu and Goebel (2018) discusses the learnings from the NASA University Leadership Initiative (ULI) which is a five-year project aimed at addressing the safety needs and the corresponding technological requirements for the next generation National Airspace System (NAS). The project is to develop an integrated fusion methodology to be used for the prognostics and safety assurance for NAS which needs to be capable of ensuring safe operation in a complex airspace by proactively detecting and resolving threats as well as providing prognostics. To develop a new algorithm for aircraft dynamics simulation, a hybrid approach, combining the physics of the dynamic system and the data driven learning is used which enhances the learning and prediction. It provides additional constraints for learning and predicting the system behaviour when the physics-based models are integrated into the data driven learning models. It also enhances the extrapolation capabilities. In this way, the training cost of the purely data driven method could be reduced. Liu and Goebel (2018) suggest that a rigorous information fusion, using both data driven, and physics-based, is needed for the complex system prognosis, where a huge amount of data that is hierarchical and dynamic in nature, both at spatial and temporal scales are involved. Big data analytics can enhance the capabilities of the organization by developing new insights leading to better decisions. (Liu & Goebel, 2018)

Norman et al. (2018), focuses on the use of big data analytics in the testing and evaluation of aircraft in the Joint Strike Fighter (JSF) program. Though, the complexity of the systems in the department of defence (DoD) evolved, the evaluation infrastructure, according to Norman et al. (2018), has not caught up with it. By making use of data driven methods, this can be bridged, making decision making faster, better, and smarter. In the DoD, there was an issue with knowledge management as the program managers used to conduct tests without realizing that the same tests where conducted in other programs and that they can learn from them. To reuse the knowledge, the new knowledge management model was conceptualised with four functional areas: data gathering, warehousing of data to make it available to users, providing analytics capability and visualization of the data. The Test Resource Management Center (TMRC) of DoD aims to use big data analytics and cloud technology to improve the evaluation efficiency and reduce the decision-making time. According to Norman et al. (2018), data analytics can help to identify issues earlier in the life cycle, saving cost and time, lowering the risk. It can also identify patterns that humans might miss, help the developers to discover 'unknown-unknowns' which is one of the highest risks in complex products acquisition. (Norman, et al., 2018)

Qin (2014) in his perspective article on process systems in chemical systems, claim that while process systems gather a vast amount of data from sensors and indirect measurements, this data is mainly stored to be used by control systems after an incident, in a reactive manner. According to Qin (2014), the current data analytics methods are also mainly based on "clean" and structured data, limiting the potential benefits from big data. Qin (2014) instead argues that process system development should incorporate data driven methods like data mining and machine learning methods which can make use of unstructured data, detect root-cause faults faster and identify quality related faults sooner than conventional techniques. To make the shift towards more data driven development, Qin (2014) takes the perspective of the different 'V' of big data. For variety, more heterogeneous sources of data should be utilized to gather data to gain information. To increase value and veracity, process system development needs to adapt more machine learning methods. To gain volume, data mining of historical data based on time series should be done to better analyse events, make decisions and identify root-causes. For velocity, system architecture should adapt to a "data-friendly" information system, which could complement rather than replace the current "control-centric" systems (Qin, 2014).

Wang (2010) presents concepts and methods that together forms the methodology of Parallel transportation Management Systems (PtMS) in the field of Intelligent Transportation Systems (ITS). The parallel management and control of systems is meant to be a data driven approach to systems development and decision making and since it is developed within the area of ITS, it considers both engineering and social aspects (Wang, 2010). According to Wang (2010), parallel, in this context, is the interaction of the "real" system, and one or several virtual or artificial versions of the system. PtMS by Wang (2010) is based on the ACP approach (artificial societies, computational experiments and parallel execution). The artificial societies are used for modelling as well as representation of the systems, the computational experiments are used to perform analysis and evaluations, and control and management is achieved through parallel execution between the real-world system and the virtual system. According to Wang (2010), the artificial systems approach can better support a complete approach to representing the system while the computational experiments can support the adaptability to solve complex problems and, learn and improve upon the solutions. The parallel execution of the system helps in the implementation of the system, as well as provide validation and evaluation on different solutions on a system level as there often in complex system is not one optimal solution to a given problem (Wang, 2010).

The main findings of this section are summarised in the table below.

Key findings
• Data can provide insights and aid decision making. • Decision making is faster compared to model-based methods. • Data helps to identify unknown relationships.

Table 15. Key findings from data driven control and optimization

4.3.2. Data driven modelling

Within the area of automotive systems engineering (ASE), Bach et al. (2017) argue that the increase of advanced technology, complexity and dependencies between features cause traditional descriptive and conceptual models to no longer be adequate on their own. Bach et al. (2017) suggest that data driven development should be incorporated together with the current practice of using the V model in ASE. Bach et al. (2017, pg.285) defines data driven development as "approaches, which utilize data and data analytics to substantially influence business or design decisions" and proposes that the use of data from real world testing, as well as from contextual data, such as in this case weather data or traffic data, can help overcome challenges in ASE. Within the requirement solicitation and analysis phase, data can provide answers to simple questions such as who, what, when, how, to whom questions. Data can also provide a better understanding of the system, the environment and the context as well as even uncover connections between parameters not previously thought of (Bach, et al., 2017). In the design phase, data can support the determination of desired characteristics of the system and provide a better understanding of dependencies, uncertainties and potential outlier errors, supporting design decisions and decreasing errors in the implementation phase. The real-world data can also help narrow down testing scenarios on a unit or subsystem level and provide initial values for further testing and analysis (Bach, et al., 2017). In implementation and validation, the recorded data and contextual data can support realistic simulation of the whole system for a multitude of different scenarios. The simulations can help focus the real-world testing and give experts an overview and a good base to validate the behaviour of the system (Bach, et al., 2017).

Bach et al. (2017b) argues that the increase in complexity especially causes difficulties for the verification and validation of the systems. Bach et al. (2017b) proposes a combination method to address these challenges. The method is called the reactive replay method, which combine dynamic simulations of the system, including the dynamic behaviour of the system and its environment, and real-world prototype-based testing, with both parts feedbacking to each other. This way, specific aspects of the system can be simulated, combined with real world data, simplifying the simulation models. According to Bach et al. (2017b), as the testing scenarios are then based on real world data, they would not require further validation. However, Bach et al. (2017b) claims that there are still issues in determining when scenarios cover the testing requirements, or with simulations, doing an excessive amount of testing. Currently, scenario selection is done based on expert knowledge. Bach et al. (2017b) instead proposes a two-step method of scenario selection to achieve sufficient coverage with the fewest amount of scenario testing. In the first step, specification-based selection, scenarios are partitioned based on system requirements, enabling the generation and combination into many sets of scenarios. However, the generated scenarios can still contain a substantial amount of overlap in situations and information. Therefore, the second step of data driven reduction is necessary. By utilizing the data from the reactive replay method, scenarios based on real world data can be generated that will sufficiently cover the testing requirements with minimum parameter overlap (Bach, et al., 2017b).

Huang et al. (2011) proposes a data driven approach for the automated selection and re-use of the model components which are pre-build and validated. Building validated models are costly as it takes time, effort, and expertise. Automatic model generation (AMG) method generates simulation models from data gathered from several sources by using data analysis and data structuring algorithms which configure the models and create them. Complex systems have a hierarchical structure where it can be decomposed into subcomponents and further into components. The behaviour of a model depends on the subcomponents and its structure. If this information can be gathered from data sources, it can aid in AMG, facilitating the re-use of pre-validated model components. Though a big model-component library would be required for AMG, it is justified according to Huang et al. (2011), considering the long-term benefits. Model-components are the building blocks which are self-contained, interoperable, reusable, and replaceable units that have a well-defined interface. A model generation algorithm, capable of selecting, structuring, and configuring the model-components is employed by AMG with the help of model selection heuristics. There are two types of applicability conditions used, behavioural and structural. Structural pre-condition is associates model component with objects whereas behavioural pre-condition define the dynamic conditions to determine the expected behaviour of the model. These conditions together define the logic of model component selection and structuring. According to Huang et al. (2011), AMG has two parts: the first part consists of using data and models for AMG and the second part consist of analysing the simulation output with available data for automatic model calibration. The data sources can be describing geometry, geomatics or topology related to the system. It can also be describing resources, orders, or demands, processes or operations, products, or services. The data is subjected to selection, pre-processing, mining, and post processing to build knowledge from it (Huang, et al., 2011).

To deal with the increasing complexity in systems Hybertson et al. (2018) proposes a shift of focus to increase the use of evidence in systems engineering drawing from areas such as law and medicine. Hybertson et al. (2018) classifies evidence as "any information or artefact that helps objectively evaluate the validity of an assertion, answer a question or resolve an issue" and argues that an evidence based process can support decisions better connected to the core issue more consistently. According to Hybertson et al. (2018), many different fields and approaches supports and contributes in the evidence-based systems engineering. Areas such as knowledge management to capture and distribute already known information including expert knowledge, model based approaches which supports analysis reasoning and learning from evidence and, analytics where through methods in data analytics, big data and IoT new evidence can be gathered which is specific for a given situation. The framework of evidence-based systems engineering by Hybertson et al. (2018) is based on five major steps. In the first step, setting the stage, a question to be solved should be stated. However, Hybertson et al. (2007) argue that before stating a question the situation must be fully understood which is not always a straightforward path in complex systems. Hybertson et al. (2018) here promotes the use of systems thinking to capture a holistic view. In the second step the evidence is gathered. This is done through both to use of a knowledge base containing experiences, theories etc. and collecting new data using models and data analytics methods. The third step of mediating the evidence supports the development of the best solution for this specific situation. Here, contextual factors are considered as well as norms and preferences. The fourth step is to apply the evidence generated to answer the initial question. The final step is to learn and evolve. This step is iterative and is carried out throughout the framework. For example, the situational data can feedback to the knowledge base or the application can feedback to research to raise new questions to be investigated (Hybertson, et al., 2018).

Opiyo (2015) proposes a framework for a pipeline for data analytics aimed at supporting the decision making in the development of complex product systems. The idea is to use the insights from data, that was unavailable before, to predict aspects like cost, assembly time etc. which can aid product development. As the systems become more complex, it is difficult to predict the aspects like performance, reliability, and cost. If the developers can predict these aspects early by accommodating the knowledge gained and resolve issues before they occur, it can result in a successful product. According to Opiyo (2015), the developmental phase consists of i) need analysis ii) component and feature identification iii) modelling and representation iiii) exploration (engineering analysis and data analysis). The data for the analysis is first acquired from many different sources, the raw data is then subjected to pre-processing and transformation and then analysed. In the acquiring and storing of raw product data, traditional data gathering methods can be combined with the latest methods such as sensor data, IoT device data etc. Once the data is received a broker hands it over to an analytics system for further processing and transformation. The pre-processed and transformation phase enriches and re-represents the data. A key requirement of this phase is to have low latency characteristics. The final step in the pipeline, the data analytics phase which captures the trends and patterns of the data. Opiyo (2015), breaks down the component and feature identification phase to three sets of features, low end complexity manifestation (CM), high end CM, and basic system manifestation features connected to the components. Advanced analytics, machine learning and statistical methods employed in the pipeline, can be used to gain insights which can be inputs for designers to influence the earlier mentioned features. It shall be noted that the predictions depend on availability of historical data sets. Opiyo (2015) suggests that the pipeline and strategies discussed will aid the developers to acquire data and insightfully explore the design space to come up with superior products.

Ding et al. (2015) claim that due to the complexity in wind power turbines it is not possible to explain the behaviour through analytical expressions, rather data driven methods are crucial in the development. By using large amounts of data generated from sites of interest for wind power, it is then possible for developers to give an accurate estimation of variables. The data can also support prediction of in this case, the power production of the turbine as well as the effects of changes in the environment on the system (Ding, et al., 2015). Ding et al. (2015) also discusses the concept of local and global models for data analytics. Local models use many parameters and are very useful for capturing features of a specific portion of the data. However, local data models only look at the data within one portion, ignoring characteristics and strengths of other portions of the data which, according to Ding et al. (2015), is an issue with current methods in the wind industry. Global data models instead require fewer parameters and, opposite from local models, when a change is made somewhere in the data model, it will have an effect in other areas as well. Global models are also easier to scale but lack the flexibility of local models. Ding et al. (2015) suggest that data analytics models should fall between local and global models, being able to utilize benefits from both.

Mount et al. (2016) explores the opportunities that data driven methods offer for the socio-hydrology and the associated challenges. Socio-hydrology model is an emerging model from the traditional hydrology model where human, social, economic and infrastructure components are added. Due to these factors, the complexity is associated with the total system increases which results in the system's transformation from the conventional model to a socio-hydrologic model. The hydrological modellers face a challenge due to this, but it also provides them with opportunities to analyse new type of information to understand how interactions happen between the social domain and the hydrology domain. Mount et al. (2016) suggests that in the future, the data driven methods will find more use in the socio hydrology field. As data driven models can discover the model structure from the data, dependency of a priori model is minimized, if not completely avoided. The conventional physical models focus on hypothetical knowledge which is assembled from several hypothetical models. In the hydrology models, some amount of data is used as it is one of the constituents of the model itself. However, in the new generation of hydrology models, a more comprehensive representation, improved estimation, and prediction capabilities are desired. According to Mount et al. (2016), the capacity to develop complex socio-hydrology models can surpass the hypothetical knowledge that is available at hand. Here the data driven methods can help in hypothesis development. This is because they can identify the influence of the variables on the system responses. The hybrid methods which combine the advantages of both data driven and hypothetical models, deliver better predictions. It also helps to gain more hypothetical insights through heuristic explorations as well as data driven structures and behaviours (Mount, et al., 2016). Mount et al. (2016) identifies some challenges in adopting the data driven methods. A need for rigorous model development protocols and a shift in perspective of the data driven modelling community from the software solution-based thinking. Generation of large number of training data is expensive and simulation efficiency is reduced when a full emulation model is used. Despite the limitations, Mount et al. (2016) finds that data driven methods are complementing the conceptual/physical models and are set to be essential in socio-hydrologic systems owing to the increased system complexity and decreased system understanding.

The main findings of this section are summarised in the table below.

Key findings
• Building validated models require expertise, effort and consume time. • For model-based methods, complexity causes difficulty in verification and validation. • Data driven methods can provide better understanding of the system behaviour. • Data can give feedback and insights to earlier phases of development. • Generation of large set of training data can be expensive.

Table 16. Key findings from data driven modelling

4.3.3. Data driven monitoring and faults diagnostic

As a result of the increasing complexity of systems and the cost of these systems, the demand on safety and failure prevention is increasing, placing higher requirements on fault detection and diagnosis (FDD) methods (Dai & Gao, 2013). Dai & Gao (2013) describe in the analysis of literature that there are three different categories of FDD methods. Model-based FDD methods which are based on online data, Signal-based FDD methods based on sensor data or online data, and knowledge based FDD methods which use historical data and smart computing. There are also many FDD methods using an integration of methods. All three categories are according to Dai & Gao (2013) considered data driven as data and the understanding of the data are fundamental parts of the process and they are all based on the notion of information redundancy. The redundancy is based on either checking the data against a model, knowledge or checking the consistency of the data. The model based FDD is based on identifying inconsistencies in state variables (Dai & Gao, 2013). These variables are commonly related to physical values of the system, making it rather simple the check for changes from the nominal value, if the data fits the model. However, building the model to process the data takes considerable effort (Dai & Gao, 2013). The signal based FDD methods instead do not require a model to process the input data. These methods identify faults through the analysis of patters from the data, from sensors or from online sources. Faults can then be found through the correlation of specific patters to certain faults in the system (Dai & Gao, 2013). When the system is too complicated to use explicit models, or analyse patterns from signal, Dai & Gao (2013) argue that knowledge based FDD methods are necessary. Knowledge based methods identify faults and the health of the systems by "learning" from a large amount of data, using AI to uncover the knowledge of which the data consists and identify inconsistencies in the system. However, Dai & Gao (2013) argues that for complex systems, the best solution is some combination of the three categories.

According to Mosallam et al. (2015), the increased demand on maintenance has paved way for a strategy which focuses on predictive maintenance capabilities and estimation of remaining useful life (RUL). A data driven approach is gaining popularity in this field when defining the system through accurate physical models is not possible. Mosallam et al. (2015), discusses the path towards data driven prognostics and builds a generic component based prognostic methodology. The data driven methods try to map the relationship between variables and can also represent uncertainty in a probabilistic form. RUL mapping can be done both cumulative way and direct way. In cumulative model, the RUL is predicted using the empirical models which map out the degradation evolution aiding the calculation of health status. In direct method, the sensor data is collected to calculate the end of life (EOL) directly without calculating the health status. Monitoring can be done at system level and component level. The first step in adopting data driven approach is to identify the critical components. Qualitative/quantitative hazard analysis, failure mode effect analysis (FMEA) and fault tree analysis (FTA) are some tools that are used in this regard. Physical parameters are selected based on criteria such as speed, temperature, position etc. Selection of monitoring sensors are important. Aspects such as parameter to measure, accuracy, reliability, range, resolution, characteristics, and cost are the important factors here. Sensors can be used to capture both event data as well as condition monitoring data. After data acquisition from the sensors, data pre-processing must be conducted to prepare it for better analysis. Handling missing data, noise reduction, normalization and smoothing are the different categories in this phase. Health indicators are constructed from the component degradation evolution model and the health status generated in a probabilistic format help in decision making. According to Mosallam et al. (2015), this method is quite generalizable for prognostics of complex systems.

Galar et al. (2015) argue for a more context aware approach to assess the health of a system, and to predict future faults and maintenance needs. According to Galar et al. (2015), to accurately assess the health of complex systems, data collected from independent systems needs to be integrated to form an aggregated data set. This aggregated data can also be used to gain new insights about the system. A context aware system is according to Galar et al. (2015) commonly very complicated, including information and data from many various sources such as data gathered from the internet, internal work orders, sensor data, external monitoring systems, etc. Galar et al. (2015) argues for a hybrid modelling approach to a context aware health diagnosis and prediction. The hybrid model consists of symbolic models, physics-based model and data driven models. The symbolic models are used to capture expert knowledge, the physics-based models capture the mathematical degradation of the system and, the data driven models utilize sensory and historical data. Through the combination of models, the data driven models can find and confirm physics-based issues in the system, and previous knowledge from the symbolic models can support the data driven models when sufficient data cannot be provided (Galar, et al., 2015).

Earlier, Galar et al. (2013) proposed a methodology that predicts railway vehicle breakdown by using the information from track side (infrastructure side) and from the onboard systems side (equipment side). Empirical relationships described in words are the basis of the symbolic model, which can be found in work orders and maintenance reports, mostly handwritten by the maintenance personnel. However, they are less effective in building complicated dependencies and time varying behaviour. Work orders from both rolling stock and infrastructure need to be integrated to build the scenarios under which the breakdowns occur. Data driven methods depend on the relationships built using a training data taken from the system. Mathematical models or physical models are based on physics of failure, first principles or empirical relationships. Physics based models are useful in understanding dynamics of the system in time varying conditions. Building relationships between the variables in system degradation context is challenging. However, in the railway engineering, there are physical models for rolling stock and track degradation, e.g. models for degradation of wheel and track. By combining these methods, the accuracy of fault prediction improves. (Galar, et al., 2013)

Villarejo et al. (2016) follows the hybrid model by Galar et al. (2013), to develop a contextual awareness hybrid model to calculate RUL to manage the life cycle of railway equipment. The disparate data sources are integrated using a software called OPTIRAIL® which interfaces with the other connected systems, storing data in a configured database. Data collected consist of data from assets as well as from events, along with the timestamp. The context engine makes use of this to build relationships between events and assets. (Villarejo, et al., 2016)

According to Villarejo et al. (2016), data driven fault diagnostic models are time invariant, i.e. the deterioration of system do not vary with respect to time. This is a limitation for the method which can be overcome to an extend by using more data to increase the accuracy of the model. Due to the complexity of the railway systems, it is difficult to base it on a single model and hence data driven methods are suitable. The physics-based methods are capable to handle dynamic and time varying nature of the system but the effort to validate the model can be costly. Developing constitutive relationships among the complementary variables is a challenge for the physics-based degradation models. Physics based models are also weak in handling unknown failure modes. Hybrid models can bring the strengths of both methods, including that of symbolic models. Villarejo et al. (2016) argues that the hybrid models are ideal for prognosis when the historic data is low and expert classification is limited.

Liu et al. (2018), proposes a framework to implement AI in high speed railway (HSR) systems, by creating a cyber twin for the physical systems/subsystems to enable real time condition monitoring and improve the decision making. According to Liu et al. (2018), the connection between the cyber space and physical space is made by using a cyber physical interface (CPI) which makes it possible to have a cyber twin model. Advanced signal processing is employed for failure prediction and prevention, which is robust and capable of incorporating domain expertise. Heterogeneous data is made into a structured entity for further processing, through a concept called 'time machine', where the snapshots of data is taken only at discrete time. A method called adaptive clustering is used to structure data into similar categories to generate the awareness about the physical system status. Edge computing is employed, which extends the traditional cloud-based computing to the edge of the network, as it is suitable to the transport sector. In edge computing, the less complex / urgent activities are performed at the edge and less urgent / resource demanding processes are performed in the cloud. A peer to peer comparison step helps to prioritize maintenance activities within the fleet where there are many networked machines. (Liu, et al., 2018)

Xu et al. (2013) argues that the method with highest probability to be able to solve issues in reliability, maintainability, and availability of systems, is Prognosis and Health Management (PHM). PHM includes the functions, condition monitoring, health assessment, faults diagnostics, failure progression analysis, prognosis, and maintenance decision support. According to Xu et al. (2013), within the aircraft industry, the prognosis, and the estimation of the systems Remaining Useful Life (RUL) is the most important task of PHM. Xu et al. (2013) believes that for a complex system, such as an aircraft engine, for prognostics and estimation of the RUL, it is not suitable to use a model-based approach owing to the non-linearity and difficulty in representing the behaviour of the system through analytical models. However, neither using experience-based methods based on previous analysis, or data-based methods alone is according to Xu et al. (2013) enough. Rather, it requires an integrated approach of these methods.

Brown et al. (2007) present the concept and methodology used in the joint strike fighter air vehicles for health management and prognosis in the system. The concept by Brown et al. (2007) is based on using an on-board system on the vehicle and an off-board system which is based on data gathered from the on-board system. The off-board system is also integrated with the maintenance and logistics systems to seamlessly be able to produce work orders which minimizes the systems downtime based on the data. The onboard system for PHM is developed to, based on sensor data, automatically identify and isolate faults in the system which are critical to system function (Brown, et al., 2007). Through the integration with behavioural models of the system, the onboard PHM system can track fault and events in the system to each component. The off-board system handles noncritical functions and has a larger focus on making improvements to the system. According to Brown et al. (2007), the collaboration between the on-board and off-board systems, coupled with the maintenance and logistics systems, allows for effective health management of the system. The integration with data analytics methods in the on and off-board systems, together with a systems model, will according to Brown et al. (2007), increase the diagnostics capabilities of the system, as well as, provide better prognosis capabilities of the RUL and the assessment of component conditions through modelling of fault progression.

Elattar et al. (2018) discusses about the prognostic health monitoring using data driven methods through a case study of aircraft turbofan engine. Data driven method is used to calculate the useful remaining life of the engine in this case from the data collected using several sensors. The model-based approach uses mathematical methods to study the physics of the failure mode to predict the wear and tear in the system. However, the availability of high-fidelity models is very important, and it is costly and time intensive activity to create them. In comparison, data driven methods using historical data can create models for predicting the remaining useful life (RUL). There are offline and online approaches to the calculation of RUL (Elattar, et al., 2018). Offline approaches are used when the computation is resource intensive and cannot be done onboard. Collecting data from sensors and predicting RUL in real time is resource heavy and hence a challenge if it is to be incorporated onboard due to the intensive resource requirements. Though model-based prognostics have their advantages, creating the model of a complex system is according to Elattar et al. (2018) difficult. Data driven prognostic model is resource intensive and requires a robust algorithm to process the huge amount of data to give accurate RUL estimation. It should be noted that the data used in training stage and testing stage are different. In the testing stage the accuracies of HI prediction and RUL prediction are evaluated. The last stage consists of implementation of the algorithm and its deployment on Raspberry Pi 2, single board computer. Elattar et al. (2018) shows through this case how this method can give accurate results without being heavy on computational resources which is beneficial for real time onboard applications.

Sankavaram et al. (2009) Propose a process for diagnostics and prognostics, specifically with focus on automotive and electronic systems. The process uses both model-based, and data driven methods for determining there are RUL of the system. The general methodology of Sankavaram et al. (2009) consist of 6 steps; modelling the system with both graphical and analytical models; sensing, to ensure accurate diagnostics and prognostics, develop different test procedures and updating the tests with feedback to the system model; infer, Where different sensors/reasoners integrate; adaptive learning where the model is updated with new faults from the infer stage, and finally; predict, where the remaining in service life is determined. According to Sankavaram et al. (2009), the methodology is made to have the ability to experiment with either model based, or data driven methods, especially in the predict stage.

The main findings of this section are summarised in the table below.

Key findings
• Hypothetical methods are based on physics of failure / degradation and are weak in handling unknown failure modes. • Data driven methods are gaining popularity in predicting Remaining Useful Life (RUL) of systems. • Data driven methods are resource intensive. • Hybrid methods can combine advantages of physics based and data driven methods.

Table 17. Key findings from data driven monitoring and faults diagnostics

5. Integrated analysis

The purpose of the integrated analysis is to use aspects from different categories of the descriptive analysis to integrate and contrast different viewpoints. While there was no focus on from which category the arguments and points where taken from, it was the goal to attempt and include aspects of different perspectives in the integrated analysis. While all articles analysed in the descriptive framework will not make an appearance in the integrated analysis, all articles helped add towards an understanding, validation and holistic view of the different concepts and methodologies on the three main topics, complex products and systems, model-based systems engineering and data driven methods.

5.1. MBSE, Data driven methods and CoPS innovation

MBSE has emerged as a suitable methodology to develop CoPS as it provides a holistic view of the system, captures the system characteristics using models, enables traceability and handles the whole system development life cycle (Ramos, et al., 2012). By using the 'system model' MBSE captures the system requirements and subsequently aids in their decomposition to the lower levels, till the component level, while helping in finding design alternatives (Estefan, 2008).

However, the findings of the literature review highlight some limitations of MBSE. When it comes to data driven methods, the findings suggest several challenges and opportunities. The pros and cons associated with both the methods can influence their integration and the role they play in the innovation life cycle of CoPS. These issues will be discussed in the following sections.

5.1.1 Future challenges of MBSE.

With the use of MBSE methodologies, CoPS development has become more efficient when compared to the document-based methods. Though many authors highlight the benefits of applying MBSE methodology (Eg. Kaslow et al. (2017), Aleina et al. (2016), Fusaro et al. (2017) and Claver et al. (2014)), it faces several challenges as the complexities increase in CoPS development. According to Do and Cook (2014) there is a need for integration between tools. This integration is vital for building more executable models. Fisher et al. (2014), Bajaj et al. (2016) and Li et al. (2019) also highlight that the integration of models and tools is among one of the key challenges due to the multidisciplinary nature of CoPS development. Lindblad et al. (2018)

highlights the lack of a structured data base might be limiting the ability of model-based methods to look at optimal solutions using engineering data and optimization tools. Rhodes and Ross (2010) argue that MBSE addresses only the structural and behavioural aspects of CoPS. To capture the dynamic nature of system and uncertainties that it faces, additional - contextual, temporal and perceptual- dimensions need to be considered (Rhodes & Ross, 2010).

As CoPS become more complex, there is a strong indication that MBSE needs to evolve from its present state. According to Huang et al. (2011), it takes time, effort, and expertise to build validated models. Mount et al. (2016) found that the hypothetical knowledge has limitations when the complexity increased. According to Opiyo et al. (2015), the prediction of several key performance aspects of the system becomes difficult as the system becomes more complex. Bach et al. (2017) observe that verification and validation can become difficult in such a situation. According to Li et al. (2019), the V&V still is a challenge for MBSE due to lack of verification of inconsistencies and accurate simulations, and that the simulations are still carried out in separate domains, creating dependency in physical tests. Schluse et al. (2017) and Fisher et al. (2014) also observe that simulations of various domains are not well integrated which affects the verification and validation. Zeigler et al. (2018) argue that simulations have limitations in predicting the emergent behaviour seen in CoPS and that integration of big data with models and experimentation could improve this ability. Li et al. (2019) also highlight the need to incorporate data through the 'triple v' model.

Several authors propose the need for an integrated model that can manage the model heterogeneity and model life cycle (E.g. Bajaj et al. (2016); Fisher et al. (2014); Gausemeier et al. (2013)). This is comparable to the 'central model' proposed by Ogren (2000). Although the concept looks ideal, the benefits arising from the integration of models is debatable as it still has weakness associated with models-based methods. According to Li et al. (2019), because the models are developed in specific domains, they inherit domain specific characters which makes it difficult to integrate them. Also, the lack of a robust API or standard/ custom meta models that can interface with tools, as highlighted by Fisher et al. (2014), limit the scope of tool integration in modelling and simulation. These issues are reflected in the observations of Bajaj et al. (2016) as potential tensions while creating an integrated model, the most important one being the need to have flexibility of a multi-disciplinary approach while maintaining the rigidity of a complete system model. Though Bajaj et al. (2016) proposes the 'Total System Model' as a possible way to integrate the models using SysML language as the key integrator, the inherent weaknesses of SysML could affect Total System Model. According to Do and Cook (2014), SysML has weaknesses related to tool integration, design and analysis, and traceability. Given these findings, accomplishing a complete integration of tools and models appear to be difficult.

5.1.2 Data driven methods and MBSE

The challenges faced by MBSE in the face of increased complexity highlight its need for an upgrade and, as per many authors, data driven methods could potentially complement MBSE. The data driven method proposed by Wang (2010) in the case of intelligent management systems can be considered as an early step in the direction of maintaining a virtual representative of the system, providing multiple solutions to complex issues. The concept of digital twins, as discussed by Madni et al. (2019) and Schluse et al. (2017), highlight the new possibilities of bringing models and data together while creating synergy between simulations. It can also make use of environmental data to bring the simulations closer to the real-world scenario. Data driven methods can increase the feedback and according to Bach et al. (2017), they can help in identifying previously unknown relationships. According to Nightingale (2000), one of the main reasons for uncertainties in CoPS development is the inability to predict the emergent properties. With the support of data driven methods, the increased feedback may result in uncertainties reduction. This can also address, to some extent, the additional dimensions – contextual, temporal, and perceptual - in CoPS that Ross and Rhodes (2010) discuss.

According to Ren et al. (2019) big data can influence the design and testing of components. A similar observation is made by Bach et al. (2017), that data driven methods can contribute in the design of system characteristics and identify previously unknown relationships. Opiyo (2015), in the pipeline concept aims to gain design related insights from the data analytics for product superiority. Ding et al. (2015), Huang et al. (2011) and Norman et al. (2018) highlight the benefit of data analytics and visualization in aiding the re-use of knowledge in the design and development activities. As data driven approaches are less dependent on expert knowledge as compared to model-based methods, they can be useful in situations when building a physical model is costly or challenging or time consuming (Villarejo, et al., 2016) (Elattar, et al., 2018). From the above discussions, it can be deduced that data driven methods help to bring specifications and requirements closer, at the same time improve the feedback loops and reduce the uncertainties.

Though data driven methods have benefits, several authors highlight their limitations. According to Galar et al. (2013), data driven methods build the relationships using the training data while Mount et al. (2016) and Xu et al. (2013) highlight the need of a large set of training data for the methods to be effective. Villarejo et al. (2016) add that most of the data driven methods are time invariant and depend on training data to overcome this limitation, whereas the physics-based methods are capable of handling time varying dynamic system analysis. In a situation where training data is not readily available or is very limited, the data driven methods can be less effective as they cannot learn from it. According to Hobday (1998), the degree of customisation in CoPS is high. A key question about the data driven methods, in this context, could be about the data availability, especially in the early developmental phases. It would be challenging to gather system specific data before the system is implemented. Hence, data driven methods can be viewed as a reactive way of gaining insights and may not be able to contribute to the initial phase of system development where important decisions regarding system architecture needs to be taken.

According to the opinion of several authors, instead of adopting a data driven method, a hybrid method is a better option. Bach et al. (2017) suggests that the use of real-world data for testing and for contextual purposes can complement existing model-based methods. Uraikul et al. (2015) and Xu et al. (2013) consider the integration of methods as a beneficial. Liu and Goebel (2018) and Mount et al. (2016) found that when physics-based methods and data driven methods were combined, the learning and prediction capabilities increased, while the training cost of data driven methods reduced. Brown et al. (2007), Galar et al. (2013), Galar et al. (2015) and Villarejo et al. (2016) highlight that integrated methods / hybrid methods are effective in prognostics and in RUL assessment. Hence it can be deduced from the literature that as stand-alone method, data driven methods have their limitations, but they can be complimentary to the model-based methods.

5.1.3. Innovation in CoPS and role of data driven and model-based methods

According to Nightingale (2000), the product characteristics affect the innovation process in CoPS. Innovation life cycle in CoPS is different as compared to the mass-produced goods and so are the CoPS design rules and decision criteria (Bonaccorsi & Giuri, 2000) (Hobday, et al., 2000). Hobday (1998) attributed the complexity of CoPS development to its architecture and hierarchical product structure and highlighted that the ability to manage architectural changes is crucial. Hobday (1998) also observes that due to the complexity, the system architecture can have many design alternatives. An important capability of MBSE is the ability to capture the system requirements from the business, mission and stakeholder requirements and construct the architectural model, from which the behavioural and structural models are developed (Estefan, 2008) (Madni & Sievers, 2018). This aids in providing a holistic system view as well as the traceability within the system. According to Davies (1997), architectural innovation is one important type in the CoPS innovation, which concerns the interconnection of subsystems and components and their functions. There is a strong indication from the literatures that in the architectural innovation, model-based methods could be vital. As the system specific data may not be ready at the early phases of development, the influence of data driven methods could be limited.

The component and systemic innovation, as per Davies (1997), are limited to changes in the technical aspects at component level and functional aspect in a subsystem level, respectively. As CoPS tend to show nonlinear behaviour from one generation to other, the impact of change of one part/subsystem can be profound on the connected systems (Hobday & Rush, 1999). Data driven methods are shown to contribute to this type of innovation as they gather information that can be used to gain useful insights for the design and development, as discussed by Opiyo (2015) and Brown et al. (2017).

From the findings, the data driven / integrated methods contribute in predictive health and maintenance activities which are important in the operational phase of CoPS. According to Hobday et al. (2000), the product life cycle lasts for a very long period for CoPS and the innovation continues even after the CoPS becomes operational. The data driven methods and the integrated methods have the potential to support the continued innovation through the insights gained from data analysis. This is shown by Ren et al. (2019) on how the data can aid in conditional monitoring, potential failure detection, economic tool replacement, etc. Xu et al. (2013) explains that PHM is crucial for availability of the systems as they do a variety of connected activities such as function-condition monitoring, fault prediction, failure progression analysis, decision support etcetera. As CoPS are business critical for the customers according to Hobday (2000), the PHM activities are crucial for their functioning and profitability. Though data driven methods have limitations to influence the architectural phase, in the subsequent systemic, component as well as the continued innovation, they can be valuable. Thus, having the integration of model based and data driven methods can be beneficial to the innovation life cycle of CoPS as the methods can complement each other.

5.2. Management implications for systems integrators

The role of the systems integrator is twofold. They must first create the network of stakeholders that will develop the CoPS, and they must manage the tasks between them (Rutten, et al., 2009). To perform these tasks, the system integrator must have strong project capabilities (Davies & Brady, 2000) and an outward, market focus, while still maintaining technical capabilities (Liu & Su, 2014).

The integration of data-driven methods with MBSE will bring several managerial implications to system integrators. Both challenges and opportunities. Through the review of the literature, three different management areas in the development of CoPS where identified where the integration of data and models could potentially have an impact on current methods. These areas are general project management, decision making and project to project learning.

5.2.1. Project management

The findings indicate that the integration of data-driven methods with MBSE will not require a completely new project management approach. Rather, both Bach et al. (2017) and Li et al. (2019), suggest the continuous use of the 'V' model even with the increase of data-driven method, which is already an established model for CoPS development (Estefan, 2008). Although, there seem to be more emphasis on the iterative aspects with the integration of data driven methods. Hybertson et al. (2018) argue for increased iterations from the knowledge base and data analytics, to the derived solution, which could reduce uncertainties in development. Bach et al. (2017) argue for shorter iteration cycles, especially within software development and testing. However, there are also indications that an integration of methods could lead to an overall reduction of iterations in development. For example, the objective of CLOSE by Di Maio et al. (2018) is similar to Bach et al. (2017), to reduce cycle times, but also to decrease the amount of iterations. Bach et al. (2017b) also showed how, with the influence of data driven methods, they managed to reduce the number of real-world scenario testing. With the integration of model based and data driven methods, there appears to be potential for increased and/or shortened iteration cycles in certain parts of development. At the same time, there appears to be an opportunity for an overall decrease of iterations. Further studies are needed to establish the iterative aspect of an integrated method.

The iterative aspect of development phases could potentially draw project management closer to the "spiral method" described by Estefan (2008) and closer to an agile model, such as LITHE that was proposed by Ramos et al. (2013), along with an integration of data driven methods.

MBSE has so far proven to be a holistic method of developing complex systems (Ramos, et al., 2012), however, several limitations persist which hinders the methodology from representing all aspects of the system accurately (Fisher, et al., 2014). Several authors have mentioned how data driven methods can increase the holistic view of the system. Li et al. (2019) argue that the addition of a big data section to the 'V' model could better support a holistic representation in the development. Both Schluse et al. (2017) and Di Maio et al. (2018) suggest that the integration of simulation capabilities through tools such as digital twins, supports a holistic model of the system. It would therefore seem from the findings that data driven methods together with MBSE can add towards a more holistic representation of the system.

5.2.2. Decision-making

MBSE arguably supported managers in their decision making by showcasing a more holistic view of the system. By creating relationships between both the structural and the behavioural aspects of the system through a common systems model as described by Ogren (2000) and Ramos et al. (2012), both traceability and reliability of the development was increased. However, several points of criticism still exist for MBSE when it comes to managers abilities to use the methodological aspects for making decisions. The poor integration capabilities between domain knowledges identified in MBSE can lead to a lack of confidence in the decision making (Di Maio, et al., 2018). Developing verified models which are of high fidelity is also considered very time consuming, as well as requiring a high amount of expertise within the specific industry (Huang, et al., 2011) (Elattar, et al., 2018). MBSE has also shown to struggle in areas where there is low accumulated knowledge or in newly established fields (Li, et al., 2019). MBSE methodologies and the system model itself often becomes very complex to manage in the development of complex systems (Ding, et al., 2015) (Lindblad, et al., 2018).

Data driven methods can according to Liu & Goebel (2018) support finding new insights as well as aid manager in making better decisions. However, data driven methods still suffer some drawbacks. Data driven methods can become very resource intensive in terms of computing power required for them to operate (Elattar, et al., 2018). The data driven methods are often limited by the amount of available data to give an accurate result (Opiyo, 2015) (Villarejo, et al., 2016). Still, both Norman et al. (2018) and Qin (2014) suggests that data driven methods often decreases the decision-making time compared to model-based methods.

The findings therefore seem to support that data driven methods can enable managers to make timelier decision, based on a more accurate representation of the system. If there is an abundance of data, managers can base their decision on the data with support from the MBSE methodology and the system model. However, this is based on the computational resources and ability to gather, process, and analyse the data. If there is not a sufficient amount of data, decisions could be based on the MBSE methodology and system model, with support from the data driven methods. Although this would require higher knowledge and experience of the system. The need for such a large amount of data for data driven methods to be reliable may cause them to be less effective in earlier stages of development. As CoPS are heavily customized systems (Hobday, 1998), gathering data early in development could prove difficult. At later stages when more data can be gathered, especially through tools such as digital twins as suggested by Heber & Groll (2017), Madni et al., (2019), Schluse et al., (2017) and Di Maio et al., (2018), data driven methods could be more influential for decision making in CoPS development.

5.2.3. Project-to-project learning

The potential for an integrated methodology to feedback new knowledge and learnings from the data driven methods to MBSE and the system model, as suggested by Brown et al. (2007) and Sankavaram et al. (2009), could lead to higher potential for project-to-project learning. By utilizing the new knowledge from the data driven methods in the model-based methods, new generations of CoPS, or as previously mentioned, updates to the current systems, can be improved. This is especially important in CoPS as Hobday (2000) show that project based, or project lead organisational structures are the best suited structures for CoPS development.

The increased collection and use of contextual/environmental data as proposed by several authors (e.g. Bach et al., (2017); Ding et al., (2015); Galar et al., (2015); Villarejo et al., (2016)) could also have the potential to improve project-to-project learning. The findings from the contextual data could be utilized in new product generations. The data could as well potentially be used in other CoPS and potentially even in cross sector learning as several different CoPS and sectors could benefit from, for example, terrain data. However, the gathering and use of contextual data also places a need for system integrators to develop their project-to-project learning capabilities to achieve the potential benefits from an integrated method.

The feedback from data driven methods to model-based methods, with the increase of contextual and environmental data, could better support economics of repetitions which Davies & Brady (2000) argue, is how organisations improve their ability to execute CoPS projects. An integrated method could therefore better support CoPS projects and the potential learning that system integrators can gain from them.

5.3. Data as the baseline

The change in focus from models to data is promoted by several authors. By instead creating relationships between the data rather than the models, Lindblad et al. (2018) suggest that changes to the system will be more accurately portrayed and MBSE will be more than a “visualisation” of the system and will more accurately represent the real system. The concept by Hybertson et al. (2018) instead shifts focus to “evidence” produced through data as well as mode-based methods, achieving a more objective result. Li et al. (2019) suggest that a big data branch to the ‘V’ model would act as the connection between MBSE and the real system. Zhan et al. (2015) argue that a focus on the data elements of which the models are built on will give a more accurate representation of the system.

As Bonnet et al. (2015) argues that with the introduction of a new methodology, a cultural change is often important, as shown in the change from the document focused system development to MBSE. It is therefore not unlikely that a similar cultural change would take place with the implementation of a new methodology of data driven methods along with MBSE.

6. Conclusion

As the development of CoPS becomes more complex, the existing MBSE methodologies face new challenges. With the advent of data driven methods supported by technologies such as IoT and big data, new possibilities are emerging for developing CoPS. The findings of the literature review show certain limitations of model-based methods as well as data driven methods when used independently. As it takes time to develop validated models, decisions in the early phases may be delayed due to the dependency on model-based methods. The lack of integration among models / simulations of different disciplines can increase the chances of missing vital information. This can also affect the subsequent verification and validation phase. In data driven methods, lack of training data can reduce the usefulness of the method. Dependence on data driven methods solely without the expert knowledge can increase the associated risks.

Data driven methods are faster in comparison to model-based methods but in the early phases of development, when the availability of system specific data is often limited, the development depends mostly on model-based methods. In capturing the requirements and decomposing them to the component level designs, MBSE plays a crucial role. An integration of the methods can be advantageous to the development process, according to the analysis, where the methods can complement each other. Design as well as simulations are found to benefit from the data driven methods, due to the increased feedback as well as insights about the external conditions. This can improve the system reliability through improved verification and validation. It can also contribute to reduction in uncertainties. As model-based methods are weak in predicting emergent behaviour, integrating data driven methods can improve its ability to anticipate such behaviour.

An integrated method could better support a holistic decision making in CoPS development and, with the implementation of data driven methods, more timelier decisions. However, to fully utilize both methods, a high amount of expert knowledge is needed for the model-based methods. For the data driven methods, a substantial amount of data is needed to reliably base decision on them. In an integrated methodology, when either resource is missing, the other could be used as the main input for the decision, with support of the other.

Among the different types of innovation in CoPS, model-based methods play a major role in the architectural innovation while the data driven methods have the potential to contribute to the component and systemic innovations. The findings suggest that in the continued innovation, data driven methods play an important role by gathering insightful information during the use phase of the system. Considering the longer lifespan of CoPS, this is area where data driven methods can be very valuable. Integration of methods can aid in transferring the insights generated by data driven methods to the system model. Considering the whole innovation life cycle in CoPS, both model-based and data driven methods can complement each other, though their influence is varied in degree depending on the development phases.

The findings also seem to support that an integrated methodology could increase project-to-project learning capabilities in system integrators. This is especially important as the organisation structure of CoPS suppliers are predominantly project based. With the mentioned feedback of learnings from data driven methods in the use phase, to the model-based methods, system integrators have the opportunity to also transfer learnings to new projects, and, in doing so, build their economy of repetition. However, this will certainly create challenges for system integrators to develop their project-to-project learning capabilities, as well as identify how, and when in the development of new projects, to incorporate the data driven learnings with MBSE.

The general project management methodologies applied in CoPS with MBSE, do not appear to be substantially altered while applying an integrated methodology of MBSE and data driven methods. However, the project management process could become more iterative at certain stages than it currently is, more aligned with an agile methodology. System integrators may therefore need to develop their capabilities to handle shorter iteration cycles which could be a challenge specifically in a CoPS context where projects may last decades. Although, even with the increased capability challenges, the findings support that data driven, and model-based, methods can complement, rather than impede, each other in several aspects in the development of CoPS.

7. Future research areas

One aspect that the findings support is that in order to fully achieve the benefits from an integrated method, the models in MBSE need to be based on data. Several authors from the findings point towards this need and proposes different concepts on how to drive the models with data. However, more research is needed to identify how this could be achieved in practice.

Though large amounts of system specific data may not be available in the early phases of development, the project-to-project learning and accumulated data may still be useful in contributing to the development. However, the analysis shows a lack of research on how data driven methods can be utilized in early stages of development. More research is therefore needed to explore how effectively data driven methods can contribute to the early development phases as it can result in faster decision making.

Although the established project management processes may not drastically change, the integration of model-based and data driven method may require new ways of collaboration to facilitate innovation in CoPS development. This might require new ways of organizing the work and new types of organizational structures. Further research could look more in depth of how organisational and project structures could be optimised when using an integrated methodology.

The effect of iterations within the development of CoPS have been addressed by some authors pointing towards shorter and increased iteration cycles at different part of the process. However, authors also suggest that the integration of methods could lead to an overall reduction of iteration. More research is needed to identify the impacts of iteration in certain stages, and overall, in CoPS development.

Data driven methods are shown to be complementing the model-based methods. However, the impact of incorporating them on the system development cost needs to be studied along with the overall system life cycle cost. It would be interesting to see if the integrated methods result in a higher or lower development cost as compared to the present developmental costs. More research is required to find out the economic benefits of the integrated methods.

Though this thesis has added toward the conceptual development phase of Lynham's general method of theory framework by identifying the key elements and their relations, more research is needed to achieve a fully informed conceptual model. Further research could also start looking towards the second phase of the framework, operationalize, where the focus instead would lie on establishing a connection of the conceptual aspects, and practices within CoPS.

References

Acha, V., Davies, A., Hobday, M. & Salter, A. J., 2004. Exploring the Capital Goods Economy: Complex Product Systems in the UK. *Industrial and Corporate Change,* 13(3), p. 505–529.

Aleina, S. C., Ferretto, D., Stesina, F. & Viola, N., 2016. *A Model-Based Approach to the Preliminary Design of a Space Tug Aimed at Early Requirement's Verification.* Mexico, International Astronautical Federation, IAF.

Anderson, P. & Tushman, M. L., 1990. Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change. *Administrative Science Quarterly,* 35(4), pp. 604-633.

Bach, J. et al., 2017. *Data-driven Development, a Complementing Approach for Automotive Systems Engineering.* Vienna, IEEE.

Bach, J. et al., 2017b. *Test Scenario Selection for System-level Verification and Validation of Geolocation-dependent Automotive Control Systems.* Funchal, IEEE.

Bajaj, M. et al., 2016. MBSE++ — Foundations for Extended Model-Based Systems Engineering Across System Lifecycle. *INCOSE International Symposium,* 26(1), pp. 2429-2445.

Bell, E., Bryman, A. & Harley, B., 2017. *Business Research Methods.* 5 ed. Oxford: Oxford University Press.

Bonaccorsi, A. & Giuri, P., 2000. When shakeout doesn't occur - The evolution of the turboprop industry. *Research policy,* 29(7-8), pp. 847-870.

Bonnet, S., Voirin, J.-L., Normad, V. & Exertier, D., 2015. Implementing the MBSE Cultural Change: Organization, Coaching and Lessons Learned. *INCOSE International Symposium,* 25(1), pp. 508-523.

Brown, E. R., McCollom, N. N., Moore, E.-E. & Hess, A., 2007. *Prognostics and Health Management A Data-Driven Approach to Supporting the F-35 Lightning II.* Big Sky, IEEE.

Clark, K. B. & Wheelwright, S. C., 1992. Organizing and Leading "Heavyweight" Development Teams. *Californian Management Review,* 34(3), pp. 9-28.

Claver, C. F. et al., 2014. *Systems Engineering in the Large Synoptic Survey Telescope Project: An Application of Model Based Systems Engineering.* Montréal, Modeling, Systems Engineering, and Project Management for Astronomy VI.

Dai , X. & Gao, Z., 2013. From Model, Signal to Knowledge: A Data-Driven Perspective of Fault Detection and Diagnosis. *IEEE Transactions on Industrial Informatics,* 9(4), pp. 2226-2238.

Davies, A., 1997. The Life Cycle of a Complex Product System. *International Journal of Innovation Management,* 1(3), pp. 229-256.

Davies, A. & Brady, T., 2000. Organisational capabilities and learning in complex product systems: towards repeatable solutions. *Research policy,* 29(7-8), pp. 931-953.

Davies, A., Brady, T., Prencipe, A. & Hobday, M., 2011. Innovation in Complex Products and Systems: Implications for Project Based Organizing.. In: G. Chattani, S. Ferriani, L. Frederiksen & F. Täube, eds. *Project-Based Organizing and Strategic Management (Advances in Strategic Management, Vol.8).* Bingley: Emerald Group Puplishing Limited, pp. 3-26.

Dedehayir, O., Nokelainen, T. & Mäkinen, S. J., 2014. Disruptive innovations in complex product systems industries : A case study. *Journal of Engineering and Technology Management,* Volume 33, pp. 174-192.

Di Maio, M. et al., 2018. *Closed-Loop Systems Engineering (CLOSE): Integrating Experimentable Digital Twins with the Model-Driven Engineering Process.* Rome, IEEE.

Ding, Y., Tang, J. & Huang, J. Z., 2015. *Data Analytics Methods for Wind Energy Applications.* Montreal, International Gas Turbine Institute.

Do, Q. & Cook, S., 2012. *An MBSE Case Study and Research Challenges.* Rome, Italy, INCOSE, pp. 1531-1543.

Elattar, H. M., Elminir, H. K. & Riad, A. M., 2018. Conception and implementation of a data-driven prognostics algorithm for safety–critical systems. *Soft Computing,* Volume 23, pp. 3365-3382.

Estefan, J. A., 2008. *Survey of Model-Based Systems Engineering (MBSE) Methodologies,* California: INCOSE MBSE Initiative.

Falagas, M. E., Eleni, I. P., Malietzis, G. A. & Pappas, G., 2008. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strength and Weaknesses. *The FASEB Journal,* Volume 22, pp. 338-342.

Fisher, A. et al., 2014. *Model lifecycle management for MBSE.* Las Vegas, NV, USA, INCOSE, pp. 207-229.

Freidenthal, S., Steiner, R. & Moore, A., 2014. *A Practical Guide to SysML : The Systems Modeling Language.* 3 ed. Waltham: Elsevier.

Fusaro, R., Ferretto, D. & Viola, N., 2017. *MBSE Approach to Support and Formalize Mission Alternatives Generation and Selection Processes for Hypersonic and Suborbital Transportation Systems.* Vienna, IEEE.

Galar, D., Kumar, U., Villarejo, R. & Johansson, C.-A., 2013. Hybrid prognosis for railway health assessment: An information fusion approach for PHM deployment. *Chemical Engineering Transactions,* Volume 33, pp. 769-774.

Galar, D., Thaduri, A., Catelani, M. & Ciani, L., 2015. Context Awareness for Maintenence Decision Making: A Diagnosis and Prognosis Approach. *Measurement,* Volume 67, pp. 137-150.

Gann, D. M. & Salter, J. A., 2000. Innovation in project-based, service-enhanced firms: the construction of complex products and systems. *Research Policy,* 29(7), pp. 955-972.

Gao, Z., Saxén, H. & Gao, C., 2013. Guest editorial: Special section on data-driven approaches for complex industrial systems. *IEEE transactions on industrial informatics,* 9(4), pp. 2210-2212.

Gausemeier, J., Gaukstern, T. & Christian, T., 2013. *Systems Engineering Management Based on a Discipline-Spanning System Model.* Atlanta, Procedia Computer Science.

Geffner, H., 2018. *Model-free, model-based, and general intelligence.* Stockholm, IJCAI, pp. 10-17.

Geraldi, J. G., 2009. Reconciling order and chaos in multi-project firms. *International Journal of Managing Projects in Business,* 2(1), pp. 149-158.

Gubbi, J., Buyya, R., Marusic, S. & Palaniswami, M., 2013. Internet of Things (IoT): A vision, architectural elements and future directions. *Future Generation Computer Systems,* 29(7), pp. 1645-1660.

Hansen, K. L. & Rush, H., 1998. Hotspots in complex product systems: Emerging issues in innovation management. *Technovation,* 18(8/9), pp. 555-561.

Hardstone, G. A. P., 2004. Capabilities, structures and strategies re-examined: incumbent firms and the emergence of complex product systems (CoPS) in mature industries. *Technology analysis & strategic management,* 16(2), pp. 179-196.

Heber, D. & Groll, M., 2017. *Towards a digital twin: How the blockchain can foster E/E-traceability in consideration of model-based systems engineering.* Vancouver, Canada, Proceedings of the 21st International Conference on Engineering Design (ICED17).

Herzig, S. J., Qamar, A. & Paredis, C. J., 2014. Approach to Identifying Inconsistencies in Model-based Systems Engineering. *Procedia Computer Science,* Volume 28, pp. 354-362.

Hobday, M., 1998. Product complexity, innovation and industrial organisation. *Research Policy,* 26(6), pp. 689-710.

Hobday, M., 2000. The project-based organisation: an ideal form for managing complex products and systems?. *Research Policy,* 29(7-8), pp. 871-893.

Hobday, M. & Rush, H., 1999. Technology management in complex product systems (CoPS)-ten questions answered. *International Journal of Technology Management,* 17(6), pp. 618-638.

Hobday, M., Rush, H. & Tidd, J., 2000. Innovation in complex products and systems. *Research Policy,* 29(7-8), pp. 793-804.

Holland, O. T., 2015. Model-Based Systems Engineering. In: M. L. Loper, ed. *Modeling and Simulation in the Systems Engineering Life Cycle: Core Concepts and Accompanying Lectures.* 1 ed. London: Springer, pp. 299-306.

How, D. N., Hannan, M., Lipu, H. M. & Ker, P. J., 2019. State of Charge Estimation for Lithium-Ion Batteries Using Model-Based and Data-Driven Methods : A Review. *IEEE,* 7(IEEE Access), pp. 136116-136136.

Huang, Y., Seck, M. D. & Verbraeck, A., 2011. *From data to simulation models: component-based model generation with a data-driven approach.* Phoenix, Proceedings of the 2011 Winter Simulation Conference (WSC), pp. 3719-3729.

Huenteler, J., Schmidt, T. S., Ossenbrink, J. & Hoffmann, H. V., 2016. Technology life-cycles in the energy sector — Technological characteristics and the role of deployment for innovation. *Technological Forecasting and Social Change,* Volume 104, pp. 102-121.

Hu, H., Wen, Y., Chua, T.-S. & Li, X., 2014. Toward Scalable Systems for Big Data Analytics: A Technology Tutorial. *IEEE Access,* Volume 2, pp. 652-687.

Hybertson, D. et al., 2018. Evidence-based systems engineering. *Systems engineering,* 21(3), pp. 243-258.

INCOSE, 2015. *INCOSE Systems Engineering Handbook: A Guide for Systems Life Cycle Processes and Activities.* 4 ed. Hoboken, New Jersey: John Wiley & Sons, Incorporated.

Jabareen, Y., 2009. Building a Conceptual Framework: Philosophy, Definitions, and Procedure. *International Journal of Qualitative Methods,* 8(4), pp. 49-62.

Kaslow, D. et al., 2017. *A Model-Based Systems Engineering (MBSE) approach for defining the behaviors of CubeSats.* Montana, IEEE Aerospace Conference.

Lee, C., Yeung, C. & Cheng, M., 2015. *Research on IoT Based Cyber Physical System for Industrial Big Data Analytics*. Singapore, IEEE International conference on Industrial Engineering and Engineering Management.

Lehtinen, J., Aaltonen, K. & Rajala, R., 2019. Stakeholder Management in Complex Product Systems: Practices and Rationales for Engagement and Disengagement. *Industrial Marketing Management,* Volume 79, pp. 58-70.

Lindblad, L., Witzmann, M. & Bussche, S. V., 2018. *Data-driven systems engineering: Turning MBSE into industrial reality.* Glasgow, The european space agency & University of Strathclyde.

Li, Q., Wei, H., Yu, C. & Wang, S., 2019. *Model and data driven complex product development: from V, double Vs to triple Vs.* Chongqing, IEEE.

Liu, J. & Su, J., 2014. Market orientation, technology orientation and product innovation success: insights from CoPS. *International journal of innovation management,* 18(4), p. 1450020 (25 pages).

Liu, Y. & Goebel, K., 2018. *Information fusion for national airspace system prognostics: A NASA ULI project.* Philadelphia, Annual Conference of the PHM Society.

Liu, Z. et al., 2018. *Industrial AI Enabled Prognostics for High-speed Railway Systems.* Seattle, IEEE, pp. 1-8.

Lynham, S. A., 2002. The General Method of Theory-Building Research in Applied Disciplines. *Advances in Developing Human Resources,* 4(3), pp. 221-241.

Madni, A. M., Madni, C. C. & Lucero, S. D., 2019. Leveraging Digital Twin Technology in Model-Based Systems Engineering. *Systems,* 7(7).

Madni, A. M. & Sievers, M., 2018. Model-based Systems Engineering: Motivation, Current Status, and Future opportunities. *Systems Engieering,* 21(3), pp. 172-190.

Marjani, M. et al., 2017. Big IOT Data Analytics: Architecture, Opportunities, and Open Research Challenges. *IEEE Access,* Volume 5, pp. 5247-5261.

Mauro, A. D., Greco, M. & Grimaldi, M., 2016. A formal definition of big data based on its essential features. *Library review,* 65(3), pp. 122-135.

Miller, R., Hobday, M., Leroux-Demers, T. & Olleros, X., 1995. Innovation in Complex Systems Industries: The Case of Flight Simulation. *Industrial and Corporate Change,* 4(2), pp. 363-400.

Moody, J. B. & Dodgson, M., 2006. Managing Complex Collaborative Projects: Lessons from the Development of a New Satellite. *Journal of Technology Transfer,* Volume 31, pp. 567-588.

Mosallam, A., Medjaher , K. & Zerhouni, N., 2015. *Component based data-driven prognostics for complex systems: Methodology and applications.* Beijing, IEEE, pp. 1-7.

Mount, N. J. et al., 2016. Data-driven modelling approaches for socio-hydrology: opportunities and challenges within the Panta Rhei Science Plan. *Hydrological Sciences Journal,* 61(7), pp. 1192-1208.

Naghizadeh, M., Manteghi, M., Ranga, M. & Naghizadeh, R., 2017. Managing integration in complex product systems: The experience of the IR-150 aircraft design program. *Technological Forecasting and Social Change,* Volume 122, pp. 252-261.

Nightingale, P., 2000. The product-process-organisation relationship in complex development projects. *Research policy,* 29(7-8), pp. 913-930.

Norman, R. et al., 2018. *Using Big Data Analytics to Create a Predictive Model for Joint Strike Fighter.* Seattle, IEEE, pp. 3590-3596.

Ogren, I., 2000. On principles for model-based systems engineering. *Systems engineering,* 3(1), pp. 38-49.

Opiyo, E. Z., 2015. *Data analytics pipeline for prediction and decision making in complex products and systems development.* Houston, Proceedings of the ASME 2015 International Mechanical Engineering Congress and Exposition.

Park , T.-Y. & Kim, J.-Y., 2014. The capabilities required for being successful in complex product systems: case study of Korean e-government. *Asian Journal of Technology Innovation,* 22(2), pp. 268-285.

Peltoniemi, M., 2011. Reviewing Industry Life-Cycle Theory : Avenues for Future Research. *Internatinonal Journal of Management Reviews,* 13(4), pp. 349-375.

Qin, J. S., 2014. Process Data Analytics in the Era of Big Data. *AIChE Journal,* 60(9), pp. 3092-3100.

Ramos, A. L., Ferreira, J. V. & Barceló, J., 2012. Model-based systems engineering: An emerging approach for modern systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* 42(1), pp. 101-111.

Ramos, A. L., Ferreira, V. J. & Barceló, J., 2013. LITHE: An Agile Methodology for Human-Centric Model-Based Systems Engineering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* 43(3), pp. 504-521.

Ranjbar, M. S., Park, T.-Y. & Kiamehr, M., 2018. What happened to complex product systems literature over the last two decades: Progress so far and path ahead. *Technology Analysis & Strategic Management,* 30(8), pp. 948-966.

Ren, Y. et al., 2019. The Application of Downhole Vibration Factor in Drilling Tool Reliability Big Data Analytics - A Review. *ASME Journal of Risk and Uncertainitiy in Engieering Systems. Risk Uncertainty Part B : Mechanical Engineering,* 5(1), p. 010801.

Ren, Y.-T. & Yeo, K.-T., 2006. Research challenges on complex product systems (CoPS) innovation. *Journal of the Chinese Institute of Industrial Engineers,* 23(6), pp. 519-529.

Rhodes, D. H. & Ross, A. M., 2010. *Five Aspects of Engineering Complex Systems Emerging Constructs and Methods.* Sand Diego, CA, IEEE International Systems Conference.

Roehrich, J. K., Davies, A., Frederiksen, L. & Sergeeva, N., 2019. Management Innovation in Complex Products and Systems: The Case of Integrated Project Teams. *Industrial Marketing Management,* Volume 79, pp. 83-93.

Rush, H., 1997. *Managing Innovation in Complex Product Systems (CoPS).* London, UK, IEE Colloquium on EPSRC Technology Management Initiative (Engineering and Physical Sciences Research Council).

Rutten, M. E., Dorée, A. G. & Halman, J. I., 2009. Innovation and interorganizational cooperation: a synthesis of literature. *Construction innovation,* 9(3), pp. 285-297.

Sankavaram, C. et al., 2009. *Model-based and Data-driven Prognosis of Automotive and Electronic Systems.* Bangalore, IEEE.

Sargent, R. G., 2015. Types of Models. In: M. L. Loper, ed. *Modeling and Simulation in the Systems Engineering Life Cycle: Core Concepts and Accompanying Lectures.* 1 ed. London: Springer, pp. 51-55.

Saunders, M., Lewis, P. & Thornhill, A., 2019. *Research Methods For Business Students.* 8th ed. Harlow: Pearson Education Limited.

Schluse, M., Atorf, L. & Rossmann, J., 2017. *Experimentable digital twins for model-based systems engineering and simulation-based development.* Montreal, IEEE.

Snyder, H., 2019. Literature review as a research methodology: An overview and guidelines. *Journal of Business Research,* Volume 104, pp. 333-339.

Storberg-Walker, J., 2003. Comparison of the Dubin, Lynham, and Van de Ven Theory-Building Research Methods and Implications for HRD. *Human Resource Development Review,* 2(2), pp. 211-222.

Tonnquist, B., 2012. *Project Management.* 2nd ed. Stockholm: Sanoma Utbildning AB.

Torraco, R. J., 2016. Writing Integrative Literature Reviews: Using the Past and Present to Explore the Future. *Human Resource Development Review,* 15(4), pp. 404-428.

Uraikul, V., Chan, C. W. & Tontiwachwuthikul, P., 2007. Artificial Intelligence for Monitoring and Supervisory Control of Process Systems. *Engineering Applications of Artificial Intelligence,* 20(2), pp. 115-131.

Villarejo, R. et al., 2016. Context-driven Decisions for Railway Maintenance. *Proceedings of the Institution of Mechanical Engineers,* 230(5), pp. 1469-1483.

Wang, F.-Y., 2010. Parallel Control and Management for Intelligent Transportation Systems: Concepts, Architecture, and Applications. *IEEE Transaction on Intelligent Transportation Systems,* 11(3), pp. 630-638.

Whittemore, R. & Knafl, K., 2005. The integrative review: updated methodology. *Journal of Advanced Nursing,* 52(5), pp. 546-553.

Xu, J., Wang, Y. & Xu, L., 2013. PHM-Oriented Integrated Fusion Prognostics for Aircraft Engines Based on Sensor Data. *IEEE Sensors Journal,* 14(4), pp. 1124-1132.

Zeigler, B. P., Mittal, S. & Traore, M. K., 2018. MBSE with/out Simulation: State of the Art and Way Forward. *Systems,* 6(4), p. 40.

Zhan, G., Ge, B., Li, M. & Yang, K., 2015. A Data-Centric Approach for Model-Based Systems Engineering. *Journal of systems science and information,* 3(6), pp. 549-560.

Zhang, W. & Igel, B., 2001. Managing the product development of China's SPC switch industry as an example of CoPS. *Technovation,* 21(6), pp. 361-368.

Appendix A – Articles included in the literature review

Year	Authors	Title	Category	Citations
1997	Rush, H.	Managing innovation in Complex Product Systems (CoPS)	Key capabilities in Complex Products and Systems	4
1997	Davies, A.	The life cycle of a complex product system	Innovation life cycle of Complex Products and Systems	67
1998	Hobday, M.	Product complexity, innovation and industrial organisation	Characteristics of Complex Products and Systems	584
1998	Hansen, K. L. & Rush, H.	Hotspots in complex product systems: Emerging issues in innovation management	Characteristics of Complex Products and Systems	57
1999	Hobday, M. & Rush, H.,	Technology management in complex product systems (CoPS)-ten questions answered	Characteristics of Complex Products and Systems	193
2000	Hobday, M.	The project-based organisation: An ideal form for managing complex products and systems?	Organization structure in Complex Products and Systems	703
2000	Gann, D. M. & Salter, J. A.	Innovation in project-based, service-enhanced firms: The construction of complex products and systems	Key capabilities in Complex Products and Systems	725
2000	Davies, A. & Brady, T.	Organisational capabilities and learning in complex product systems: Towards repeatable solutions	Key capabilities in Complex Products and Systems	356
2000	Nightingale, P.	The product-process-organisation relationship in complex development projects	Key capabilities in Complex Products and Systems	92
2000	Hobday, M. Rush, H. & Tidd, J.	Innovation in complex products and systems	Innovation life cycle of Complex Products and Systems	343
2000	Bonaccorsi, A. & Giuri, P.	When shakeout doesn't occur – The evolution of the turboprop industry	Innovation life cycle of Complex Products and Systems	44
2000	Ogren, I.	On principles for model-based systems engineering	Traditional Model-based systems engineering	28
2001	Zhang, W. & Igel, B.	Managing the product development of China's SPC switch industry as an example of CoPS	Key capabilities in Complex Products and Systems	11
2004	Acha, V, Davies, A. Hobday, M. & Salter, A. J.	Exploring the capital goods economy: Complex product systems in the UK	Characteristics of Complex Products and Systems	66

2004	Hardstone, G. A. P.	Capabilities, structures and strategies re-examined: incumbent firms and the emergence of complex product systems (CoPS) in mature industries	Key capabilities in Complex Products and Systems	20
2006	Moody, J.B & Dodgson, M.	Managing complex collaborative projects: Lessons from the development of a new satellite	Organization structure in Complex Products and Systems	11
2007	Brown, E. R., McCollom, N. N. Moore, E.-E. & Hess, A.	Prognostics and health management a data-driven approach to supporting the F-35 lightning II	Data driven monitoring and fault diagnostics	34
2007	Uraikul, V. Chan, C. W. & Tontiwachwuthikul, P.	Artificial intelligence for monitoring and supervisory control of process systems	Data driven control and optimization	146
2008	Estefan, A. J.	Survey of model-based systems engineering (MBSE) methodologies	Traditional Model-based systems engineering	761
2009	Rutten, M. E. Dorée, A. G. & Halman, J. I.	Innovation and interorganizational cooperation: A synthesis of literature	Characteristics of Complex Products and Systems	59
2009	Geraldi, J.G.	Reconciling order and chaos in multi-project firms	Organization structure in Complex Products and Systems	24
2009	Sankavaram, C. Pattipati, B. Kodali, A. Pattipati, K. Azam, M. Kumar, S. & Pecht, M.	Model-based and data-driven prognosis of automotive and electronic systems	Data driven monitoring and fault diagnostics	47
2010	Rhodes, D. H. & Ross, A. M.	Five aspects of engineering complex systems emerging constructs and methods	Emerging cultural and technical challenges in Model-based systems engineering	24
2010	Wang, F.-Y.	Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications	Data driven control and optimization	485
2011	Davies, A. Brady, T. Prencipe, A. & Hobday, M.	Innovation in complex products and systems: Implications for project-based organizing	Organization structure in Complex Products and Systems	51
2011	Huang, Y. Seck, M. D. & Verbraeck, A.	From data to simulation models: component-based model generation with a data-driven approach	Data driven modelling	17
2012	Ramos, A. L., Ferreira, J. V. & Barceló, J.	Model-based systems engineering: An emerging approach for modern systems	Traditional Model-based systems engineering	109
2012	Do, Q. & Cook, S.	An MBSE Case Study and Research Challenges	Case studies on Model-based systems engineering	3
2013	Gausemeier, J. Gaukstern, T. & Christian, T.	Systems engineering management based on a discipline-spanning system model	Emerging cultural and technical challenges in Model-based systems engineering	13
2013	Ramos, A. L. Ferreira, V. J. & Barceló, J.	Lithe: An agile methodology for human-centric model-based systems engineering	Emerging cultural and technical challenges in Model-based systems engineering	11
2013	Dai , X. & Gao, Z.	From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis	Data driven monitoring and fault diagnostics	229

2013	Xu, J. Wang, Y. & Xu, L.	PHM-oriented integrated fusion prognostics for aircraft engines based on sensor data	Data driven monitoring and fault diagnostics	65
2013	Galar, D. Kumar, U. Villarejo, R. & Johansson, C.-A.	Hybrid prognosis for railway health assessment: An information fusion approach for PHM deployment	Data driven monitoring and fault diagnostics	15
2014	Liu, J. & Su, J.	Market orientation, technology orientation and product innovation success: Insights from CoPS	Key capabilities in Complex Products and Systems	8
2014	Park, T.-Y. & Kim, J.-Y.	The capabilities required for being successful in complex product systems: case study of Korean e-government	Key capabilities in Complex Products and Systems	8
2014	Dedehayir, O. Nokelainen, T. & Mäkinen, S. J.	Disruptive innovations in complex product systems industries: A case study	Innovation life cycle of Complex Products and Systems	34
2014	Fisher,A. Nolan, M. Friedenthal, S. Loeffler, M. Sampson, M. Bajaj, M. VanZandt, L. Hovey, K. Palmer, J. & Hart, L.	Model lifecycle management for MBSE	Emerging cultural and technical challenges in Model-based systems engineering	12
2014	Claver, C. F. Selvy, B. M. Angeli, G. Delgado, F. Dubois-Felsmann, G. Hascall, P. Lotz, P. Marshall, S. Schumacher, G. & Sebag, J.	Systems engineering in the Large Synoptic Survey Telescope project: An application of model-based systems engineering	Case studies on Model-based systems engineering	6
2014	Herzig, S. J. I. Qamar, A. & Paredis, C. J. J.	An approach to identifying inconsistencies in Model-Based Systems Engineering	Model-based systems engineering and data	26
2014	Qin, S. J.	Process data analytics in the era of big data	Data driven control and optimization	190
2015	Bonnet, S. Voirin, J.-L. Normand, V. & Exertier D.	Implementing the mbse cultural change: Organization, coaching and lessons learned	Emerging cultural and technical challenges in Model-based systems engineering	6
2015	Zhan, G. Ge, B. Li, M. & Yang, K.	A data-centric approach for model-based systems engineering	Model-based systems engineering and data	3
2015	Galar, D. Thaduri, A. Catelani, M. & Ciani, L.	Context awareness for maintenance decision making: A diagnosis and prognosis approach	Data driven monitoring and fault diagnostics	62
2015	Mosallam, A. Medjaher, K. & Zerhouni, N.	Component based data-driven prognostics for complex systems: Methodology and applications	Data driven monitoring and fault diagnostics	8
2015	Ding, Y. Tang, J. & Huang, J. Z.	Data analytics methods for wind energy applications	Data driven modelling	3
2015	Opiyo, E. Z.	Data analytics pipeline for prediction and decision making in complex products and systems development	Data driven modelling	1
2016	Huenteler, J. Schmidt, T. S. Ossenbrink, J. & Hoffmann, V. H.	Technology life-cycles in the energy sector - Technological characteristics and the role of deployment for innovation	Innovation life cycle of Complex Products and Systems	125

2016	Bajaj, M. Zwemer, D. Yntema, R. Phung, A. Kumar, A. Dwivedi, A. & Waikar, M.	MBSE++ — Foundations for Extended Model-Based Systems Engineering Across System Lifecycle	Emerging cultural and technical challenges in Model-based systems engineering	9
2016	Aleina, S. C. Ferretto, D. Stesina, F. & Viola, N.	A model-based approach to the preliminary design of a space tug aimed at early requirement's verification	Case studies on Model-based systems engineering	5
2016	Villarejo, R. Johansson, C.-A. Galar, D. Sandborn, P. & Kumar, U.	Context-driven decisions for railway maintenance	Data driven monitoring and fault diagnostics	2
2016	Mount, N. J. Maier, H. R. Toth, E. Elshorbagy, A. Solomatine, D. Chang, F.-J. & Abrahart, R. J.	Data-driven modelling approaches for socio-hydrology: opportunities and challenges within the Panta Rhei Science Plan	Data driven modelling	37
2017	Naghizadeh, M. Manteghi, M. Ranga, M. & Naghizadeh, R.	Managing integration in complex product systems: The experience of the IR-150 aircraft design program	Key capabilities in Complex Products and Systems	10
2017	Kaslow, D. Ayres, B. Cahill, P. T. Hart, L. & Yntema, R.	A Model-Based Systems Engineering (MBSE) approach for defining the behaviors of CubeSats	Case studies on Model-based systems engineering	3
2017	Fusaro, R. Ferretto D. & Viola, N.	MBSE approach to support and formalize mission alternatives generation and selection processes for hypersonic and suborbital transportation systems	Case studies on Model-based systems engineering	3
2017	Heber, D. & Groll M.	Towards a digital twin: How the blockchain can foster E/E-traceability in consideration of model-based systems engineering	Model-based systems engineering and data	8
2017	Schluse, M. Atorf, L. & Rossmann, J.	Experimentable digital twins for model-based systems engineering and simulation-based development	Model-based systems engineering and data	15
2017	Bach, J. Langner, J. Otten, S. Sax, E. & Holzapfel, M.	Test scenario selection for system-level verification and validation of geolocation-dependent automotive control systems	Data driven modelling	7
2017	Bach, J. Langner, J. Otten, S. Holzapfel, M. & Sax, E.	Data-driven development, a complementing approach for automotive systems engineering	Data driven modelling	3
2018	Zeigler, B. P. Mittal, S. & Traore, M. K.	MBSE with/out Simulation: State of the Art and Way Forward	Emerging cultural and technical challenges in Model-based systems engineering	11
2018	Lindblad, L. Witzmann, M. & Bussche, S. V.	DATA-DRIVEN SYSTEMS ENGINEERING: TURNING MBSE INTO INDUSTRIAL REALITY	Model-based systems engineering and data	2
2018	Di Maio M. Kapos, G.-D. Klusmann, N. Atorf, L. Dahmen, U. Schluse, M. & Rossmann, J.	Closed-loop systems engineering (close): integrating experimentable digital twins with the model-driven engineering process	Model-based systems engineering and data	1
2018	Liu, Z. Jin, C. Jin, W. Lee, J. Zhang, Z. Peng, C. & Xu, G.	Industrial AI Enabled Prognostics for High-speed Railway Systems	Data driven monitoring and fault diagnostics	3
2018	Liu, Y. & Goebel, K.	Information fusion for national airspace system prognostics: A NASA ULI project	Data driven control and optimization	3

2018	Norman, R. Bolin, J. Powell, E. T. Amin, S. & Nacker, J.	Using Big Data Analytics to Create a Predictive Model for Joint Strike Fighter	Data driven control and optimization	0
2018	Hybertson, D. Hailegiorghis, M. Griesi, K. Soeder, B. & Rouse, W.	Evidence-based systems engineering	Data driven modelling	1
2019	Roehrich, J. K. Davies, A. Frederiksen, L. & Sergeeeva, N.	Management innovation in complex products and systems: The case of integrated project teams	Organization structure in Complex Products and Systems	3
2019	Lehtinen, J. Aaltonen, K. & Rajala, R.	Stakeholder management in complex product systems: Practices and rationales for engagement and disengagement	Key capabilities in Complex Products and Systems	6
2019	Li, Q. Wei, H. Yu, C. & Wang, S.	Model and Data Driven Complex Product Development: from V, Double Vs to Triple Vs	Model-based systems engineering and data	0
2019	Madni, A. M. Madni, C. C. & Lucero, S. D.	Leveraging digital twin technology in model-based systems engineering	Model-based systems engineering and data	51
2019	Elattar, H. M. Elminir, H. K. & Riad, A. M.	Conception and implementation of a data-driven prognostics algorithm for safety–critical systems	Data driven monitoring and fault diagnostics	0
2019	Ren, Y. Wang, N. Jiang, J. Zhu, J. Song, G. & Chen, X.	The Application of Downhole Vibration Factor in Drilling Tool Reliability Big Data Analytics - A Review	Data driven control and optimization	3