

Doctoral theses at NTNU, 2022:47

Nanda Anugrah Zikrullah

Contributions to the safety of novel subsea technologies – Methods and approaches to support the safety demonstration process

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Engineering
Department of Mechanical and Industrial
Engineering



Norwegian University of
Science and Technology

Nanda Anugrah Zikrullah

Contributions to the safety of novel subsea technologies – Methods and approaches to support the safety demonstration process

Thesis for the Degree of Philosophiae Doctor

Trondheim, February 2022

Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU

Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Engineering

Department of Mechanical and Industrial Engineering

© Nanda Anugrah Zikrullah

ISBN 978-82-326-5459-8 (printed ver.)

ISBN 978-82-326-6561-7 (electronic ver.)

ISSN 1503-8181 (printed ver.)

ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2022:47

Printed by NTNU Grafisk senter

RAMS

Reliability, Availability,
Maintainability, and Safety

Contributions to the safety of novel subsea technologies – Methods and approaches to support the safety demonstration process

Nanda Anugrah Zikrullah

October 2021

PhD THESIS

Department of Mechanical and Industrial Engineering

Faculty of Engineering

Norwegian University of Science and Technology

Supervisor 1: Professor Mary Ann Lundteigen

Supervisor 2: Associate Professor Hyungju Kim

Supervisor 3: Functional Safety Researcher Meine J.P. van der Meulen

Preface

Personally, the decision to do PhD was pivotal in my life. The past three years of my research journey impacted my personal and professional life. I have gained numerous experiences full of ups and downs. There were moments where they were so unbearable that I really wanted to quit and looked for 'easier' alternatives. However, I felt lucky that I managed to struggle until the end, obviously with the supports from everyone around me. In the end, I can say that "I have finished my PhD". Indeed, I do not want to repeat it. Nevertheless, the experience is irreplaceable. If I get a chance to talk with my past self, I will encourage myself to take this once in a lifetime chance.

Formally, this thesis results from a PhD project at the Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU). It is submitted to NTNU for partial fulfilment of the Doctor of Philosophy (PhD) degree requirements. The work has been fully funded by the Safety 4.0 project partners and the research council of Norway. The work was carried out from August 2018 to September 2021.

This PhD is supported through close collaborations with my main supervisor, Professor Mary Ann Lundteigen, at the Department of Engineering Cybernetics (NTNU), previously employed at the same department as me before switching during my second year of PhD. Her contributions are reflected in several articles I produced during my PhD. In addition, my co-supervisor, Associate Professor HyungJu Kim from the Department of Maritime Operations, University of South-Eastern Norway (USN), has contributed with his expertise in safety assessment and maritime engineering. Also, my second co-supervisor, Meine J.P. van der Meulen, from the Group Technology and Research, Det Norske Veritas (DNV), provides excellent industrial insights and networking ability to support my research.

I am glad that my PhD has granted me an opportunity for making contributions to a field in which I take great interest, namely functional safety. I learnt this topic from a practical perspective during my period of employment in Indonesia. When coming here to Norway to pursue my Master degree, I thought I would only gain more knowledge, obtain my degree, and return to Indonesia to continue working. Instead, I am glad that I have contributed to developing new knowledge, ranging from concepts to methods, which I hope will lead to safer design and operations of such systems. I also hope that this experience could assist me in future professional employment.

Trondheim, 2021-09-27

Nanda Anugrah Zikrullah

Acknowledgment

Supports from others contributed greatly to the completion of my PhD. While it is not possible to mention everyone, I would, to my best ability, acknowledge those that have significant impacts on the process.

First, I would like to thank my main supervisor, Professor Mary Ann Lundteigen, for her support, guidance, and encouragement. She is conscientious and always encouraged me to be critical, precise and clear to improve the quality of my work. I must admit that it was always challenging to fulfil her expectation. Nevertheless, it felt rewarding when getting acknowledged. Looking back from now, she was the best supervisor that I could have dreamt of. If I ever get to supervise a junior in any occupation, she would be my role model. I hope that I can continue to have new and exciting collaborations with her in the future.

I would also like to thank my co-supervisor, Associate Professor HyungJu Kim, for sharing his expertise and becoming an inspiration for improving my way of doing research. Unfortunately, our close collaborations may have been short since he needs to relocate from Trondheim to Borre for a new opportunity at USN. Nevertheless, the guidance I received from you since my Master period has shaped my PhD journey. I felt lucky to have you as one of my co-supervisors.

My industry contact is the second co-supervisor, Functional Safety Researcher Meine J.P. van der Meulen from DNV. I am happy to have intelligent discussions and close collaborations, practically on all my results. While he was expected to provide industrial inputs, his current focus in research provides new insights and perspectives to the problems. It was fun working with him, and I must also acknowledge his help in my future professional life. His networks and advice allow me to land my next job as a functional safety consultant for DNV. I hope that we can still maintain contacts and collaborations, especially since we are a colleague.

On two articles, I have collaborated with researchers from Equinor, Adjunct Professor Gunleiv Skofteland, which is also employed at NTNU. He has shared his extensive and detailed knowledge of subsea processing systems. I am thankful for his time and efforts in supporting my works.

I want to thank all the project partners from Safety 4.0 and SUBPRO. My PhD project would not have even started without their continuous supports and interest in the topic. Also, the funding provided by Safety 4.0 partners and the research council of Norway has allowed this project to flow. I would also like to mention the funding support from SUBPRO for ensuring the work-life balance during my PhD. Even though I am not officially under SUBPRO, they allow me to join all the formal and informal activities, leading to new networks. This could not happen without the support of Professor Sigurd Skogestad and the administration team, Gro Mogseth, Esma Benzaim, and Pål J. Aune.

Within the Safety 4.0 projects, I am grateful for my DNV colleague's support, especially my project manager, Tore Myhrvold. He is punctual and systematic, and his assistance ensures a smooth process of the project. I would also like to mention my other colleagues I

have discussions with, Odd Ivar Haugen, Frank Børre Pedersen, Andreas Hafver, Christine L. Berner, Ketil F. Hansen, Siegfried Eisinger, and Kenneth Kvinnesland. From partners of Safety 4.0, I should also mention my gratitude for Christoffer Lassen from Equinor, Øyvind Rokne from TechnipFMC, Cato Bratt and Geir Lund from ABB, Carsten Mahler from OneSubsea, Ellen M.S. Lycke and Jone Sigmunstad from Aker Solutions, Rory Mackenzie from Total, and Eirik Duesten from Petroleum Safety Authority. Their differing backgrounds expertise led me to achieve greater heights in the quality of the results.

I would also like to thank the administrative staff at the Department of Mechanical and Industrial Engineering for smoothing the work process. In addition, I want to acknowledge the discussions' opportunities with Professor Jørn Vatn, Professor Antoine Rauzy, Associate Professor Nicola Paltrinieri, Associate Professor Yiliu Liu, and Associate Professor Bjørn Axel Gran. Their knowledge helped me to shape the interesting research topics to focus on during my PhD.

At NTNU, I would like to extend my humblest gratitude to several colleagues that share the same PhD journey with me that is Aibo Zhang, Abu M.D.A. Islam, Bahareh Tajiani, Behnaz H. Davatgar, Federico Ustolin, Ewa M. Laskowska, Himanshu Srivastav, I G.A.G. Angga, Jan Sramota, Jon Martin Fordal, Juntao Zhang, Lin Xie, Liu Yang, Michael Pacevicius, Muhammad M. Sabara, Renny Arismendi, Shenae Lee, Tae Hwan Lee, Tianqi Sun, Tom Ivar Pedersen, Xinge Qi, Xingheng Liu, and Yun Zhang. All the moments we share inside and outside the office meant a lot for me. For those that are still ongoing with their PhD, I know that you will follow shortly. While for those that have finished, I hope that I can follow your steps on giving further contributions in the fields.

A special thanks to my Indonesian friends in Trondheim, Aditya Wihen, D'Aqnan M.M. Pandi, Fadhil Berilyian, Mikael Y. Estuariwinarno, Harbi Qadri, Muhammad G. Alfarizi, Ressi B. Muhammad, and Robin A. Surya, for the friendship and happy moments together. I hope all of you will also be successful in the future.

Finally, I would like to thanks my parents, Dwi D. Hartadjaja and Ninik Joeniwati, for their patience and continuous supports and trust in me. I would also like to thank my wife, Salma Alkindira, for accompanying me during my PhD journey and for lifting me during the toughest period.

thank you

Abstract

This PhD thesis explains the contribution made to the safety of novel subsea technologies. It is supported by this thesis objective, which is to develop and demonstrate the application of new safety assessment methods within the scope of functional safety, which can capture and manage the complex operational behaviour of novel software-intensive systems. The objective is supported by several study cases, focusing primarily on novel subsea systems for the oil and gas industry. The novel and complex characteristics of the systems are represented in the concept of integrating the control and safety elements. The following five research questions' topics have been addressed explicitly:

- Topic I – Safe design principles. This study clarifies several safe design principles that are derived from the design approaches in several industries. It is found that the governing functional safety standard, IEC 61508, is aligned with the safe design principles. These principles have been applied to the study case. The implication is that some processes need to be adapted for novel technology involving software-intensive systems with complex operational behaviours.
- Topic II – Solution-specific safety requirements. Two hazard analysis approaches that are often considered well suited for hazard analysis of novel technology, i.e., functional hazard analysis (FHA) and systems-theoretic process analysis (STPA) has been compared. The authors investigate the characteristics of both methods in more detail by performing study cases on an equipment protection system in subsea processing applications. It is concluded that STPA is more suitable based on various factors, including the method's approach, modelling coverage, and analysis capability. The study also provides recommendations for the improvement of both methods.
- Topic III – Alternative concepts. The study proposes a new classification method to distinguish different integration types, from complete independence to complete integration. STPA is then performed several times on systems with different levels of integration at the logic solver. The study also proposes a modelling technique in STPA to capture the different integration types. The result found that integration does not necessarily change the system's functionality, but it may introduce new interactions leading to hazards. Nevertheless, the magnitude of risk for the hazard is unknown.
- Topic IV – Effect on risk. The study proposes a modelling pattern to quantify the hazardous scenarios' frequency. A text-based finite-state automata modelling pattern implemented in Altarica 3.0 has been developed. The authors demonstrate the approach capability by performing a study case on the STPA results from the topic III study. It is found that the method is capable of capturing dependencies while also highlighting the inefficiency of STPA caused by unnecessary requirement productions. The study also discusses the method's limitation if compared with other quantification processes recommended in the standards.

- Topic V – Safety argumentation. This topic summarises all the preceding results to clarify the link between the developed methods and approaches with the safety argumentation concept for novel technology. The concept is based on an argument-induced evidence model. While the PhD works do not cover all aspects of the safety argumentation concept, this PhD highlights the current state of the research and the required further works to build a complete safety demonstration framework for novel technology.

The overall implications of the framework and methods developed in this PhD thesis are that the engineers or analysts now have more assurance during the safety demonstrations process of novel technology involving software-intensive systems. The overall development process for the framework has been explained in this thesis and scrutinized through a systematic peer review process. This thesis also serves as an input for the ongoing joint industry research project Safety 4.0, which aims to enable and accelerate the uptake of novel subsea solutions by developing a standardized safety demonstration framework.

Structure of the report

This doctoral thesis is written in a collection of articles format, commonly known as thesis compilation. The thesis is split into two parts:

- Part I: Main Report summarizes and links the articles and research contributions within a similar context. Part I describes how most of the defined research challenges and objectives have been answered in the main results. The remaining challenges are summarized as ideas for further research works.
- Part II: Articles, which consist of standalone articles that have been published or submitted in international conferences and journals.

Readers who are interested in the overall PhD research topic may focus on part I. Readers interested in solving particular challenges within a similar research area may focus on part II.

List of publications

This thesis includes four publications that have been submitted or published in international journals and conferences. The publications are listed in Table 1 and the full texts are presented in Part II: Articles.

Table 1: Overview of articles included in this PhD thesis.

Article ID	Page	Title	Status
Article I	71	Clarifying implementation of safe design principles in IEC 61508: challenges of novel subsea technology development	Published
Article II	81	A comparison of hazard analysis methods capability for safety requirements generation	Published
Article III	105	A comparison of hazardous scenarios in architectures with different integration types	Published
Article IV	115	Finite-state automata modeling pattern of systems-theoretic process analysis results	Submitted / under review

The following details of the articles included in the thesis are presented with the Authors' contributions.

Article I

N. A. Zikrullah, H. Kim, M. J. P. van der Meulen, M. A. Lundteigen, Clarifying implementation of safe design principles in IEC61508: Challenges of novel subsea technology development, in: Proceedings of the 29th European Safety and Reliability Conference (ESREL), Research Publishing, 2019, pp. 2928–2936.

Contributions from the authors

First (I), second and fourth authors conceptualized the research idea as a starting point for the PhD. Then, I identified the research gaps and state of the art on the topic. Afterwards, I proposed a methodology for systematic assessment of the standard that the co-authors vetted. Next, I, third, and fourth authors formulated the study case, and the co-authors validated the results. Finally, I wrote the manuscript, and the co-authors reviewed the work.

Article II

N. A. Zikrullah, H. Kim, M. J. P. van der Meulen, G. Skofteland, M. A. Lundteigen, A comparison of hazard analysis methods capability for safety requirements generation, Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability. 2021, 1748006X211003463.

Contributions from the authors

Research problems were an ongoing work at DNV, and based on the findings from article I, the first (I) and third authors conceptualized as research ideas. First, I identified state of the art, which filters the appropriate methods for the topic. Then, I proposed the research methodology vetted by the second, third, and fifth authors. A tool from database software has been developed by me to record the analysis results. The fourth author and I formulated the study case based on the pump part of a subsea gas compression system. The co-authors verified the results that I produced. Finally, I wrote the manuscript, and the co-authors reviewed the work.

Article III

N. A. Zikrullah, M. J. P. van der Meulen, G. Skofteland, M. A. Lundteigen, A comparison of hazardous scenarios in architectures with different integration types, in: Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference (ESREL 2020 PSAM15), Research Publishing Services, 2020, pp. 4001–4008.

Contributions from the authors

Based on the results of article II, the first (I) and second authors challenged the application of the method for the novel use cases for integration of process control and safety system. I identified state of the art for use cases, and together with the third authors, formulated the study case based on a compressor part of a subsea gas compression system. Then, I proposed a modelling approach for different types of architectures from the study case, utilized the tool developed during the work of article II for the analysis, and compared the statistics of the results. The second and fourth authors validated the results. Finally, I wrote the manuscript, and the co-authors reviewed the work.

Article IV

N. A. Zikrullah, M. J. P. van der Meulen, M. A. Lundteigen, Finite-state automata modeling pattern of systems-theoretic process analysis results, Reliability Engineering & System Safety. (n.d.), under review.

Contributions from the authors

First (I) and third authors conceptualized the research idea. Then, I identified the state of the art and research gaps. Afterwards, a methodology is proposed, and I frame the modelling assumptions that the second and third authors vetted. Next, the study case is developed by me from article III and obtained research data from SINTEF and the literature. Then, I produced simulation results and developed a simplified model to validate the results based on an alternative method. The results are verified and validated by the second and third authors. Finally, I wrote the manuscript, and the co-authors reviewed the work.

I found it challenging to reproduce early work from the literature during the research work as not all relevant data were available. These challenges hindered my progress for several months. Besides, I also realized that not all of my results could be published in written form. Therefore, my supervisors and I agreed not to repeat the same problems by publishing most of the valuable research data for others in the future. My published research data is listed as follow:

Research Dataset I

N. A. Zikrullah, M. J. P. van der Meulen, G. Skofteland. M. A. Lundteigen, “Systems-Theoretic Process Analysis results for system with different integration types”, Mendeley Data, V1, (2021), DOI: 10.17632/prwtzmt3kg.1

Contributions from the authors

Research dataset I result from the analysis of a compressor part of a subsea gas compression system from article III. While article III focuses only on the findings, the current dataset provides all the raw results. In addition, the datasets were recorded in a tool from database software that I developed from article II that other users can also use for similar analysis.

Research Dataset II

N. A. Zikrullah, M. J. P. van der Meulen, M. A. Lundteigen, Systems-theoretic process analysis - finite state automata (STPA-FSA) modeling approach source code, Mendeley Data V1, (2021), DOI: 10.17632/39g8ywj7j.1.

Contributions from the authors

Research dataset II is the source code of the modelling pattern in the method proposed in Article IV. First (I) author developed the code in Altaria 3.0, and complete validation results of the methods by comparison with a simplified model based on an alternative method are attached.

There were several other works done during this PhD from direct collaboration with the industry partners. These publications were the products of the articles listed above. The contributions were significant but are not included in this thesis. The reasons for the exclusion are explained in the following.

Technical report I

N. A. Zikrullah, M. J. P. van der Meulen, Use case 2 report, Safety 4.0 Internal Technical Report, (2020).

Contributions from the authors

The technical report is developed based on a PhD project plan, article I, and other works to clarify the integration concept. The first (I) and second authors conceptualized the content of the report. The authors performed discussions, workshops, and literature reviews to build this report. Both authors split the responsibility to write the manuscript. A cross-review process was performed by the authors and vetted by the industry partners of Safety 4.0.

Reasons for exclusion

This technical report aims to communicate the progress of the use case (UC) 2’s works as part of the Safety 4.0 project of which this PhD is a part. The report focuses on high-level clarification of the UC 2 concept against the regulations and standards, proposes alternative architectures for integration used for works in the PhD, and describes a systematic plan to solve the UC 2 challenges. I decided to exclude this technical report because the contents are insufficient as a separate scientific publication, and it is only an internal report that has not been peer-reviewed by the scientific community. However, relevant information from

the report is included as part of the thesis.

Technical report II

N. A. Zikrullah, M. J. P. van der Meulen, Use case 2 – Identifying constraints and comparing design alternatives using STPA, Safety 4.0 Internal Technical Report, (2021).

Contributions from the authors

This technical report was developed based on article II, article III, and research dataset I. The first (I) and second authors conceptualized the content of the report. The second author was responsible for writing the manuscript, and I was responsible for reviewing the content. The industry partners of Safety 4.0 vetted the report.

Reasons for exclusion

This technical report aims to communicate the results of work done for UC 2 as part of the Safety 4.0 project of which this PhD is a part. The report focuses on demonstrating the usefulness of STPA and the challenge of having integration for the system's hazards. The target audience is the industry participants. Thus, the report has more extensive content than the articles to include more examples from the dataset. However, the in-depth discussions still referred to the published articles. I decided to exclude this technical report because the contents are too similar to articles II and III. However, relevant information from the report is included as part of the thesis.

Book chapter I

N. A. Zikrullah, M. J. P. van der Meulen, M. A. Lundteigen, Chapter 14: Towards safe integration, in: Demonstrating Safety of Novel Solutions – With examples from subsea electric technology (working title), (n.d.), to be submitted.

Contributions from the authors

This book chapter covered use case 2 results in Safety 4.0 and was developed based on all the articles, technical reports, and this PhD thesis. The first (I) author conceptualized the chapter's content that the second and third authors vetted. The first author performed a literature review to supplement the content. I wrote the manuscript, and the co-authors reviewed the work. Finally, the industry partners of Safety 4.0 vetted the content.

Reasons for exclusion

This book chapter aimed to communicate the scientific knowledge obtained from UC 2. The report focuses on the integration concept and the practical challenges of having integration at a higher level. The target audiences are people with sufficient expertise, i.e., industry experts or researcher. Although the book chapter has a different focus than this PhD thesis, I decided to exclude this technical report due to the considerable similarity of content.

Book chapter II

A. Hafver, D. Kostopoulos, N.A. Zikrullah, Chapter 9: Assessing safety from a systemic and life-cycle perspective, in: Demonstrating Safety of Novel Solutions – With examples from subsea electric technology (working title), (n.d.), to be submitted.

Contributions from the authors

This book chapter reused article II methods and findings as one subsection of the con-

tent. The first author conceptualized the chapter's content that the second author vetted. The first author performed a literature review to supplement the content and wrote the manuscript, while the co-authors (including me) reviewed the work. Finally, the industry partners of Safety 4.0 vetted the content.

Reasons for exclusion

This book chapter aimed to communicate the scientific knowledge obtained from the work of safety 4.0. The report focuses on the systemic and life-cycle perspectives when assessing safety. The target audiences are people with sufficient expertise, i.e., industry experts or researcher. I decided to exclude this technical report since it only reused parts of the article that I published.

Contents

Preface	iii
Acknowledgment	v
Abstract	vii
Structure of the report	ix
List of publications	xi
Part I: Main Report	1
1 PhD Project Background	3
1.1 Safety 4.0 background and structure	3
1.2 PhD project focus	5
1.3 Terms and definitions	6
1.4 Main report structure	7
2 Industrial Status and Challenges	9
2.1 Subsea systems	9
2.1.1 Subsea production systems	10
2.1.2 Subsea processing systems	11
2.1.3 Subsea control systems	11
2.2 Subsea risk picture	13
2.3 Selected use case	14
2.4 Regulations, guidelines, and standards	15
2.5 Technology qualification	17
2.6 Functional Safety	17
2.7 Challenges of introducing novel subsea technologies	19
2.7.1 Challenge I – Increase in complexity due to integration of software-inten- sive systems	20
2.7.2 Challenge II – The non-compliance of technology against the local reg- ulations and standards	20
2.7.3 Challenge III – Lack of safety demonstration process framework for the complex software-intensive systems	21
3 Academic status and gaps	23
3.1 Complexity?	23
3.2 Hazard and risk	25
3.2.1 Risk identification	26

3.2.2	Risk analysis and evaluation	26
3.2.3	Risk treatment	26
3.3	Safety assessment and demonstration	27
3.3.1	Requirement generations	28
3.3.2	Evidence and assumptions	29
3.3.3	Argument and justification	30
3.4	Gaps in academia	30
3.4.1	Gap I – Unavailability of the safe design principles	30
3.4.2	Gap II – Ambiguous safety requirement generations methods for complex system	31
3.4.3	Gap III – Unknown classification methods for the integration concept	31
3.4.4	Gap IV – Need of a reliability performance’s modelling approach to generate evidence for the complex system’s safety	32
3.4.5	Gap V – Need of clarification of framework for safety demonstration	32
4	Research Questions, Objectives, and Delimitation	33
4.1	Research questions	33
4.1.1	Topic I – Safe design principles	33
4.1.2	Topic II – Solution-specific safety requirements	35
4.1.3	Topic III – Alternative concepts	35
4.1.4	Topic IV – Effect on risk	36
4.1.5	Topic V – Safety argumentation	36
4.2	Research objectives	36
4.3	Delimitation	37
5	Research Methodology	39
5.1	Research motivation	39
5.2	Classification of research	40
5.3	Research approaches	40
5.4	Challenges and lessons learned	42
6	Key Results and Contributions	45
6.1	Contribution I – Safe design principles	45
6.1.1	Identification of the safe design principles	46
6.1.2	Alignment of the requirements in IEC 61508 part II with the safe design principles	46
6.1.3	Challenge of the safety demonstration process of novel technology	47
6.2	Contribution II – Solution-specific safety requirements	47
6.2.1	Identification of the most suitable methods for hazard analysis of novel and complex software-intensive solutions	47
6.2.2	Findings on the hazard analysis methods’ capability for identifying functional hazards	48
6.2.3	Findings on the hazard analysis methods’ capability to provide systemic perspective for the analysis	48
6.2.4	Identification of the produced solution-specific safety requirements’ characteristics	48
6.2.5	Recommendation on the improvement of the hazard analysis methods from the lessons learned	49
6.3	Contribution III – Alternative concepts	49

6.3.1	Proposal of the integration concept classification	49
6.3.2	Proposal of hierarchical control structure modelling approach considering the integration	50
6.3.3	Challenges for implementing different types of integration concepts	50
6.4	Contribution IV – Effect on risk	52
6.4.1	Improvement of a modelling approach for STPA's loss scenarios based on finite-state automata modelling type	52
6.4.2	Discussion on the model capability to address dependency	52
6.4.3	Model's capability to identify unnecessary requirements	53
6.4.4	Identification of the proposed model limitations when compared to the available modelling approaches	53
6.5	Contribution V – Safety argumentation	54
6.5.1	Clarification for the safety argumentation concept	54
6.5.2	Relevance assessment for the available methods in the framework	55
6.5.3	Remaining aspects of safety argumentation concept	56
7	Summary and Recommendations for Further Work	57
7.1	Summary and Conclusion	57
7.2	Recommendation for Further Works	58
7.2.1	Generic application of the framework	58
7.2.2	Uncertainty management of the results	58
7.2.3	Management of software and systematic safety integrity	58
7.2.4	Testing of the safety argumentation concept	59
	Bibliography	61
	Part II: Articles	69
	Article I	71
	Article II	81
	Article III	105
	Article IV	115

List of Figures

1.1	PhD project interaction with other work packages and partners of Safety 4.0 project.	5
1.2	The overall structure of this PhD thesis.	8
2.1	Simplified illustration of subsea systems for an oil and gas field. (Note, *: Subsea production systems; **: Subsea processing systems).	10
2.2	A typical process for gas lift (adapted from [17]). Grey box represents example elements that can be used for subsea processing systems.	11
2.3	Element of subsea control systems utilizing electrohydraulic technology.	12
2.4	A control loop.	13
2.5	Subsea risk picture (adapted from [18]).	14
2.6	Example of process control system and safety instrumented system on a subsea compression system having integration on the logic solvers.	15
2.7	Hierarchy of acts, regulations, standards, and operator specification in Norway.	16
2.8	Functionals safety lifecycle process based on IEC 61508.	18
2.9	Treatment of loss scenarios derived from IEC 61508 process.	19
3.1	The metamodel of complexity as combination of a variety of models (adapted from [52]).	25
3.2	Illustration of hazard, hazardous event, and barrier.	27
3.3	Safety argumentation concept.	28
3.4	Effect of different requirement types on design process (adapted from [64]).	29
3.5	Relations between the design principles, technology, and formalized requirements.	31
4.1	PhD scope based on simplified safety lifecycle of IEC 61508 [7].	34
5.1	The research process, activity, and results.	41
6.1	Link between academic gaps, research questions, contributions, and PhD objectives.	45
6.2	Overview of PhD contributions.	46
6.3	Example of solution-specific safety requirement (or controller constraint) produce by STPA.	49
6.4	Generic architecture for control and safety logic solvers in for horizontal separation.	50
6.5	Proposed hierarchical control structure model considering different integration concept for STPA.	51
6.6	Comparisons of loss scenarios for system with different integration level.	52
6.7	Example of generic model in STPA-FSA.	53

6.8 Example of sensitivity analysis results. 53

6.9 Example of safety argument. 54

6.10 Linking the proposed methods with the safety lifecycle. 55

A.1 Revision of 'Fig. 7. Number of loss scenarios for system with different integra-
tion types on Article III [92]. 105

List of Tables

1	Overview of articles included in this PhD thesis.	xi
1.1	Glossary of key terms.	6
2.1	Technology novelty categorization (adapted from [35]).	17
5.1	Criteria of research and how it is achieved.	40
5.2	RQ-specific research approaches.	42

List of Acronyms

API American Petroleum Institute

ARP Aerospace Recommended Practice

BBN Bayesian Belief Network

CCPS Center for Chemical Process Safety

CENELEC European Committee for Electrotechnical Standardization

CESM Composition, Environment, Structure, Mechanism

DDR&E Director, Defense, Research, and Engineering

DNV Det Norske Veritas

DOI Digital Object Identifier

DRD Director, Research Directorate

E/E/PE Electrical/ Electronic/ Programmable electronic

EN European Standards

ESREL European Safety and Reliability Conference

EUC Equipment Under Control

FHA Functional Hazard Analysis

FMEA Failure Mode and Effect Analysis

FMECA Failure, Mode, Effect, and Criticality Analysis

FPSO Floating Production, Storage, and Offloading

FRAM Functional Resonance Analysis Method

FSA Finite State Automata

FTA Fault Tree Analysis

GSN Goal Structuring Notation

GSPN Generalized Stochastic Petri Net

HAZOP Hazard and Operability Study

- HSE** Health, Safety, and Environment
- IEC** International Electrotechnical Commission
- IEEE** Institute of Electrical and Electronics Engineers
- ISBN** International Standard Book Number
- ISO** International Organization for Standardization
- ISSN** International Standard Serial Number
- LoA** Level of Abstraction
- mA** mili Ampere
- MIT** Massachusetts Institute of Technology
- N.D.** No date
- NCS** Norwegian Continental Shelves
- NOG** Norwegian Oil and Gas
- NORSOK** The Norwegian Shelf's Competitive Position
- NTNU** Norwegian University of Science and Technology
- NUREG** US Nuclear Regulatory Commission Regulation
- OECD** The Organisation for Economic Co-operation and Development
- OG21** Oil and gas strategy for the 21st century
- OTC** Offshore Technology Conference
- PAS** Publicly Available Specification
- PCS** Process Control System
- PhD** Doctor of Philosophy
- PRA** Probabilistic Risk Assessment
- PSA** Petroleum Safety Authority
- PSAM** Probabilistic Safety Assessment and Management
- R&D** Research and Development
- RAM** Reliability, Availability, and Maintainability
- RAMS** Reliability, Availability, Maintainability, and Safety
- RBD** Reliability Block Diagram
- RIDM** Risk-Informed Decision Making

RP Recommended Practice

RQ Research Question

SAE The Society of Automotive Engineers

SAS Safety and Automation System

SFI Senter for Forskningdrevet

SIL Safety Integrity Level

SIS Safety Instrumented System

SPN Stochastic Petri Net

SSFMEA Software System Failure Mode and Effect Analysis

STPA Systems-Theoretic Process Analysis

SUBPRO Subsea Production and Processing

TQP Technology Qualification Program

TR Technical Report

TRA Technology Readiness Assessment

TRL Technology Readiness Level

UC Use Case

UIS University of Stavanger

URL Uniform Resource Locator

USN University of South-Eastern Norway

V Volt

WP Work Package

Part I

Main Report

Chapter 1

PhD Project Background

This chapter provides the background for this PhD project. The PhD project has been part of a Safety 4.0 project, a joint industry project headed by DNV focusing on demonstrating the safety of novel subsea technologies. First, this chapter starts by providing a brief introduction to Safety 4.0, including the background and structure in section 1.1. Then, section 1.2 describes the focus of the PhD project and its interaction with Safety 4.0. It is followed by a summary of frequently used terms and definitions in this thesis in section 1.3. Finally, section 1.4 outlines the structure of the main report.

1.1 Safety 4.0 background and structure

Safety 4.0 is a project lead by DNV, and it provides the primary resources for the research activities. From academia, the Norwegian University of Science and Technology (NTNU) and the University of Stavanger (UiS) provide resources focusing on the PhD project and the Postdoc project, respectively. This project is also supported by industry partners such as Equinor Energy, Neptune Energy Norge AS, Lundin Norway AS, and TotalEnergies that provide expertise from the operator perspective, and ABB, TechnipFMC, Aker Solutions, and OneSubsea that provide expertise from the manufacturer's perspective. Finally, Petroleum Safety Authority Norway (PSA) is a regulatory body that serves as an observer for the project. The collaborations from the major oil and gas industry players have initiated Safety 4.0 to tackle recent challenges from the industries related to novel technologies.

Oil and gas represent a significant export industry for Norway. Subsea technology is developed to acquire more hydrocarbons in places that were unreachable before. 75% of the recent discoveries on the Norwegian continental shelf are shared between the wellhead or tie-back subsea developments [1]. Cost savings, increased production efficiency, environmental challenges (water depth pressure and temperature), and accessibility are some of the recurring subsea challenges when developing new or existing fields. Hence, they drive the emergence of novel subsea technologies that can tackle the challenge above more efficiently.

One example of novel technology is the all-electric subsea production system. This technology allows continuous monitoring of the well's condition that cannot be obtained before. Another example is the integration between control and safety systems for the subsea processing system. This concept may reduce physical complexity for subsea applications, reducing the cost of installation.

The novel concepts are, however, not free from risk. The risk may affect either the financial, environment or safety, as there is no or limited experience about the performance. For example, an all-electric subsea production system requires changes in the operation and

maintenance philosophy [2]. These changes may introduce new or different errors during operations. On the integration example, improper implementations may compromise the performance of the safety system [3]. Hence, making the system more vulnerable. This would be problematic for the industry, as several accidents have occurred in the past years, e.g., Philadelphia refinery explosion (2019), Deepwater horizon (2010), and Texas city refinery explosion (2005). Due to this, safety represents one of the critical properties of many novel systems. Therefore, it is vital to demonstrate that the novel technology would behave as intended.

Unfortunately, the current safety demonstration process is inefficient due to the lack of support from relevant standards and regulations [4]. Therefore, further research and development (R&D) are necessary to tackle this issue. The focus on R&D by Norway's oil and gas industry participants has led to several research activities and project collaborations, both by industry and academia. One example of a research centre in academia is SFI SUBPRO (Subsea Production and Processing) under NTNU. SFI refers to Senter for Forskningsdrevet Innovasjon (Norwegian) or Centre for Research-based Innovation. The research council of Norway supports SFI as a long-term initiative to build up research groups in important areas. The conducted research requires commitment since results and findings can lead to subsequent research projects.

Safety 4.0 project utilized one of the SFI SUBPRO project's results named 'new safety and control philosophy for subsea' to pinpoint the gap in the safety demonstration process. Officially, the project title is 'Safety 4.0 – Demonstrating safety of novel subsea technologies' [5]. The safety 4.0 project started in mid-2018 for three years and is funded by the Petromaks 2 program [grant number 281877/E30] and the project partners.

The Safety 4.0 project aims to enable and accelerate the uptake of novel subsea solutions by developing a standardized safety demonstration framework. The framework is based on governing framework principles: adaptive, argument-based, modular, uncertainty-based risk perspective, systems perspective, and life-cycle perspective. The objective is further divided into seven work packages (WPs):

- WP1 Project framing and mapping of needs. The WP1 objective is to frame the project by developing a detailed plan based on mapped gaps, challenges, opportunities (improvement potentials), and needs (wanted outcome of the Safety 4.0 project).
- WP2 Framework development. The WP2 objective is to develop a framework (work processes, methods, and tools) for standardized demonstration of safety for novel subsea technologies
- WP3 Tests & demonstrations. The WP3 objective is to utilize three relevant use cases (UC) to exemplify and address the research challenges in the Safety 4.0 project.
- WP4 Ensuring functional safety of novel technologies (NTNU PhD project). The WP4 objective is to develop and demonstrate the application of new safety assessment methods that can identify requirements and capture the safety behaviour of novel and complex subsea systems. Additional details are briefly mentioned in the next section and further discussed in Chapter 4.
- WP5 Knowledge-based approaches and methods for risk-informed safety demonstration (UiS Postdoc project). The WP5 objective is to: i.) develop scientifically well-founded methods and practical guidance on using better and reflect the knowledge dimension when assessing risk related to novel solutions, and ii.) develop principles

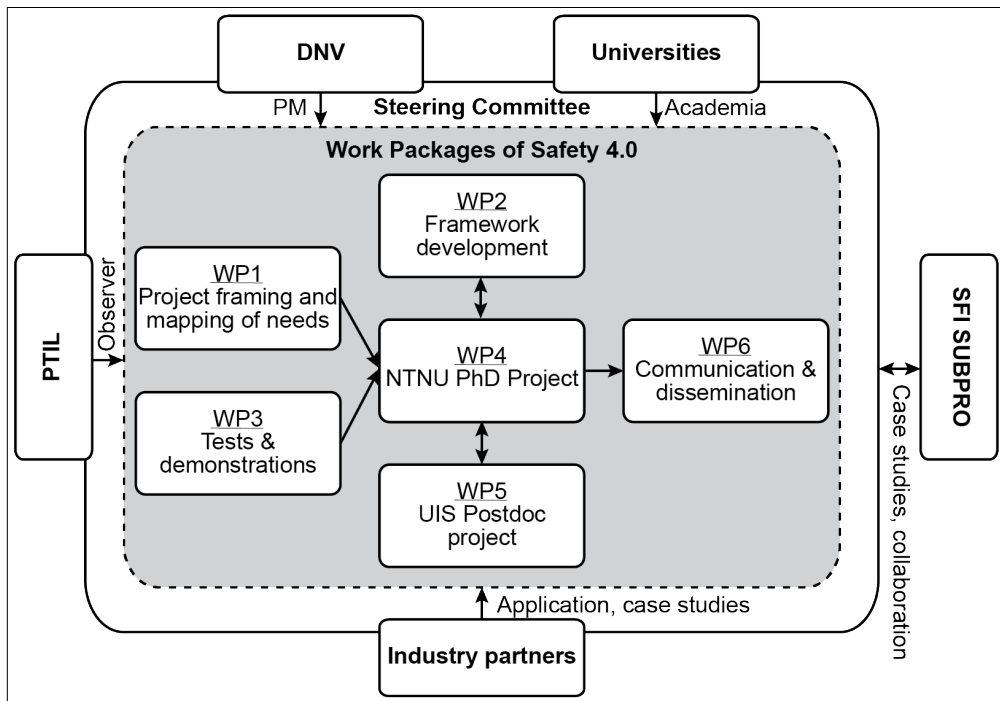


Figure 1.1: PhD project interaction with other work packages and partners of Safety 4.0 project.

for using knowledge characterizing risk descriptions for decision-making, including the use of requirements/risk acceptance criteria in planning and decision-making process.

- WP6 Communication and dissemination. The WP6 objective is to communicate and disseminate research results through various activities for general audiences with relevant expertise.
- WP7 Project management and administration. The WP7 objective is to manage the project within the available project time frame and budget.

1.2 PhD project focus

As part of WP4, this PhD project has an initial objective 'to develop and demonstrate the application of new safety assessment methods that can identify requirements and capture the safety behaviour of novel and complex subsea systems. This objective has been revised for clarity and precision and is discussed later in Section 4.2'. In brief, this PhD research covers the functional safety topic and focuses on the complex operational behaviour of novel software-intensive systems. The revision arises from clarifications of industrial challenges and status and gaps from academia. The PhD project started in Fall-2018 and has a duration of three years. It has close relations with other work packages. See Figure 1.1.

The gaps and needs from WP1, together with the use cases (UC) from WP3, guide WP4's focus. The link is described in Chapter 2. The PhD project was designed to be use-case

driven, meaning that the starting point was to solve local challenges of the specific UC and then generalized the proposed contribution to be applicable at a general framework level. Furthermore, I have interacted with the people involved in framework development in WP2 and the postdoctoral work in WP5 for knowledge exchange during the work process. The interaction was performed to supplement the ongoing work and to allow discussions from a different perspective. Finally, the resulting contributions from the project have been disseminated to the general audience through collaboration with WP6 as described in Chapter 5. WP7 is not illustrated in the figure since it involves the overall process of all the WPs. The PhD was responsible for following the proposed project schedule and report for any delays. The outer layer indicates the involvement of the project participants with the WPs. The PhD project is affiliated with SFI SUBPRO to allow possible research collaborations.

1.3 Terms and definitions

An important starting point before going deeper into technical discussions is to define the frequently used terms. This is to avoid ambiguity of the meaning for the terms since they may have different definitions depending on the referred literature. Table 1.1 presents the summary of the terms that are used in the Thesis, based on vocabulary definitions' list by the International Electrotechnical Commission (IEC) [6, 7], Institute of Electrical and Electronics Engineers (IEEE) [8], and International Organization for Standardization (ISO) [9–11]. The following chapters may present additional terms and definitions that are used within the context of the chapter.

Table 1.1: Glossary of key terms.

Terms	Definitions	Ref.
Error	Discrepancy between a computed, observed, or measured value or condition and the true, specified, or theoretically correct value or condition.	[6]
Failure	Loss of ability for an item to perform as required.	[6]
Functional safety	Part of the overall safety relating to the equipment under control (EUC) and the EUC control system that depends on the correct functioning of the E/E/PE safety-related systems and other risk reduction measures.	[7, part 4]
Hazard	Potential source of harm.	[9]
Maintenance	Combination of all technical and administrative actions, including supervisory actions, intended to retain an item in or restore it to a state in which it can perform as required.	[6]
Process	Set of interrelated or interacting activities that transforms inputs into outputs.	[6]
Qualification test	Procedure to verify conformance to the requirements of a specification.	[6]
Random hardware failure	Failure occurring at a random time, which results from one or more of the possible degradation mechanisms in the hardware.	[7, part 4]
Reliability	Ability to perform as required, without failure, for a given time interval, under given conditions.	[6]
Requirement	Need or expectation that is stated, generally implied, or obligatory.	[11]
Risk	Combination of the probability of occurrence of harm and the severity of that harm.	[9]
Safety	Freedom from risk, which is not tolerable.	[9]
Safety function	Function to be implemented by an E/E/PE safety-related system or other risk reduction measures that is intended to achieve or maintain a safe state for the EUC, in respect of a specific hazardous event.	[7, part 4]

Table 1.1 Continued: Glossary of key terms.

Terms	Definitions	Ref.
Safety integrity	Probability of an E/E/PE safety-related system satisfactorily performing the specified safety functions under all the stated conditions within a state period of time.	[7, part 4]
Software	Intellectual creation comprising the programs, procedures, data, rules, and any associated documentation pertaining to the operation of a data processing system.	[7, part 4]
Software-intensive system	Any system where software contributes essential influences to the design, construction, deployment, and evolution of the system as a whole.	[8]
State-transition diagram	Diagram showing the set of possible states of a system and the possible single step transitions between these states.	[6]
System	A set of interrelated items that collectively fulfil a requirement.	[6]
Systemic failure	Failure at system level which cannot be simply described from the individual component failures of the system.	[10]
Systematic failure	Failure, related in a deterministic way to a certain cause, which can only be eliminated by a modification of the design or of the manufacturing process, operational procedures, documentation, or other relevant factors.	[7, part 4]
Verification	Confirmation, through the provision of objective evidence, that specified requirements have been fulfilled.	[6]
Validation	Confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled.	[6]

1.4 Main report structure

The remainder of Part I: main report of this thesis is organized as follows. First, Chapter 2 presents the industrial challenges that became the background of the project. The industrial challenges are anchored against the current status and gaps from academia as presented in Chapter 3. These resulted in the formulation of the PhD project research questions and objectives in Chapter 4. Then, a scientific research methodology was defined to answer the questions in Chapter 5. Next, Chapter 6 discusses the key results and contributions made to answer the research questions. Finally, Chapter 7 summarizes and concludes the PhD works and provides remaining challenges that need to be solved by further research. The interrelations between each chapter are illustrated in Figure 1.2.

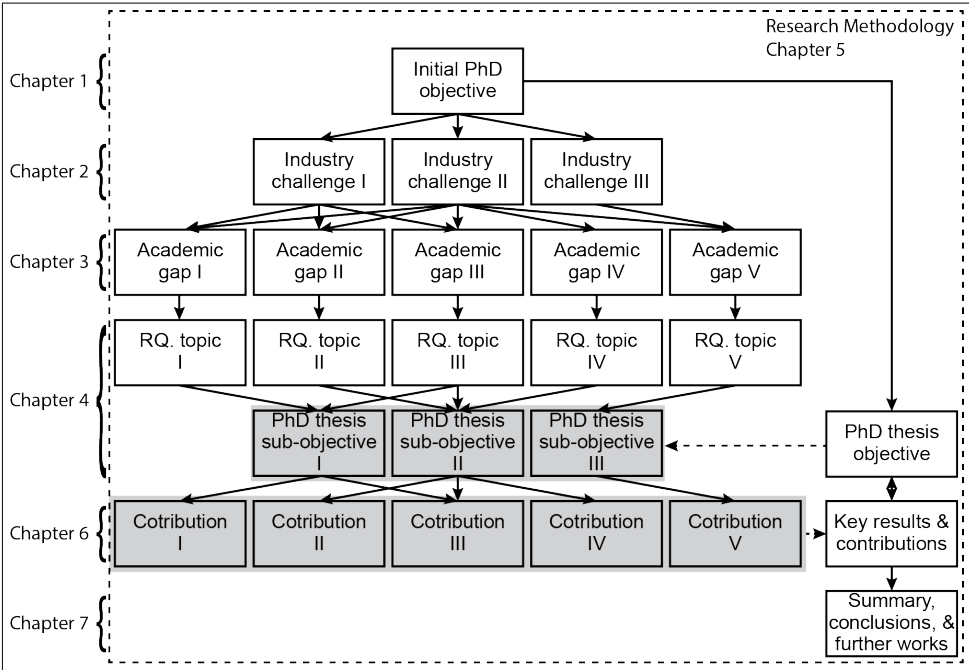


Figure 1.2: The overall structure of this PhD thesis.

Chapter 2

Industrial Status and Challenges

This chapter provides relevant status and challenges from the oil and gas industry that shapes the PhD's research topic. The chapter starts by introducing subsea systems for those that are not familiar with such systems in section 2.1. Next, a subsea risk picture is discussed in section 2.2 to understand the potential impact of introducing novel technology to subsea applications. Then, section 2.3 gives most attention to the systems that have been selected as use cases for this PhD project. It is followed by the introduction of important regulatory frameworks and industry standards that are relevant to safety systems and technology qualifications in section 2.4. Afterwards, the technology qualification program and the functional safety practices that have been established in the industry are clarified in section 2.5 and 2.6 respectively. Finally, section 2.7 concludes this chapter by highlighting the challenges of introducing novel subsea technology in Norway, especially when the selected use case has new/different characteristics.

2.1 Subsea systems

Subsea systems are part of the offshore oil and gas industry that utilizes technology developed to recover hydrocarbons in deep water areas. The first successful installation was for well production in the Gulf of Mexico in 1961. Nowadays, subsea systems are present in various water depths that can be classified into shallow water (<200 m), deepwater (200-1500 m), and ultra-deepwater (>1500 m) areas.

Oil and gas strategy for the 21st century (OG21) [12] for the Norwegian continental shelf outlines several priorities to ensure the competitiveness of the petroleum industry with the growing global energy market. Recent research [13] confirms the high relevance of strategy with the current global condition and emphasizes that subsea technologies are vital for the realization of marginal fields and to increase the efficiency of the offshore facilities. One of the strategic objectives of OG21 is to develop innovative technologies that aim to achieve, e.g., production optimization, digitalization, and protection of the external environment. These technologies are represented in the company vision for research and development (R&D), e.g., all subsea [14] or the subsea factory [15].

The depth of the water, size of the fields, and characteristics of the hydrocarbons would determine the technologies utilized to recover hydrocarbons to the land. For example, a well could be connected directly or via manifold for multiple wells to either an offshore topside facility via riser, floating production, storage, and offloading (FPSO) via riser, or onshore facility via an export pipeline [16]. Figure 2.1 illustrates standard technology in subsea systems. This technology may include a combination of subsea production systems and subsea

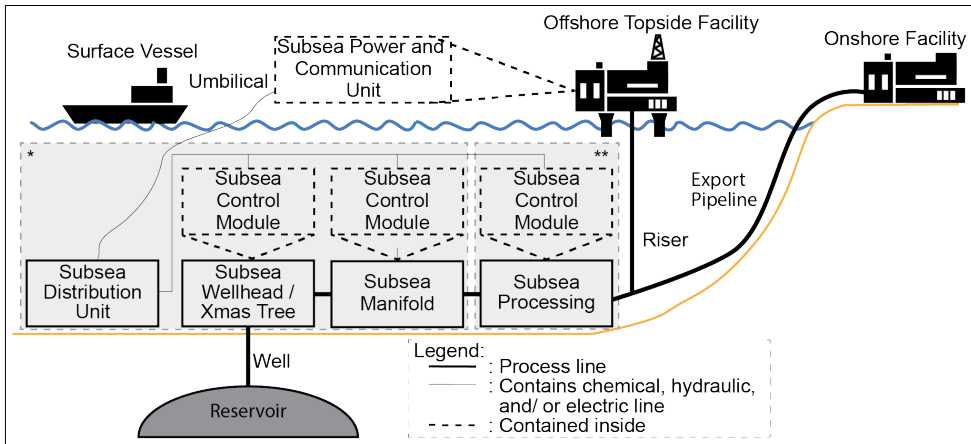


Figure 2.1: Simplified illustration of subsea systems for an oil and gas field. (Note, *: Subsea production systems; **: Subsea processing systems).

processing systems explained in the following subsections.

2.1.1 Subsea production systems

Subsea production systems, see Figure 2.1 in the left grey box area, consist of [17] subsea drilling systems¹, subsea wellheads and Xmas trees, subsea manifolds and jumper systems, a subsea distribution unit, tie-in¹ and flowline system, umbilical, subsea installation¹, and subsea control systems. Before production, subsea drilling systems are deployed to drill the seabed into the prospective reservoir area. If the drilling process is successful, subsea installations are implemented to build the required production elements, i.e., subsea wellheads, Xmas trees and manifolds. During operations, hydrocarbons are recovered from the reservoir through subsea wellheads and transferred to the manifolds for collection and distribution. A Xmas tree is installed on top of a wellhead to assemble valves, pipelines, and some subsea control systems equipment to regulate the hydrocarbons flow from the wellhead. Jumper systems are short pipe connectors linking the subsea equipment, e.g., between Xmas trees and manifolds. The subsea control systems operate the hydrocarbons' flowlines in the Xmas trees and manifolds through an individual subsea control module. Power supply and communication signals for the operation of subsea control modules are obtained from the topside subsea power and communication unit. The subsea control systems are explained in more detail in section 2.1.3.

There are several alternative architectures for well assembly: satellite, clustered, template, and daisy chain. Satellite wells are one or more individual wells that are located remotely and connected to a tie-in system. It has flexibility for tailored design and operation. For comparison, clustered wells are an arrangement of several wells located on a central subsea system. While it has lesser operational flexibility, clustered well architecture can share most of the subsea components [17]. Hence, reducing the cost of installation and operation. These clustered well may also be modularized under a well template for more reduction in installation time. Finally, daisy chain architecture joins either one or more satellite-satellite

¹Not depicted for simplicity

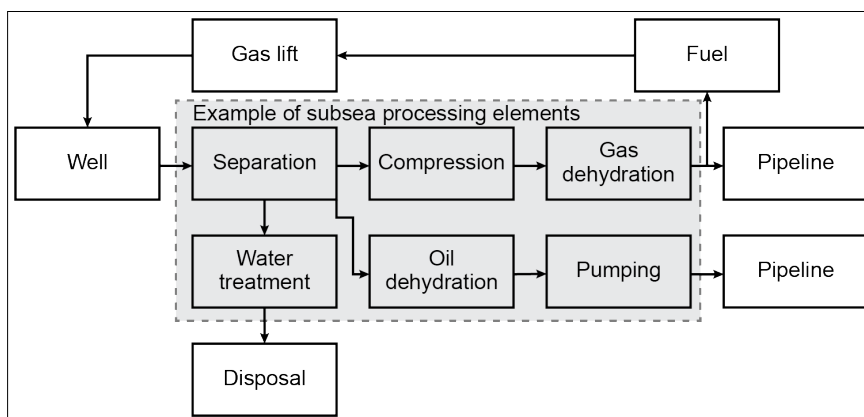


Figure 2.2: A typical process for gas lift (adapted from [17]). Grey box represents example elements that can be used for subsea processing systems.

or satellite-clustered/templates into a series structure to allow the combined use of flow-lines.

2.1.2 Subsea processing systems

Recent technology allows the migration of some topside processing equipment to subsea, named the subsea processing systems. Subsea processing is utilized to preprocess the hydrocarbons at the seabed before delivering them to the facilities. The concept is attractive since it may reduce the capital expenditure of a topside facility and improve flow management while enabling marginal field developments [17]. For example, Figure 2.2 shows a processing facility for a gas lift typically located topside. The gas lift process is deployed as an artificial recovery of hydrocarbons from the reservoir. Some of the equipment, enclosed in a grey box, could be deployed subsea.

In general, the required elements of subsea processing systems depend on the field's process requirements and needs. Subsea processing systems may include subsea liquid boosting, separation, gas compression and treatment, solids management, heat exchanger, and chemical injection. Subsea boosting is utilized to boost the low flow pressure for liquid fluid from the reservoir. A separator is utilized if the hydrocarbons need to be separated for transportation into either two (i.e., gas-liquid) or three phases (i.e., gas-water-oil). Subsea gas compression and treatment is used similarly to subsea boosting. However, it is limited to gas fluid. Hence a treatment process is required to remove liquid from the gas to prevent damage to the equipment. Solids management, heat exchanger, and chemical injection are utilized for flow assurance of the fluid. Each of them would have their control system through a subsea control module. The first commercial success for subsea processing system was the Tordis field in 2007, which utilized subsea separation technology operated by Statoil (now Equinor).

2.1.3 Subsea control systems

Subsea control systems, see Figure 2.3, are utilized to operate the subsea production and subsea processing systems. They are connected through an umbilical and are distributed to

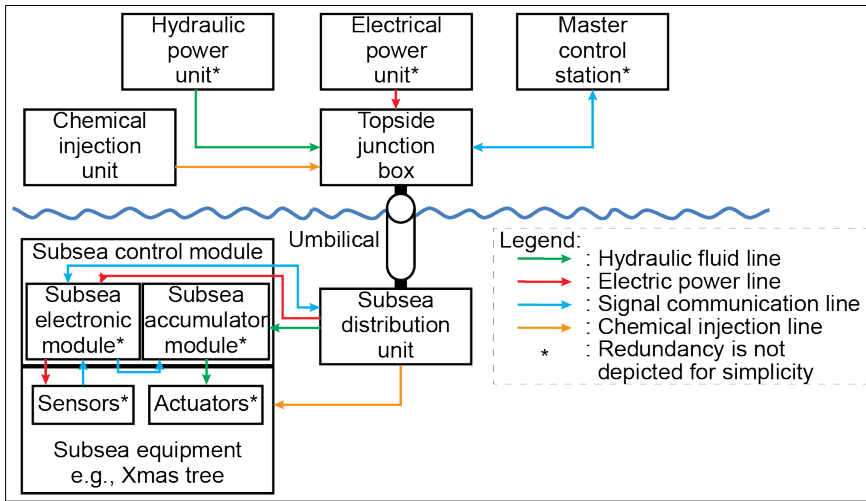


Figure 2.3: Element of subsea control systems utilizing electrohydraulic technology.

each subsea control module via a subsea distribution unit. The subsea distribution unit distributes hydraulic power for valve operation, chemical injection for flow assurance, electrical power to operate the electrical equipment, and signals for information and control. The subsea control systems have two purposes: to achieve the most optimum production, performed by the process control systems (PCSs), and ensure the safety of humans, equipment, and environment, performed by the safety instrumented systems (SISs). PCSs and SISs may functionally consist of a sensing device, logic processing device, actuating device, and transmission device connected to a controlled process forming a control loop (see Figure 2.4). The functions are often performed by redundant components, i.e., A and B, to ensure the availability of the system's performance.

Subsea sensors perform the sensing function and may consist of different sensor types, e.g., to detect pressure, temperature, or valve position. First, they are put at important locations, designed to detect the actual process condition. Then, the information is converted into electronic signals, e.g., 4-20 mA or 1-5 V, via transducers for transmission in the control loop.

There are two levels of logic processing devices for subsea equipment: topside and subsea. A master control station is located topside and is designed as the central processing unit for operation. It can be controlled manually by the operator via a human-machine interface. Some of the responsibilities of the master control station are valve control, interlocks, alarm management, emergency shutdown, and trend/historical data reporting. The subsea control modules perform the logic processing function subsea. Each subsea equipment has its control module for faster response time. The control module is connected to the power supply (electronic power or hydraulic fluids) from the topside via umbilical. Typically, a subsea control module is only responsible for transmitting the information from the subsea to the topside and lets the master control station determine the appropriate control action. However, during a loss of connection to the topside, it is responsible for bringing the subsea equipment to a safe state to prevent or mitigate the escalation of damage.

The actuating function is typically performed by an actuator connected to, e.g., valves, variable sensing devices, or circuit breakers. At first, the actuation of the valves is supported only by hydraulic fluids. Unfortunately, it has a slow response time and limited oper-

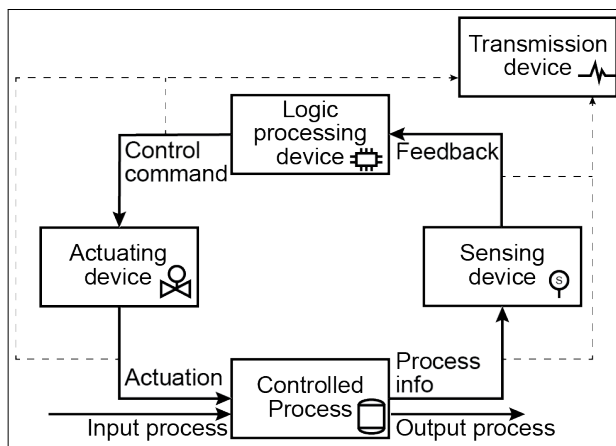


Figure 2.4: A control loop.

ational flexibility [17]. Currently, the most common control operation is supported by electrohydraulic technology. It combines the control system operation via electrical/electronic/-programmable electronic (E/E/PE) equipment to actuate the accumulated hydraulic fluids for subsea operation, see Figure 2.1.3. The most recent technology, called the all-electric, aims to replace all the hydraulic operated equipment with electronics, i.e., battery for power supply. An all-electric well, operated by Total, has been a pilot project in the Dutch North sea since 2016.

SISs are implemented to perform a safety function, ensuring to protect systems with high criticality. SISs are developed with higher integrity than PCSs through a more comprehensive design process to ensure reliable operation. Often, the components of SIS is redundant to ensure high availability upon demand for activation. Also, the safety system operation is not allowed to be overwritten by the operator. It is common to have a PCS and an SIS for protecting the same system in layers. However, they are designed to be independent of each other to prevent unwanted interaction at the SIS.

2.2 Subsea risk picture

Hydrocarbons, as the main product of the subsea oil and gas industry activity, contain hazardous substances. The release of hydrocarbons would affect the ongoing operation of the subsea systems. For example, in the 2021 'Eye of Fire' incident event, the leakage of underwater hydrocarbon gas [19] results in an underwater fire that harms the surrounding sea environment. This example is only a glimpse of the subsea risk picture for the subsea systems.

Kim et al. [18] clarifies precisely where the hazardous events can occur in subsea systems installation, as shown in Figure 2.5. The events of concern are located in different areas and consist of mainly: hydrocarbon leakage, equipment damage, and blowout at the topside facility. These hazardous events can be caused by external hazards (e.g., trawling, ship anchors, dropped objects), long-term hazards (e.g., failure mechanism, material defects, structural stress, or erosion and corrosion), or inherent hazards (e.g., well pressure or pressure build-up due to process). These hazards may have risks to the environment, destroying the ecosystem, to the equipment, affecting the productions and assets, or to the humans, which

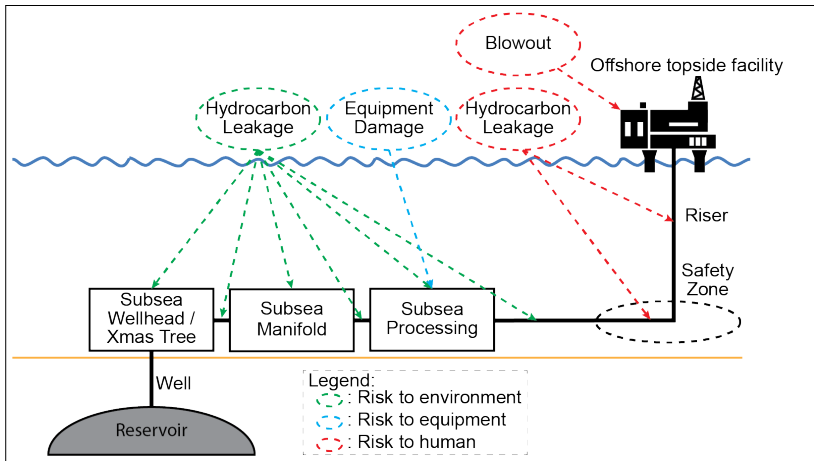


Figure 2.5: Subsea risk picture (adapted from [18]).

leads to injury or death.

Uncertainties in the introduction of novel solutions may affect the subsea risk picture. For example, they may introduce new or different mechanisms that can be another source of hazards. Hence, it is vital to identify precisely the source of uncertainty and predict the consequence to ensure that the novel solutions are safe for operation.

2.3 Selected use case

Safety 4.0 project has utilized novel subsea technology concepts and challenges that industry partners have proposed as use cases (UC): 1) all-electric safety systems, 2) integration of process control and safety, 3) Demonstrating safety of novel subsea technology based on API RP 17V [20]. This PhD project focuses on UC 2, which can be described as follow.

UC 2 is a technical solution where PCS and SIS are fully or partially integrated by utilizing the capability of software-intensive systems. Full integration refers to the complete sharing of any redundant devices in a control loop, e.g., sharing the logic solvers for PCS and SIS operations. While partial integration only covers sharing either one of the redundant components or integration in the hardware but not in the software. Integration may be realized at different system levels (e.g., component, subsystem, and system) and applied to different component device functions. Some example applications of the integration concept are shared sensors through split transmission line [21], shared logic solvers with separation in software [7] (see illustration in Figure 2.6), shared valves with separated actuators [22], or shared transmission line with separation in the data priority [3]. In Figure 2.6, it can be seen that all the redundant components of sensors and actuators can still be kept separate. However, since the hardware of the logic solver is shared, the PCS and SIS are now partially integrated. This integration may lead to new interactions that cannot be seen before.

Integration would reduce complexity in physical architecture, leading to cost efficiency [23]. However, the integration concept is still not widely used due to some practical implications such as increased software complexity, different operational and maintenance requirement, and difficulty to obtain evidence of safety [3, 23]. Further research and development, including the qualification process, are vital to resolve these practical issues. In this PhD

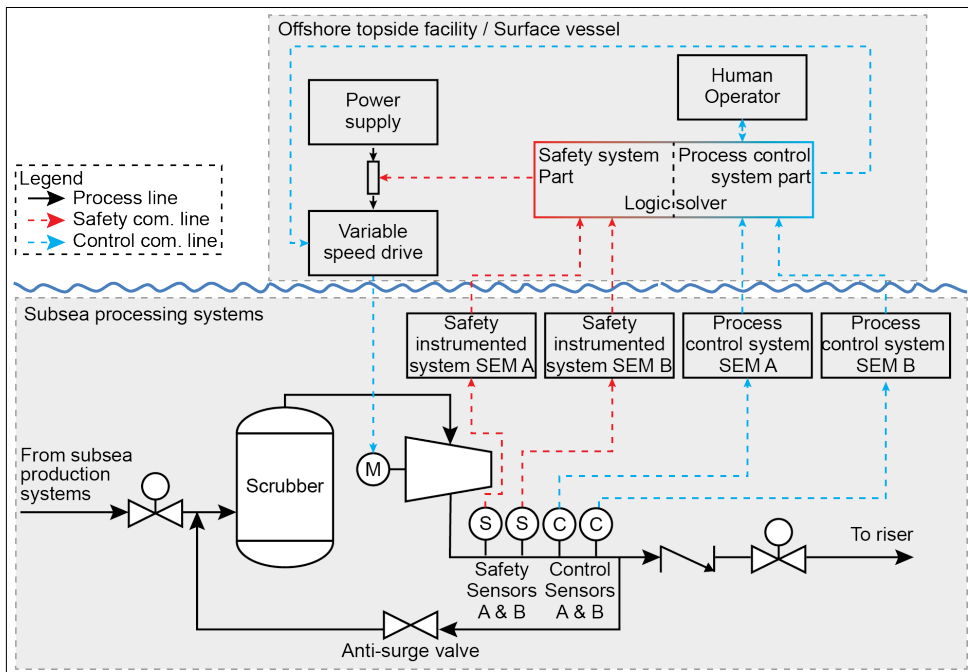


Figure 2.6: Example of process control system and safety instrumented system on a subsea compression system having integration on the logic solvers.

project, the integration concept application is limited for the subsea processing systems focusing on risk to equipment where PCSs and SISs are implemented [4].

2.4 Regulations, guidelines, and standards

The development of approaches, methods, and technologies is aimed to solve particular problems. They have been developed according to the specifications, technology feasibility, available resources, and budget. The specifications shall follow the local acts, regulations, international and local industry standards, and guidelines. Hierarchically, the relations between them are shown in Figure 2.7.

The acts and regulations are legally binding and depend on the country where the technology would be applied. They are developed to communicate the intention of the local authority. For example, Petroleum Safety Authority (PSA) manages the petroleum activity within the Norwegian continental shelf. The regulations' guideline is developed to clarify the regulations and may also provide the recommended standards to be followed for each topic. For petroleum activity, the regulations and the guidelines include:

- The Framework regulations [24], providing a framework for petroleum activities related to health, safety, and environment (HSE).
- The Activities regulations [25], regulating the policy for various activities.
- The Facilities regulations [26], governing the design and outfitting of facilities.

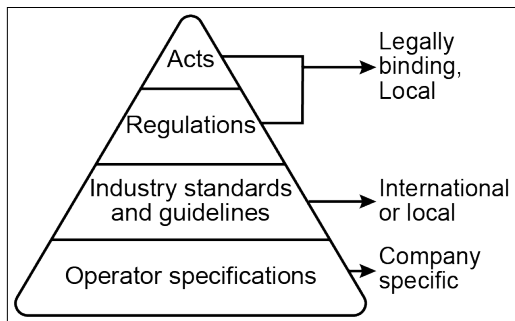


Figure 2.7: Hierarchy of acts, regulations, standards, and operator specification in Norway.

- The Management regulations [27], managing the requirements relating to HSE.

Standards encode the recommended practice best on experience in the industry. Hence, they could only be built when technology reaches a specific maturity level. International standards are generic enough to allow different or new practices on managing and developing specific topics, e.g., safety, in different countries. Local guidelines and standards may be developed to clarify the intention of the standards and are tailored based on the local regulations. The relevant standards and guidelines for safety demonstration of novel technologies in the oil and gas industry are:

- Generic facility design, the local standards are: NORSOK P-002 [28], for the design of process piping and equipment on offshore production facilities and NORSOK I-001 [29] and I-002 [30], for technical and functional design requirements of field instrumentation and safety and automation system.
- Subsea system design: NORSOK U-001 for local standards, API RP 17A [16], and ISO 13628 [31] for the international standards.
- Risk: NORSOK Z-013 [32] for local standard and ISO 31000 [33] for international standards in risk management. Currently, risk management is a mandatory starting point for any oil and gas industry project in Norway [27, 32].
- Technology qualification: API RP 17N [34] or DNV-GL-RP-A203 [35].
- Safety and functional safety: NORSOK S-001 [22], for technical safety and NOG 070 [36] for guidelines on functional safety application in Norway. As reference, the international standards are: IEC 61508 [7], IEC 61511 [37], and API RP 17V [20], a recommended practice for the design of safety systems for subsea applications.
- Reliability: ISO/TR 12489 [10] for modeling and calculation, and ISO 14224 [38] for data collection.

Individually, each operator or company may derive additional specifications based on its criteria and policy. However, the specifications may not be available for a wider audience and can only be obtained when the operator performs specific projects.

In the following sections, more attention is given to technology qualification and functional safety topics, including the standard practice, due to its relevance with the expected industry challenges.

Table 2.1: Technology novelty categorization (adapted from [35]).

Application area	Degree of novelty of technology		
	Proven	Limited field history	New or Unproven
Known	1	2	3
Limited knowledge	2	3	4
New	3	4	4

1) no new technical uncertainties; 2) new technical uncertainties;
3) new technical challenges; 4) demanding new technical challenges;

2.5 Technology qualification

A technology qualification program (TQP) is required to prove that the functionality of the novel technology is reliable. It means that the probability of failure should be as low as possible while having low uncertainty. TQP is an iterative program that can be performed at a particular level of technology development, e.g., concept evaluation, pre-engineering, and detailed engineering. Each result of TQP would end up as a milestone of the project development. These milestones signified the technology maturity and are measured by the technology readiness level (TRLs) that is advocated by the US department of defense [39] and API RP 17N [34]. For each TRL, the developer would face a decision gate, determining whether follow-up research and additional investments should be made or if the technology is deemed unprofitable and should be aborted. This process ensures that the remaining uncertainties are known, accepted, and managed during the operations.

Novelty is defined as 'the quality of being new or unusual'. [40, Novelty]. From the definition, using the term quality means that each technology has a different degree of novelty. In fact, not every novel technology is completely novel. Novelty can still be achieved even if only some parts of the proven technology from the other industry are integrated with the technology already used in practice. Completely novel technology would require significant R&D efforts before implementation.

Novelty in technology may require novel assessment, installation, operation, or maintenance, leading to uncertainties. For example, an operator needs to apply a novel maintenance regime for an integrated safety system to avoid unwanted shutdowns in the control system. In the oil and gas industry, DNVGL-RP-A203 [35] proposes to categorize the uncertainty brought by the novel technology based on two dimensions: application area and degree of novelty. Table 2.1 shows the resulting matrix of technical uncertainties and challenges involved with novel technology based on the DNV recommended practices. In addition, another category for novel technology has been proposed by API RP 17N [34].

2.6 Functional Safety

Functional safety is achieved by using an active safety barrier relying on electrical/electronic/programmable electronic (E/E/PE) components. An example of a safety system is a high fluid pressure detection system that prevents overpressure by opening a discharge valve. Passive safety barrier, e.g., pressure-resistant pipe, is not covered by the functional safety process. However, the integration of both systems is still managed by the functional safety concept.

The process to ensure functional safety is systematically described in the safety lifecycle [7] as shown in Figure 2.8. In practice, demonstration of safety is achieved by producing

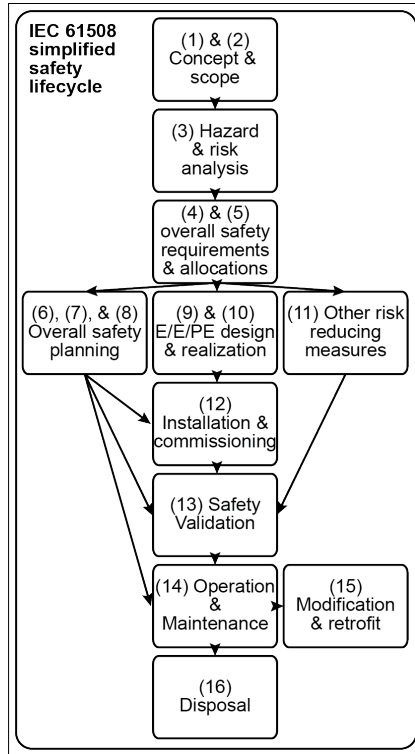


Figure 2.8: Functionals safety lifecycle process based on IEC 61508.

evidence that every step in the safety lifecycle has been performed correctly. In the railway industry, all evidence is compiled systematically as a safety case [41]. Management of functional safety, performed by a neutral third party, is required to ensure the quality of the functional safety process and the produced safety systems.

The framework for functional safety is a risk-based approach, meaning that the target of implementing a safety system is to eliminate or reduce risk to a tolerable level. The requirements, translated to the required safety integrity level (SIL), are allocated based on the target risk reduction for the system. This can be done through, e.g. [37], a layer of protection analysis or risk graph method. IEC 61508 and IEC 61511 classify the SIL into four, from SIL 1 to SIL 4, with increasing reliability for a system with a higher SIL number. The SIL cover three aspects: hardware, software, and systematic. The assessment of the system's safety integrity is based on the weakest link principle. If one of the three aspects has lower SIL than the others, then the overall safety integrity of the system would follow the lowest integrity value. For example, a system with hardware SIL of two and software and systematic SIL of three would have an overall system SIL of two.

The SIL values can be obtained by qualitative or quantitative means. There are two benefits for the quantification: 1) ease of definition of the risk acceptance criteria, and 2) ease of statistical evidence gathering to comply with the criteria. In practice, not every scenario could be quantified accurately. IEC 61508 guided the treatment of scenarios based on whether it is removable or quantifiable, as shown in Figure 2.9. Scenarios able to be removed by system design would disappear during operation. Non-quantifiable scenarios require sys-

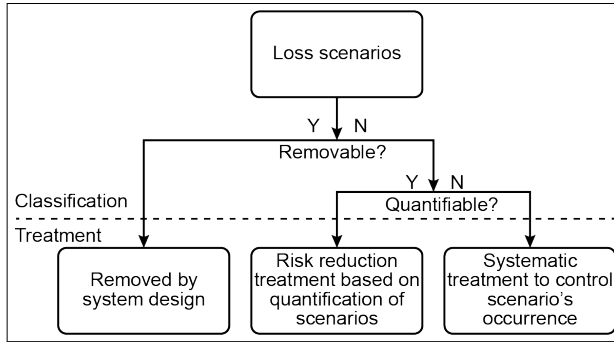


Figure 2.9: Treatment of loss scenarios derived from IEC 61508 process.

tematic treatment to control the occurrence since they may arise due to emergent properties of complex systems. It is necessary to have information on the technical details to discover the correct treatment accurately. Finally, the quantifiable scenarios are modelled based on the system's reliability, availability, and maintainability (RAM) performance.

Functional safety is governed by a generic standard IEC 61508 [7]. In addition, several sector-specific standards have been developed to fit the characteristics of each sector, like IEC 61511 [37] for the process industry, EN 50126 [41] for the railway industry, IEC 61513 [42] for the nuclear power industry, and ISO 26262 [43] for the automotive industry.

2.7 Challenges of introducing novel subsea technologies

The implementation of technologies for subsea application has several challenges concerning safety, as follows (adapted from [4, 12–14, 18, 44, 45]):

- *New/different philosophies.* Subsea technologies may implement new/different safety philosophies to accommodate environmental conditions and long tie-back to the facilities. For example, a separated PCS and SIS would have no direct interaction that affects each system's functionality. Comparatively, an integrated solution would have new interactions affecting their functions. This would result in different failure modes from the traditionally known technologies. The operator must keep up with the required expertise and adjust to new practices. Also, the new philosophies may not be supported by the current regulations and standards. A systematic procedure needs to be taken to ensure system compliance.
- *Increased digitalization of control and safety system.* Software-intensive systems may improve process control and operation, diagnostic and prognostics, and safety operations. However, the systems are more complex than the traditional mechanical systems. Complexity in the system may result in systemic hazards due to unknown interactions in the system. Therefore, safety assessment methods need to cover additional types of hazards to ensure the safety integrity of the system.
- *Remaining uncertainties in applications.* Technology qualifications aim to reduce the uncertainty on novel technology. Nevertheless, the remaining uncertainties cannot be removed due to our inability to replicate the full working conditions and the incompleteness of the assumptions. Furthermore, the different risk profiles subsea exacerbate

bate these uncertainties. Therefore, the risk assessment must consider the knowledge dimension to adapt to the consequences of assumptions' deviations when assessing the technology.

These challenges are generic for the novel subsea solutions involving software-intensive systems. However, the challenges need to be more precise to be used as the PhD research focus direction. Therefore, several specific topics have been refined from the generic industry challenges for introducing novel subsea technologies. They are explained in the following subsections.

2.7.1 Challenge I – Increase in complexity due to integration of software-intensive systems

Complexity in the system's functionality due to the integration of software-intensive systems may result in new and unpredictable systems behaviours. This unpredictability would increase the required time during the technology qualification process for discovering all possible behaviours and may reduce our perceptive confidence with the results due to remaining uncertainties. Also, a simple breakdown of the systems that are typical approaches used by the commonly used safety assessment methods cannot capture the emergent behaviours due to component interactions at the system level.

2.7.2 Challenge II – The non-compliance of technology against the local regulations and standards

Several regulations and standards in Norway advocates the design of control system and safety system to be built independently of each other to ensure that failure of one system does not affect the other system (e.g., see [36, section 8.7]), [22, section 11.4.7], and [26, section 33, 34]). This is known as the N-1 principle, which ensures the robustness of the protection system. This requirement conflicted with the integration concept in UC 2, where independence could not be assured anymore. In practice, this requirement is initially meant for topside facilities. Independence between equipment for the subsea facility may increase complexity on physical architecture, e.g., subtle dependency due to shared equipment, increased number of interlinked equipment, and the need to have multiple redundancies for each system. The architecture complexity may become another source of hazards that needs to be managed. The nonconformity of the UC 2 concept with the regulations and standards may affect the design process decision. The boundary in which the standard can accept integration level is unknown. Industry practitioners often choose to avoid the problem by having complete independent solutions. International standards, such as IEC 61508 or API RP 17V, are less restricted since they are based on risk-based requirements. Comparatively, PSA developed the regulations based on two main principles (the "duality principle"): 1.) Minimum requirements, 2.) Risk-based requirements. 1.) means that there are specific minimum requirements that need to be followed. Other non-specified solutions must be managed based on a 2.) risk-based approach. PSA emphasizes that the minimum requirements may be challenged by industry practitioners as long as the new solution is "as better as" the traditional solution. Therefore, as a starting point, research is required to clarify the non-compliance of technology with the regulations and standards.

2.7.3 Challenge III – Lack of safety demonstration process framework for the complex software-intensive systems

Assuming that the previous two challenges could be solved, providing a framework for the safety demonstration process of the novel technology is still necessary. A simple answer to resolve the problem is maybe by following the safety lifecycle approach. However, it is unknown whether the available methods fit with the increasing complexity of software-intensive systems. Hence, it is necessary to research the methods for each phase of the safety lifecycle and determine their feasibility for the safety demonstration process.

Chapter 3

Academic status and gaps

This chapter gives an academic status on safety demonstration concerning complex software-intensive systems, which is the topic given the most attention in this PhD thesis. Clarifications should be made based on the industry challenges to explain the reasoning for the topics' selection. Section 2.7 narrows the industrial challenges into three major topics: increase in complexity due to integration of the software-intensive systems, the non-compliance of technology against the regulations and standards, and lack of safety demonstration process for the complex software-intensive systems. From the wordings, all three challenges revolve around the term complexity. Complexity is one of the main characteristics of software-intensive systems. Therefore, it is reasonable to start this chapter by characterizing the complex aspects of software-intensive systems as a premise on why the topic is challenging and of interest to researchers from various areas in section 3.1.

When discussing non-compliance between the technology against the regulations and standards, it is important to pinpoint which standards and what topics they cover. Discussions in sections 2.4, 2.6 and 2.7 guide into a topic, i.e., functional safety. The functional safety concept utilizes a risk-based approach that has not been described properly before. Hence, section 3.2 follows by introducing hazards and risk assessment concepts and the available methods consolidating the concept. Afterwards, section 3.3 associates the hazards and risk assessment concept with the safety assessment and demonstration practice based on the functional safety concept. Finally, the academic discussions are narrowed toward gaps, leading to challenges that need to be solved in this PhD project in section 3.4.

3.1 Complexity?

Complexity is referred to as the main challenge of utilizing software-intensive systems [46]. In general, any technology would have a different level of complexity. Cambridge Dictionary [40, complexity] define complexity as 'the state of having many parts and being difficult to understand or find an answer to'. A system having many parts does not always mean it is complex. The complexity lies in the difficulty of understanding its unpredictable behaviour. Any system that can be understood by spending significant effort to predict the system behaviour based on the elements would be classified as a complicated system. Some examples of a complex system are the human brain, stock markets, and societies.

Johansen and Rausand [47] further elaborates the characteristics of a complex system:

- *Multiplicity and diversity.* A complex system often has many and diverse components. With a higher number of components, the system is prone to perturbations.

- *Interactivity and non-linearity.* Elements in a system interact with each other. This interaction may be caused by, e.g., physical structure, functionality, environmental conditions. Non-linear interaction in a complex system increases the difficulty in understanding the causality of phenomena produced by the system.
- *Intractability, bounded rationality, and sense-making.* The functionality of a complex system is intractable, i.e., difficult to explain. Different interpretations may arise due to limited information and resources available.
- *Emergence, self-organization, and adaptation.* The aphorism by Aristoteles describes emergence as the whole is more than the sum of its parts [48]. Emergence arises due to self-organization and adaptation at the systems level that does not present when looking at the constituting elements.
- *Teleology and migration.* Due to multiple and differing objectives that a system may have, the system design may unintentionally become complex. The stakeholder is responsible for defining the safe boundary of the system that can satisfy these objectives and manage the resulting complexity.
- *Drift toward the edge of chaos.* A complex system is dynamic and continuously evolve due to interactions in different contextual condition. Increased experience with the system may normalize the initially unacceptable condition. Hence, increasing the chance for the system to produce unknown behaviour, leading to an unwanted condition.

The characteristics above highlight the difficulty of understanding complex systems. Researchers attempt to grasp the concept of complexity by describing aspects of the systems into models. The drawbacks are that models are only partial abstractions of the systems, providing an incomplete picture. For example, a single model of system structure could not capture the emergence behaviour of the complex system. Therefore, a combination of several models [49] is required to describe complexity. Also, the models need to be at a different level of abstraction (LoA) [50] to have a holistic perspective of the system. Using LoA means that the analyst acknowledges that complex systems may exhibit different sets of behaviours at different levels of abstraction, making it easier for treating the problems separately. However, maintaining consistency in the model would still be a challenge when modelling the complexity [49].

The modelling of complexity is further described by Bunge [51] in his model that any system, including complex system, can be modelled as a quadruplet of $S = \langle C, E, S, M \rangle$. Composition (C) refers to any part building the system (e.g., human or hardware components). Environment (E) refers to any items (e.g., temperature, gravity, or other parts) outside the defined boundary of a system that may influence the system. Structure (S) refers to the link between parts of the system. The link can be physical (e.g., hardware structure) or abstract (e.g., society structure). Mechanism (M) refers to the processes that define the system behaviour (e.g., physical or functional mechanisms). These four aspects of the CESM model are interrelated, meaning that their combinations are necessary to have a systemic perspective for explaining the system behaviour. For example, when describing the systems' dynamic, information such as the timing (M) and context (E) for every related (S) element (C) in the model is required.

In this thesis, the CESM model is treated as a metamodel that captures the combination of various models at different LoA. For example, figure 3.1 shows how different models

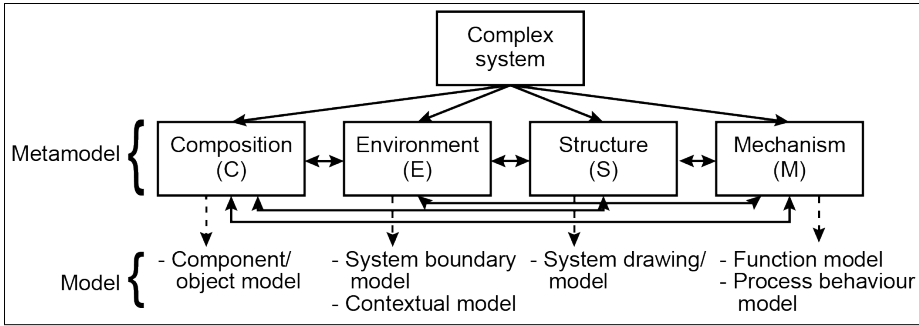


Figure 3.1: The metamodel of complexity as combination of a variety of models (adapted from [52]).

may be generated individually and can be linked to other models for explaining particular systems behaviour.

3.2 Hazard and risk

Hazards released during hazardous events may affect humans, environments, or assets/-finances. Depending on the equipment and process involved, various types of hazards, from chemical, mechanical, electrical, and noise to interaction hazards, may be present in the system. Therefore, systematic treatment measures are mandatory to prevent or reduce the harm caused by the hazards.

A concept called risk is defined to assess hazardous events systematically. Risk is formally modeled as a triplet of $R = \langle E, L, S \rangle$ [53]. It means that risk conveys information on the likelihood (L) for the occurrence of a hazardous event (E) would have a consequence severity (S) of loss. According to Aven and Reniers [54], the likelihood value is subjective, depending on the context of the assumptions and knowledge. Hence, the risk value has uncertainties.

Safety, as per definition, is also associated with risk. It is a condition that can be achieved when the risk of the system has been captured and managed. However, for complex systems, safety is paradoxical. If the complexity of the technology behaviour is not fully understood, how could we be confident in the technology's safety? Jensen and Aven [55] resolves this issue by linking risk with the knowledge dimension. They argued that despite the incompleteness in our analysis, it is still helpful to have a management process, specifically on risk, to have more knowledge of the problems. A risk management process is an activity performed by the organization to manage the risk of hazards, thus ensuring safety. It includes the process of establishing the context, risk assessment, and risk treatment process. It is crucial to acknowledge the assumptions' uncertainties at every step of the risk management process. The analyst needs to prepare for deviation, and a follow-up process is required to reduce the remaining uncertainties. The preparation can only be done if most of the complexity characteristics in a system can be captured. The following subsections discuss the aspects of risk management in more detail. The context establishment explanation is skipped since it has been partially discussed when explaining the complexity.

3.2.1 Risk identification

As part of risk assessment, risk identification is a process to identify the source of risk, i.e., hazard. It is also known as hazard analysis, which identifies hazards, hazardous events, causes, and consequences often described as scenarios. Hazard analysis is often performed by people from different backgrounds, giving diverse perspectives to the issues. The analysts are assisted with relevant information of the system, such as technical documentation, process diagram, and functional list. The hazard analysis process can be performed iteratively when new information is obtained. This iterative process ensures that the new or different scenarios caused by the updates are identified.

Several approaches have been commonly applied in the industry [56, 57] such as preliminary hazard analysis (PHA), functional hazard analysis (FHA), failure mode and effect analysis (FMEA), and hazard and operability study (HAZOP). Each method has a different focus, approach, and fitness with the system design phase. Recently, there are new methods that have been developed to address complexity in technology, such as systems-theoretic process analysis (STPA) [58] or functional resonance analysis method (FRAM) [59].

3.2.2 Risk analysis and evaluation

The risk assessment process also includes risk analysis and evaluation. A risk analysis is performed as a follow-up of the risk identification either via qualitative, semi-quantitative, or quantitative approaches. The selection of risk assessment approaches is dependent on the problems and the available information. For example, only qualitative assessment can be performed during early design since data is limited and the concept is abstract. However, when the scenarios could lead to high consequence events, a follow-up quantitative approach would be mandatory to obtain a more precise risk value [32].

Risk metrics are defined to analyze and evaluate the risk level. They are distinguished between risk to people or risk of a system. In this thesis, we focus more on the risk of a system. The most common risk metric for a system is a risk matrix. The risk matrix utilizes a pre-defined scale for its severity and likelihood value, e.g., from minor damage to catastrophic for the severity or from improbable to fairly normal for the likelihood. In some methods, additional variables may be introduced. For example, failure, mode, effect, and criticality analysis (FMECA) introduces detectability to measure the system capability to detect the hazard [56]. In addition, the risk priority number developed for STPA utilizes a level of knowledge variable to measure the uncertainty [60, 61].

The measured values may be used for evaluating whether the risk is broadly acceptable, acceptable based on tolerable risk level, or not acceptable and require significant risk reduction measures. The acceptance criteria for the risk should be predetermined before the analysis to ensure the objectivity of the results.

3.2.3 Risk treatment

Risk treatment refers to the process of treating the tolerable and not-acceptable risk. Tolerable risk level refers to the fact that the risk of some events cannot be removed or reduced efficiently. Hence, the tolerable risk would be accepted, and the operator needs to be prepared for mitigation measures. The effort for risk treatment is also affected by the selection of technology. Risk-informed decision making (RIDM) is a branch of the deterministic decision-making approach that includes the risk value together with other design criteria to choose a design solution [56, 62].

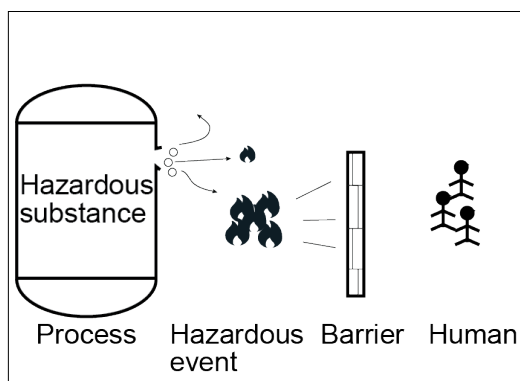


Figure 3.2: Illustration of hazard, hazardous event, and barrier.

Various measures can then be utilized to prevent or reduce the risk of hazards. It is preferable if the risk can be eliminated. When not possible, one of the most effective means to reduce the risk is a safety barrier, as illustrated in Figure 3.2. They are often designed in layers, hence defined as a layer of the protection system. A safety barrier may require an activation process (active) or not (passive). An active safety barrier requires an independent control system such as a safety system [63], while a passive safety barrier is inherently designed [64]. The implementations of safety barriers are important to maintain the safety integrity of the process. In the process industry, a control system is often credited as an additional protection layer to reduce risk [23]. The concept of functional safety governs the design of an active safety barrier.

3.3 Safety assessment and demonstration

Safety assessment is a systematic process performed to ensure the safety of the protected system. The objective is to arrive at the judgment on the adequacy of the functional safety achieved by the E/E/PE safety-related systems [7]. The process is applied to all phases throughout the overall safety lifecycle specified in Figure 2.8.

Conceptually, the results of the safety assessment are pieces of puzzles that can be arranged to demonstrate safety, see Figure 3.3. Kelly and Weaver [65] emphasizes the importance of the argumentation process to demonstrate how the safety of a system can be concluded reasonably. During the process, several questions are asked:

1. What are the main requirements for the SIS in light of the assessed hazards and risks?
2. What types of evidence can be obtained from the system?
3. What are the assumptions on the evidence?
4. Given the evidence and assumptions, how can safety be argued?
5. Is the justification for the relevance of the argumentation method sufficient?
6. What supporting requirements and sub-safety argumentation structure are required to satisfy the main requirements?

The following sections discuss each aspect in more detail.

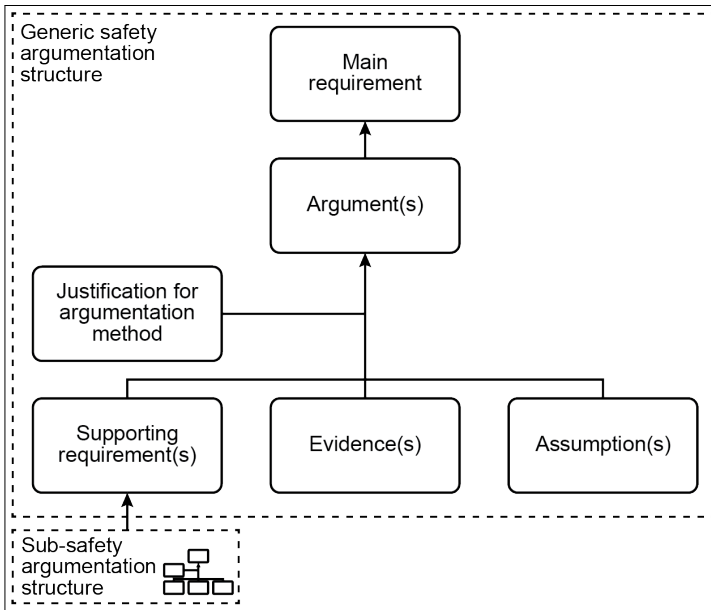


Figure 3.3: Safety argumentation concept.

3.3.1 Requirement generations

Requirements are the targets to be achieved when developing a system. They are often standardized, e.g., in regulations, standards, and company specifications, to formalize the final product and avoid recurring mistakes during system development. For new technology, requirements can be derived based on various principles, such as safety or process. When the system is large, it is common to distinguish between main requirements and supporting requirements. Main requirements are often generic and cover all elements in the system, while supporting requirements may be more detailed and sub-system/element specific to constitute their unique needs. Formally, requirements should have several properties to facilitate the design process. According to Holt et al. [66], the properties are identifiable, clear, solution-specific, have ownership, have an origin, are verifiable, able to be validated, and have priority. These properties shall limit the context where the requirements would be applicable.

There are two ways of forming a requirement: prescriptive and goal-based [64]. The implementation process of both requirements types is illustrated in Figure 3.4. Prescriptive requirements usually have a precise meaning. A safety engineer performs verification to ensure the conformity of the design. When there are deviations in the system design, a decision process is required to determine whether the prescriptive requirement is still valid or not. Goal-based requirements are indirect and require iteration and analysis processes. Hence, there would be a multitude of approaches to satisfy the goals. After sufficient experiences have been obtained, the implementation of goal-based requirements can be considered as best practices and translated as prescriptive requirements. The use of either requirement type depends on the context of the problems and the system's maturity.

Stakeholders with differing objectives and from different disciplines may affect the formulation of requirements. In this case, there is a high chance for either conflicts or dependencies in the requirements. For example, during hazardous conditions, a safety objec-

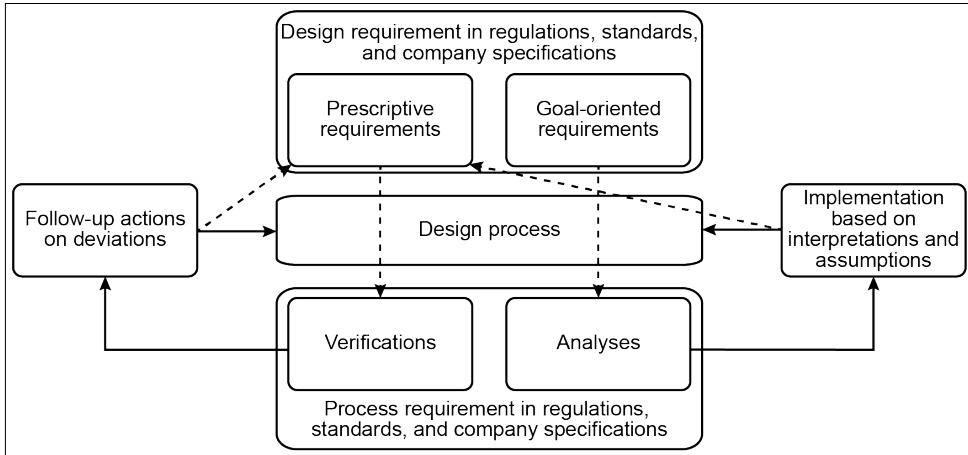


Figure 3.4: Effect of different requirement types on design process (adapted from [64]).

tive could be achieved by shutting down the equipment. However, an unwanted shutdown would affect the availability objective of the production system. Therefore, a systematic approach to manage conflicting goals is necessary to balance between various objectives.

3.3.2 Evidence and assumptions

Evidence is facts and information, proving that the requirements have been achieved. They can be obtained from the related parts of the system (e.g., human, technical, and organizational aspects). Assumptions may limit the validity of every piece of evidence within a specific context. For example, a requirement that the system shall be reliable could be proven by a reliability test report when performed during specific contextual assumptions.

Haugen [67] specifies three properties that define the evidence gathering process: intensity, rigour, and detachment. Intensity refers to the scope of the process, while rigour refers to the number and depth of the techniques used to generate the evidence. The process itself should be performed by a detached and neutral third party to ensure integrity. These three properties would have differing effort levels depending on the criticality of the system and process involved [7]. Unfortunately, evidence-gathering processes are limited by available resources and time. For example, the software has an inherent complexity that makes it impossible to test all possible scenarios [46]. Test cases are usually derived based on the worst-case scenario and are assumed to represent the other possible scenarios. This issue should be communicated with the decision-maker to ensure that the uncertainties are considered.

Nair et al. [68] proposes a taxonomy of safety evidence, consisting of at least 49 types of primary safety evidence that are classified between the process and the product (or system) involved. Examples of the evidence are system inception specification, activity planning, activity records, activity resources specifications, safety analysis results, system specifications, code, verification and validation results, and system historical service data specification. These different pieces of evidence types may result in extensive documentation [69]. Respondents in Nair et al. [70] work answered that 79% of the evidence management is still processed manually, emphasizing the challenges in the evidence management process.

3.3.3 Argument and justification

An argument links the relevance between safety requirements and evidence. Justifications methods are necessary to argue that a particular piece of evidence is relevant. For example, the reliability testing report as evidence for a specific system reliability requirement does need further justification since the evidence proves the requirement directly. Comparatively, when the requirement is imprecise, such as 'the system should be safe', a reliability testing report evidence is insufficient. Here, justification needs to be made, arguing that the safety of this particular system is achieved by having a specific reliability value. Safety arguments may be supported by several pieces of evidence and supporting requirements to be adequate. Additional techniques, such as qualitative assessment, checklist, quantitative assessment, and logic-based assessment, are required to assess the adequacy of evidence [68].

There are at least three types of techniques to structure the safety argument [68]: argumentation-induced evidence structure, model-based evidence specification, and textual templates. Several argumentation approaches utilize these techniques, such as free text, tabular structures, claim structures, traceability matrices, bayesian belief networks (BBN), and goal structuring notation (GSN) [71]. Argumentation-induced evidence structures are preferred due to their capability to show the link systematically [65, 71]. The relationships between requirements and evidence are recorded through traceability links. Changes may occur in the link due to modifications or new knowledge. For example, evidence is a 'living' artefact, with increased information and lower uncertainties in the assumptions through the system's operational time. Therefore, a follow-up process is necessary to clarify the relevance of the traceability link. This is a complex process since the intricacies and number of links involved may be too large to see the big picture of the safety arguments.

3.4 Gaps in academia

Further research on the topics covered in this chapter led to several academic gaps that need to be solved to resolve the industry challenges. The academic gaps are described in the following subsections.

3.4.1 Gap I – Unavailability of the safe design principles

Safe design is defined as 'the process of designing attributes and features into a design that enables its implementation to achieve the required level of safety' [72]. Principles are defined as the general intention of implementing the required attributes. The safe design principles are the basis for developing various safety requirements across different industrial sectors. However, these principles are seldom outlined and often lost due to years of development. Furthermore, researchers focus on developing approaches for qualification of new technologies and standards by emphasizing the process [73, 74], while some others focus more on the safety demonstration aspect of new technology [70, 75] without considering the safe design principles. Hence, it has been challenging to modify the existing requirements, which recommend particular technical solutions, without understanding their principles. Figure 3.5 shows the link between the safe design principles, relevant regulations, standards, and guidelines, and the implementation of technologies. Technology requiring new design philosophies, e.g., on the implementation of integration, would not be compliant with the current regulations and standards. Therefore, clarification on the principle of safe design beyond the detailed clauses is necessary to guide the discussion in this case.

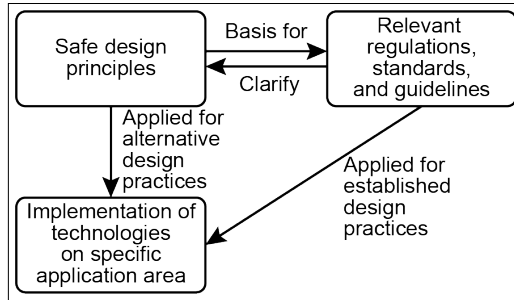


Figure 3.5: Relations between the design principles, technology, and formalized requirements.

3.4.2 Gap II – Ambiguous safety requirement generations methods for complex system

IEC 61508, for example, specifies goal-oriented safety requirements where the users shall define their acceptance criteria based on hazard analysis and risk assessment process. Hazards are typically identified using negative scenarios. First, negative scenarios are gathered by identifying unwanted system states (e.g., failure or deviation). Then, requirements to avoid, prevent, or mitigate the unwanted system state is derived according to several parameters: cause, effect, ability to detect, and criticality. Various methods utilized these approaches, such as PHA, FHA, FMEA, and HAZOP. However, the integration in UC 2 may increase the complexity of the control and safety system. As a result, the systems are becoming more vulnerable to systematic faults, which alone or in combination may result in emergent failures at the system level. A key concept to solve this lack of understanding is to identify the related scenarios [76] at a systemic level.

However, the approaches above are incapable (or not designed) of discovering systemic interaction problems [58]. They share common traits: to analyze (or break down) the system into smaller parts and consider mainly the effect of failure (or deviation). Component behaviours sometimes do not reflect the behaviour at the system level [51]. An approach that focuses more on the interaction at the system level rather than on the individual component's performance is required. Some examples of the methods are STPA and FRAM. Critical system properties may include interdependencies and cause-effects relationships. It is still unknown if the methods addressing systemic properties are truly more suitable than the traditional methods already used in the industry.

3.4.3 Gap III – Unknown classification methods for the integration concept

The safety requirements are tailored to the systems' characteristics. A technology solution is selected from several alternatives that fit best with the requirements, e.g., efficient cost, low complexity, or high safety performance. Due to various possibilities for integration, the number of architectures to be considered may grow exponentially. This may pose a problem during the safety demonstration process as the systems are difficult to classify and compare.

From the literature, we found that IEC 61508 [7], for example, classify three levels of integration based on the hardware and software involved: complete separation, integrated hardware with software separation, and complete integration. In the process industry do-

main, CCPS [3] proposes five levels of classification based on the communication networks: air-gapped system, interfaced system, integrated system with isolated networks, integrated system with a shared network, and combined system with strong dependency. In the automotive (e.g., [77]) and maritime (e.g., [78]) domains, the proposed classification of the automation level of the vehicle is based on the interaction between the systems, i.e., control, action, and system state. The ship automation level, for example, could be classified into three major categories: conventional, smart, and autonomous, focusing on the logic solvers integration for the automation process.

Unfortunately, there are no generic classification methods that distinguish the different levels of integration for every component in the control and safety system. Some of the available classifications only cover integration in one of the components [3, 7, 21] or the interactions between components [77, 78]. Therefore, there is a need to modularize the solution space by classifying them into different integration types and provide individual treatment to increase the efficiency of the process.

3.4.4 Gap IV – Need of a reliability performance’s modelling approach to generate evidence for the complex system’s safety

One example of the verification and validation results is simulation results. For safety systems, these simulations are obtained from reliability modelling of the system performance. There are three classes of modelling formalism [79]: 1. (probabilized) boolean, 2. (stochastic) finite-state automata, and 3. (stochastic) process algebra.

A system in the boolean model only has two states of working and failure. The boolean model examples are fault tree analysis (FTA) and reliability block diagram (RBD) [10]. On the other hand, the dynamics of a system in finite state automata (FSA) are modelled through a finite number of state transitions. Markovian model, Stochastic Petri Net (SPN), and FSA-based textual model are examples of the FSA [10, 80]. The process algebra model improved the FSA model with two new characteristics: it allows an infinite number of states, and the creation or destruction of components may occur after the transitions [79]. Agent-oriented modelling [81] is the only known modelling technique for process algebra. Currently, no safety standards have recommended the use of the process algebra model.

In chronological order (1-3), the modelling formalism has increased in modelling complexity and decreased expressiveness. FSA has been recommended to model the complex scenarios [82, 83]. However, FSA still has some practical challenges related to modelling uncertainty. Inaccuracy of the modelling process may lead to loss of information [62]. Furthermore, if different analysts produce different models for the same system, they lead to different results, reducing their trust in the model. Therefore, a systematic modelling approach should be developed.

3.4.5 Gap V – Need of clarification of framework for safety demonstration

Typically in IEC 61508, the framework for safety demonstration is achieved through developing arguments on textual templates, answering each requirement separately. However, Nair et al. [68] argued that this process is inefficient and error-prone, especially when the number of requirements is large. An alternative approach with easier visibility is to use the argumentation-induced evidence structure [65]. The remaining challenge is to develop and fit the identified methods and approaches with the framework.

Chapter 4

Research Questions, Objectives, and Delimitation

This chapter presents the research questions targeted in this PhD project, with a basis in the mapping of industrial challenges (section 2.7) against the academic gaps (section 3.4). Clarifications with the industrial experts, academic supervisors, and colleagues working on the same topic shape the interpretation of the research questions. The research questions have been organized in section 4.1 under the heading of five main topics, i.e., I – Safe design principles, II – Solution specific safety requirements, III – Alternative concepts, IV – Effect on risk, V – Safety argumentation. The anchoring of the five topics in the safety lifecycle of the IEC 61508 is also presented. Every topic serves as a piece that contributed to solving the puzzle on the main objective of the PhD work. As the initial PhD project objective was too generic, this thesis developed more detailed objectives in section 4.2. The PhD thesis objective contains three sub-objectives derived from the research questions' topics. Finally, section 4.3 described the delimitation of this research project.

4.1 Research questions

Figure 4.1 illustrates an adaptation of the simplified safety lifecycle phases from IEC 61508 and the scope of research for this PhD. Research questions have been derived with one to one correspondence to the academic gaps previously mentioned and explained in the following subsections.

4.1.1 Topic I – Safe design principles

Figure 4.1 shows that the (I) safe design principles, coupled with the (9) high-level safety requirements, may guide the development of (10.3) technology solutions. However, as explained previously in the gap I in section 3.4.1, the safe design principles are unavailable. Therefore, currently, the only known method for developing technology solutions are from high-level safety requirements. This situation becomes a problem when the available safety requirements are developed for traditional technology and limit the allowable novel solutions. In addition, even if it is known, the impacts of utilizing the safe design principles for developing the requirements and for demonstrating safety are unknown. Hence, the relevant research questions within this topic are:

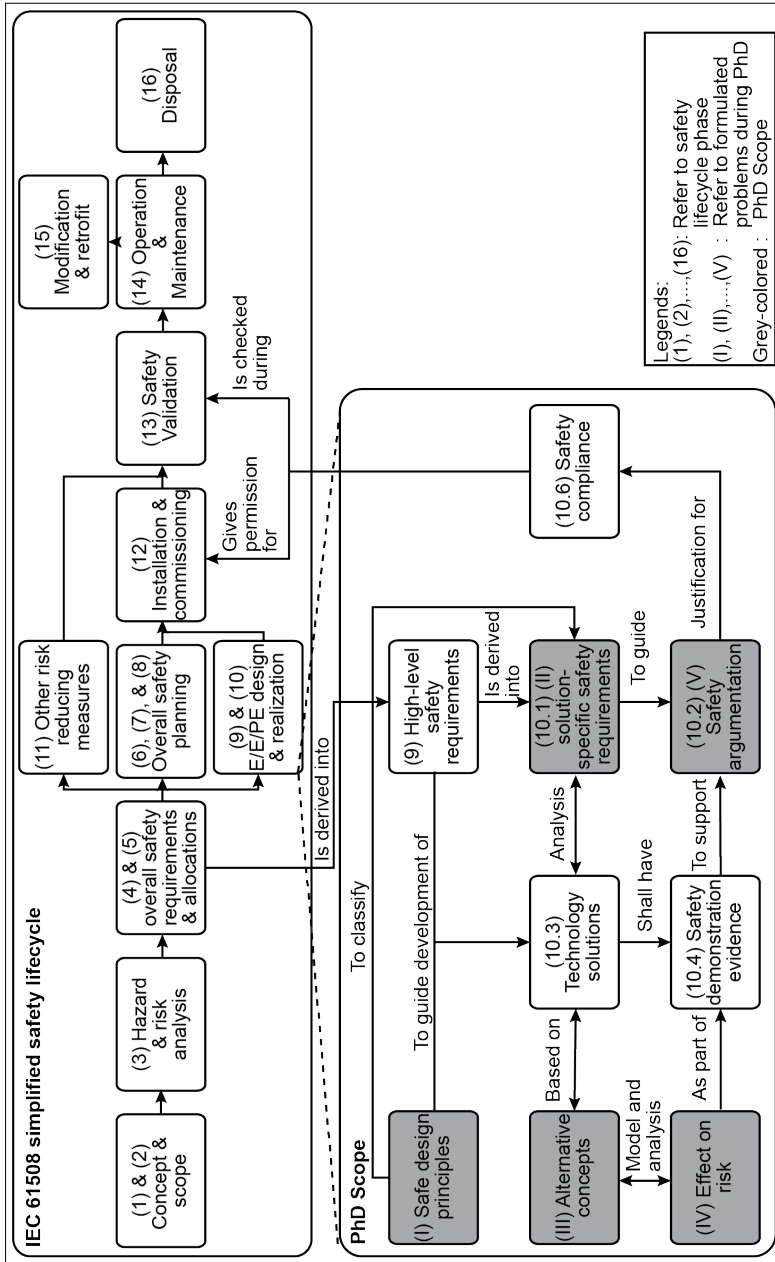


Figure 4.1: PhD scope based on simplified safety lifecycle of IEC 61508 [7].

- What are the safe design principles?
- How do the safe design principles shape the development of requirements?
- How do the safe design principles affect the safety demonstration process of novel technology?

4.1.2 Topic II – Solution-specific safety requirements

In figure 4.1, (10.1) (II) solution-specific safety requirements box is linked to (9) High-level safety requirements, (I) safe design principles, and (10.3) technology solutions. High-level safety requirements are the main requirements derived from the hazard and risk analysis process in the safety lifecycle. While, solution-specific safety requirements are supporting requirements that capture unique characteristics of the technology solutions, e.g., the interactions problem or failure modes. The safe design principles may be used to group the solution-specific safety requirements and ease the resolution process.

The long list of hazard analysis methods mentioned coupled with their respective characteristics in gap II in section 3.4.2 makes the selection of the most suitable method to generate the solution-specific safety requirements a challenge. Therefore, the following research questions are formulated to guide the selection of the method.

- How do the selected hazard analysis methods identify the same or different functional hazards?
- How do the selected hazard analysis methods provide a systemic perspective for the analysis?
- What are the main differences between the derived solution-specific safety requirements?
- What is the most suitable method to generate solution-specific safety requirements considering the complexity of the use cases?

4.1.3 Topic III – Alternative concepts

In Figure 4.1, (III) alternative concepts are choices for the proposed (10.3) technology solution within the context of (I) safe design principles and (9) high-level safety requirements. Context refers to the boundary of the intended operations as assumed during the design [72]. These alternative concepts may also have an (IV) effect on risk, explained later in the following subsection. As explained earlier in gap III in section 3.4.3, there are no classification methods to group the integration concept. This may make the follow-up safety requirement generation process inefficient and resource consuming. Hence, the following research questions are formulated to resolve the problems.

- What are the available concepts for integrating process control and safety systems in the subsea application?
- Given the available integration concepts, what are the safety requirements?

4.1.4 Topic IV – Effect on risk

Following the previous section, the selection of a (10.3) technology solution has its consequences. The (III) alternative concepts would have a different (IV) effect on the risk of the systems. The decision-makers need to balance the trade-offs between cost, complexity, and safety. The resulting (IV) effect on risk assessment would also be used as (10.4) safety demonstration evidence. However, the methods for assessing the different effects on the system's risk for the complex software-intensive system still needs further development. Considering the characteristic of UC 2, the modelling approach should also have the capability to capture the dependency effect of the integrated system as discussed earlier in gap IV in section 3.4.4. Therefore, the following research questions would be answered precisely:

- How to quantify the hazardous scenarios derived by the hazard analysis process?
- How does dependency addressed by the model?
- How does the new modelling technique performed if compared with the available methods recommended in the standards, i.e., ISO/TR 12489 [10]?

4.1.5 Topic V – Safety argumentation

For the final topic, figure 4.1 shows that (V) safety argument get its input from (9) high-level safety requirements (indirectly via (10.1)), (10.1) (II) solution-specific safety requirements, and (10.4) safety demonstration evidence. The safety argument concept combines them to build (10.6) safety compliance as part of the safety lifecycle process. However, as discussed earlier in gap V, there is a need to clarify the framework for a safety demonstration. In addition, the results from RQ topics I - IV need to be clarified for their suitability with the safety argumentation concept (see Figure 3.3). The following research questions are derived as guidance for research.

- How does safety argument support the safety demonstration process for novel technology?
- How are the methods and approaches identified from the previous research results linked with the safety argumentation concept?

4.2 Research objectives

The PhD thesis objective is 'to develop and demonstrate the application of new safety assessment methods within the scope of functional safety, which can capture and manage the complex operational behaviour of novel software-intensive systems'. The methods should account for the characteristic of the UC 2 integration of process control and safety and based on the safety assessment process identified in the functional safety standards.

The PhD thesis objective has been derived into several sub-objectives based on the research questions described earlier, as follow:

- I Study the effect of complexity in the software-intensive system's behaviour to safety. Determine the methods to capture these complex behaviours. This sub-objective is linked to RQ topics I and III.

- II Develop an approach that can demonstrate/produce evidence of the safety of software-intensive systems. Assess the scopes and limitations. This sub-objective is linked to RQ topics II-IV.
- III Propose/develop a framework to argue for the safety of complex software-intensive systems. This sub-objective is linked to RQ topic V.

The numbering is only used as a reference when discussing the objective and does not indicate any objective prioritization.

4.3 Delimitation

This PhD project is limited to developing a safety demonstration framework based on functional safety, focusing on software-intensive systems. While the framework is aimed to be generic, different focuses in the concept could produce different findings that affect the interpretation. This PhD project is also centred on developing new methods and concepts for improved hazard and risk assessment. Discussion on the technical solutions is kept minimal due to limited information.

The study case for the integration concept is focused on the subsea processing system. The system has lower risk criticality if compared with subsea production systems based on the regulations. Theoretically, the methods could be applied for any critical systems having similar characteristics. The reason for this selection is due to the closer project collaboration with the industry. The project aims to be practical, even when the industry avoids integration due to low historical experiences and stringent regulations. Therefore, the PhD project started with the concept application on a simple system with lower criticality as a pilot study.

We have derived information from international standards and previous projects shared by the industrial partners in the absence of detailed technical specifications and data. Therefore, the examples are not as extensive as an industrial level system and have been simplified and discussed at a higher abstraction level to avoid roundabout discussion on specific technical aspects. It is crucial to understand the assumptions and the abstraction level when challenging the technical results. Together with the academic supervisor, experts from the industry have vetted the assumptions and results through the project.

Chapter 5

Research Methodology

This chapter covers the research methodology utilized through the course of 3 years project. Section 5.1 starts with a general introduction to research and how it shapes my decision to do research. It is followed by classification of research in section 5.2 to show the characteristic of this PhD research project. Then, significant focus is given on the presentation research approaches used in the project, including the activity and expected project results in section 5.3. Finally, this chapter is concluded with discussions on challenges encountered during the project in section 5.4.

5.1 Research motivation

Research is defined as 'a detailed study of a subject, especially in order to discover (new) information or reach a (new) understanding' [40, research]. Essentially, the research aims to discover novelty in the subjects that it covers. This novelty may be related to either a phenomenon, theory, or problem. In addition, research may be built upon previous knowledge and replace them when new or different information is available.

The process of research is rigorous and systematic. Often, results can only be produced after numerous failures. Positively, failure during research is valuable since it contributed to verifying alternative paths and obtaining the desired results. However, the competition for fundings and citations may worsen the positive-outcome bias, which disfavors negative results [84].

This issue was the initial counterbalance for my decision to pursue PhD. I started my PhD project for two reasons: to experience research as a career and challenge myself with intellectual problems. Initially, I never considered research as a career path. I was constantly shadowed by the fear of imposter syndrome, thinking that I was not smart enough or not capable enough to do research. This is due to the emphasis on the need to produce significant breakthroughs when researching and the prevalent positive-outcome bias in my country. However, experiences with my Master supervisors (also as my current PhD supervisors), Professor Mary Ann Lundteigen and Associate Professor Hyungju Kim, made me think differently. They acknowledged all the contributions I made, small or big, and guided me to produce something I could be proud of. The information from people with different research topics and expertise might be biased and not relevant to me. I have a valid qualification for this PhD project with a Master's degree from the same field, i.e., reliability, availability, maintainability, and safety (RAMS). Coupled with the similarity of the PhD topic with my Master thesis and the interest I have in functional safety from my past working experiences, they are sufficient experiences to contribute something to the field. Honestly, the experiences differ

Table 5.1: Criteria of research and how it is achieved.

Criteria	Purpose	How it is achieved
Novelty	Obtain new findings	The originality of the results is assessed based on relevant literature, vetted by the co-authors, and is peer-reviewed by scientific peers.
Creative	Based on original concepts and hypotheses	Each contribution started by suggesting the application of new methods and concepts. They are adapted or developed from different contextual applications.
Uncertain	Uncertainty about the final results during the process	The development of methods and applications are uncertain. Although it is not reflected in the final articles, it is discussed briefly in Section 5.4.
Systematic	The research work is planned and budgeted	The PhD project is according to a well-defined research plan and adjusted every six months based on the progress status.
Transferable or reproducible	The results could be reproduced	Transfer of knowledge is performed in international conferences, scientific discussions, and seminars. Articles are published together with research data as supporting documents.

a lot (positively) from what I expected, and I feel glad that I made the right choice.

5.2 Classification of research

The coupling between research and development (R&D) aims to utilize the new knowledge to develop new technology applications. There are five core criteria of an R&D activity: novel, creative, uncertain, systematic, and transferable or reproducible. These criteria standardized the aim and process of the R&D activity, even when performed by different actors. OECD [85] distinguishes the activities into three types:

- 1 *Basic research.* This R&D activity aims to discover new fundamental knowledge through theoretical or experimental work. The application of this knowledge is not covered in the works.
- 2 *Applied research.* Similar to basic research, but it is aimed to solve particular problems or objectives.
- 3 *Experimental development.* A systematic work to utilize new knowledge or practical experiences to develop or improve new products or processes.

This PhD project is mainly applied research, aiming to improve the available safety demonstration process with new knowledge from the characteristic of novel technology. This has been done while reflecting on the core criteria of research, as presented in Table 5.1

5.3 Research approaches

Research is approached via the systematic procedure, developed based on the philosophical foundation, the research method, and the design of the process (e.g., experiments or explanatory sequential) [86]. In general, the research methods involved several steps:

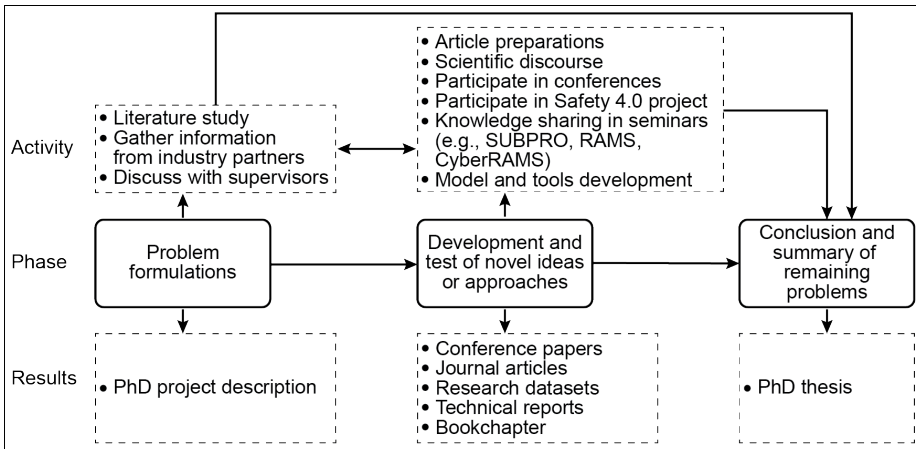


Figure 5.1: The research process, activity, and results.

- 1 Identify research questions.** Research questions are gaps identified from the literature or practical problems. A hypothesis is formed to initiate the direction of the research.
- 2 Data collection.** The data can be obtained through experiments, interviews, or literature searches.
- 3 Data analysis.** Analysis of the data can be performed through statistical inference or logical reasoning.
- 4 Interpretation.** The analysis results are compared against the state of the art of knowledge to check the initial hypothesis.
- 5 Validation.** The resulting new knowledge is validated for robustness, as it would become the foundation for further research.

There are two levels of research approach developed for this PhD project. High-level and RQ-specific. High-level refers to the overall phase of the PhD project, while RQ-specific refers to the individual approach used to solve each RQ.

The high-level approach of the project is split into three phases: problem formulation, development and test of novel ideas or approaches, and conclusion and summary of remaining problems. The approach is illustrated in Figure 5.1, including the activity and results during each phase.

Initially, the problem formulated in this PhD is obtained from discussion with industrial partners and verified through literature study. Literature is collected from several selected search engines, e.g., google scholar, science direct, Wiley online library, and Scopus. A PhD project description is then developed, clarifying the research topics, research questions, and research approaches. Finally, the supervisors vet this project description to ensure the quality and significance of the research direction.

Then, novel ideas or approaches are developed and tested to answer the research questions. The activity is based on the RQ-specific approach and differs depending on the problems to be solved, background knowledge of the researcher, and the target audience of the results. Creswell and Creswell [86] outlined three types of research approaches:

Table 5.2: RQ-specific research approaches.

Research questions	Qualitative	Quantitative	Mixed methods
RQ I. Safe design principle	X		
RQ II. Solutions-specific safety requirements			X
RQ III. Alternative concepts			X
RQ IV. Effect on risk		X	
RQ V. Safety argumentation	X		

- 1 *Qualitative*. The research aims to develop knowledge based on logical reasoning of the data.
- 2 *Quantitative*. The research aims to test objective theories based on measurable variables. Statistical procedures can be performed to analyze the data.
- 3 *Mixed methods*. The research aims to combine both qualitative and quantitative data to yields insight from different perspectives.

Table 5.2 outlined how different RQ was approached in this PhD. Supporting tools are developed whenever required. The RQ-specific approaches are processed through extensive scientific discourses from various activities. These scientific discourses include participation in international conferences (e.g., ESREL), project workshops (e.g., safety 4.0, agile software development), and seminars with the scientific communities (e.g., SUBPRO, RAMS group, and CyberRAMS group). Iteration is performed as needed to ensure high results quality. The results are published as articles in peer-reviewed international conferences and journals. In addition, the activities produced separate technical reports and a book chapter within the same topic to communicate the results for a different target audience.

Finally, the results are concluded, and the remaining problems are summarized in this PhD thesis. The PhD thesis is written in paper collection, documenting the research background, objectives, research questions, approaches, main contributions, and list of articles covered during the PhD.

The author performed the research under the guidance of the PhD supervisors through supervision meetings. The main supervisor, Professor Lundteigen, mainly monitor and discuss the research progress. Associate Professor Kim involves mainly during the work of RQ I and II. Finally, functional safety researcher van der Meulen provides an industry perspective and bridges the collaboration between this project and other work packages in Safety 4.0.

5.4 Challenges and lessons learned

Research is an attempt to go towards an unexplored area to obtain new scientific knowledge. While systematic plans and procedures have been made, there are inevitably unexpected challenges during the process. These challenges are unique for every project. Nevertheless, a systematic approach can solve the issues, hoping that similar challenges can be minimized for subsequent research works. The challenges encountered in the project and the approaches devised to resolve them are elaborated in the following.

- *Unavailability of previous research data*. This PhD is built upon current research knowledge. It is necessary to reproduce earlier works, either as a starting point or for the

validation process. However, most research data is available only as research articles, which are often unclear or insufficient. In this case, it is recommended to contact the main author to clarify their research works. However, this may become a lengthy process since they may have a long period between responses or not reply for various reasons (e.g., change of occupations). It is our responsibility as a researcher to ensure transferability and reproducibility of the research. Therefore, most of the relevant research data in this PhD project has been published. Understandably, not every research data can be shared. Therefore, we must clarify with the research stakeholder and ensure the anonymity of the data before publications.

- *Uncertainty in the research process.* Some of the causes of uncertainties specific to this PhD project are maintaining the intensity of works, the absence of technical details for the UC 2, the dynamic of stakeholders shaping the research, revision of research approach due to negative results, article preparations duration, peer-review process duration. Individually, they may have different magnitudes of effect on the research plans. However, their combination results in a need to revise the research plan every six months to adapt to the research progress. Consultation with the supervisors can be an alternative as they have more experience when dealing with similar problems. In addition, a mid-term evaluation is essential as a checkpoint to clarify the research progress.
- *Worldwide pandemic situation.* Half of the research period (early 2020 - now) is affected by the global pandemic. The pandemic resulted in shifting from physical (in-person) to online activities done from home. In this PhD project, online activities have a significant effect on scientific discourses. During this period, online activities with many participants, e.g., conferences or industry workshops, provide reduced benefits compared to the non-pandemic situation. These reduced benefits are assumed due to reduced eagerness to communicate via the online platform and the belief that other active participants would be. After experiencing different activities, it is found that if they are done in a smaller group of four to five people or done with colleagues, e.g., group seminars, the involvement from the participants are rather high. Hence it would be recommended to do future online activities in this way.

Another consequence of the increased use of online activities is the accentuating psychological effect, i.e., depression due to loneliness. Fortunately, most of the RQ-specific approach works done in this PhD project require no physical experiments. Thus, the shifting to work from home may seem to have an insignificant effect on the research work. However, it must be acknowledged that the office is not only used for work. It is also common to use it as a social activity place with the colleague. The loss of office may result in isolation that can produce negative energy hindering the work process. Weekly online meetings had been devised to be an alternative. If insufficient, it is recommended to find other social activities such as clubs or associations as a replacement.

Even when the challenges are known, and the recommendations are provided, it is unsurprising if these challenges persist differently for other research projects. What has been done as solutions in this PhD would not necessarily be suitable for others. Therefore, it is expected for a researcher to be aware of the possible challenges, be prepared, and be flexible enough to cope with the uncertainties.

Chapter 6

Key Results and Contributions

This chapter presents the summary of contributions from this PhD project. They are structured based on the published articles and linked with the research questions as shown in Figure 6.1. All the detailed processes and results are presented in the articles in part II of the thesis. One of the final contributions is unpublished. Thus, it has in-depth discussions and examples for clarification.

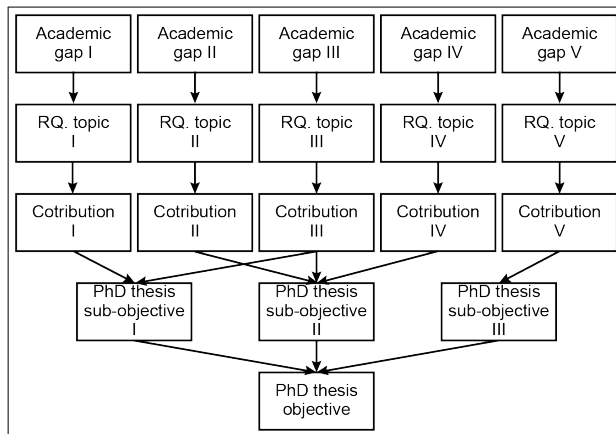


Figure 6.1: Link between academic gaps, research questions, contributions, and PhD objectives.

Figure 6.2 shows the revised version of the safety argumentation concept for novel technology and the contributions of this PhD. In addition, it clarifies the link between some elements in the safety argumentation with the published articles. The revised safety argumentation concept is explained in more detail in Section 6.5.

6.1 Contribution I – Safe design principles

The first contribution is related to the topic I – safe design principle. The detailed problem statement and scientific approaches are discussed in Gap I in section 3.4.1, research question I in section 4.1.1, and article I [87] respectively. The contributions from the article are discussed in the following subsections. They also highlight how the contributions lead to achieving sub-objective I.

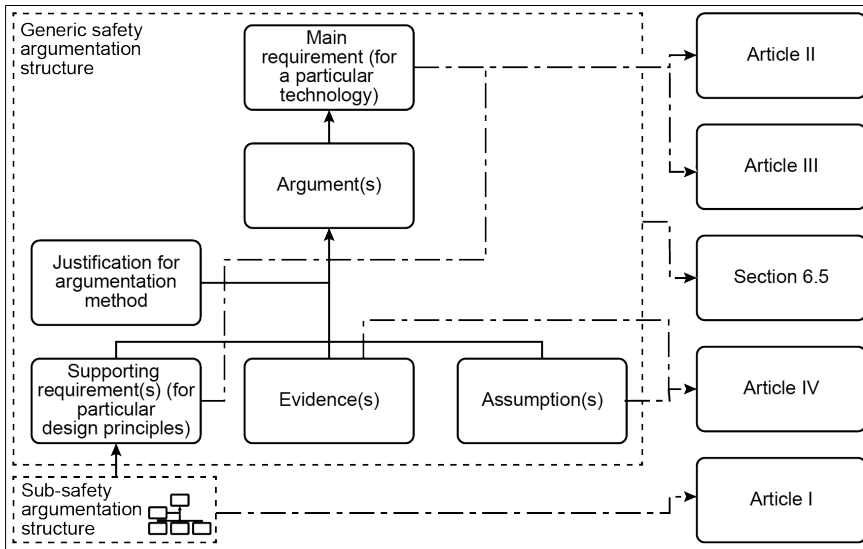


Figure 6.2: Overview of PhD contributions.

6.1.1 Identification of the safe design principles

Based on our literature review, a list of safe design principles identified by Drogoul et al. [72] are the most relevant for developing generic requirements and demonstrating the safety of the system. The safe design principles are: 1) Two levels of design, 2) Strategy against failures and errors, 3) Safety vs complexity, 4) Modularity, 5) Segregation, 6) Documentation, and 7) Demonstration of safety. These principles are derived for three reasons: to build the system safely (1-2), to select safer solutions given the recommended safe attributes (3-5), and to justify the safety of the solutions (6-7).

6.1.2 Alignment of the requirements in IEC 61508 part II with the safe design principles

IEC 61508 part II provides functional requirements that recommend several safe attributes for a system. Our study found that all these safe attributes can be classified under each safe design principle, indicating IEC 61508 alignment. This result is logically reasonable since IEC 61508 provides goal-oriented requirements that allow freedom to implement the safe attributes.

A problem not discussed in the paper is that compliance with the IEC 61508 is insufficient for regulations and standards implementing prescriptive requirements. Typically, the regulations and standards only allow specific prescribed solutions even when other alternatives fulfil the same safe design principles. However, the findings on IEC 61508 alignment with the safe design principles open an opportunity to argue the safety at the principle level. This argumentation process would be irrespective of the safe attributes and can be an alternative approach when addressing the non-compliance between the technology and the prescriptive requirements.

6.1.3 Challenge of the safety demonstration process of novel technology

Since IEC 61508 is aligned with the safe design principle, any novel technology developed according to the IEC 61508 would be capable of fulfilling the safe design principles. Thus, the challenge instead is demonstrating that the software-intensive systems' complexity and novelty would still lead to safe behaviour. However, the unique safe attributes brought by the novel technology may have several practical implementations. For example, in safe design principle 2) strategy against failures and errors, functional separation in software that may be claimed when integrating the hardware shall be tested to check for conflicting requirements or uncontrolled interactions between the control and safety function. A safety demonstration process that can capture this interaction problem is required. Resolving all these issues with the implementation of the safe attributes may increase our assurance of the safety of the novel technology.

6.2 Contribution II – Solution-specific safety requirements

The second contribution is about topic II – solution-specific safety requirements. The detailed problem statement and scientific approaches are discussed in Gap II (see section 3.4), research question II in section 4.1.2, and article II [88] respectively. The contributions from the article are discussed in the following subsections. They also highlight how the contributions lead to achieving sub-objective II.

6.2.1 Identification of the most suitable methods for hazard analysis of novel and complex software-intensive solutions

Selecting the best hazard analysis method from a large number of options is difficult. Traditional methods are improving, while new methods keep being developed. The developments are caused mainly by new systems' characteristics, e.g., the complexity of the technology involving software-intensive systems. We investigated this issue systematically and limited the in-depth discussion into two methods: functional hazard analysis (FHA) and systems-theoretic process analysis (STPA). These methods' selection is due to the need to focus on the functional assessment when technical details are unavailable during the early development of novel technology and the methods' capability to capture the complexity characteristic.

The main contribution of our study is to recommend STPA as the hazard analysis method for novel technology involving software-intensive systems during the early design phase. Specifically, advice on utilizing STPA more efficiently has been provided, e.g., using the CESM model as a reference when modelling the system's hierarchical control structure. Hence, this result fulfils the sub-objective II. The detailed reasoning leading to this conclusion is explained more in the article, while brief descriptions are discussed in the following subsections.

STPA has been increasingly used for other industrial areas. The major precedent is the adoption of the method in the standard ISO/PAS 21448 [89] for the automotive industry. Furthermore, it is expected that this method will be adopted for the oil and gas industry shortly. In Norway, for example, research in STPA has been ongoing since 2012, one year after the method was proposed. Currently, it is at the stage of applied research, where various industry stakeholders have started to have a pilot study to see its applicability. However, additional investigations of other systems with different functionalities and complexities are required to verify our claim. This is because of the differing results with other research claims

where STPA is shown to be significantly better than the compared methods [90] or is required as supplementary methods [91].

6.2.2 Findings on the hazard analysis methods' capability for identifying functional hazards

As part of our study, we investigated FHA's and STPA's capability to identify functional hazards. From the results, FHA identified a slightly higher number of functional hazards than STPA. Several differences cause these. One of them is the different approach for classifying the function types used in the model and how each method treat the function during the hazard identification process. In FHA, they analyze every system function and apply the keywords to identify the functional hazards. Comparatively, in STPA, the feedback functions in a control loop are not investigated because STPA focuses on identifying unsafe control actions. Also, the STPA process requires assigning the function to an agent performing the function, while the FHA process does not. These reasons lead to small differences in the number of identified functional hazards.

6.2.3 Findings on the hazard analysis methods' capability to provide systemic perspective for the analysis

One of the main advantages of STPA is its support to a more comprehensive modelling technique that can capture all aspects of system complexity according to the CESM model called the hierarchical control structure model. In FHA, the modelling technique is not as comprehensive, with access to only the CES aspect of the CESM model. Thus, FHA needs to be supported with other documentation to have a complete systemic perspective.

As mentioned earlier, STPA seems to lack the number of identified functional hazards compared to FHA. However, STPA is compensated during the loss scenario identification process, which includes a more comprehensive aspect of the system during its analysis. These results in significant differences in the number of identified scenarios between STPA and FHA (346 vs 206 from our study case). From a practical perspective, the type of causal factors identified are within the scope of the CESM model, highlighting FHA's and STPA's capability to capture the systemic perspective of the system.

6.2.4 Identification of the produced solution-specific safety requirements' characteristics

Both FHA and STPA are capable of generating safety requirements that satisfy most of the criteria of a requirement discussed in section 3.3.1. The safety requirements generated by both methods follow a generic structure. For example, Figure 6.3 shows the safety requirements, or controller constraints in STPA terms, produced by the STPA process. The engineer can utilize these safety requirements for the treatment process, i.e., causal factors removal or prevention. It is important to refer to the produced requirements during the safety demonstration process to ensure the safety of the systems.

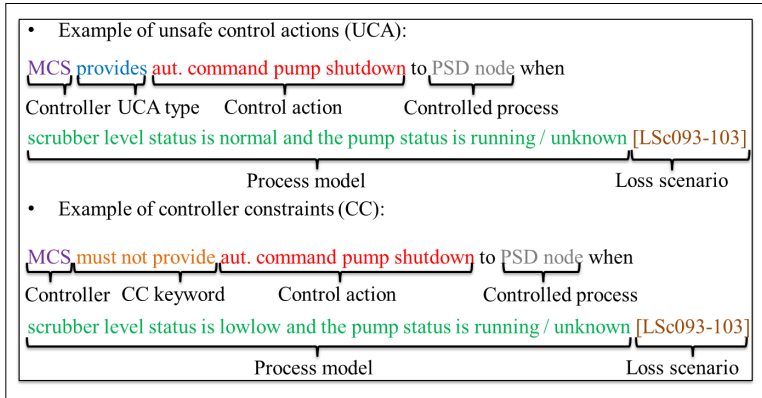


Figure 6.3: Example of solution-specific safety requirement (or controller constraint) produced by STPA.

6.2.5 Recommendation on the improvement of the hazard analysis methods from the lessons learned

Our study gave insights on recommendations to improve both FHA and STPA. For example, both methods have unique keywords that can be exchanged to increase the analysis coverage. In addition, specific to FHA, it is important to increase the capability of the modelling technique by covering all aspects of the CESP model to remove the risk of omission from the complicated process of referring to additional documentation mentioned before. Further details on the recommendation should be referred to the article.

6.3 Contribution III – Alternative concepts

The third contribution is about topic III – alternative concepts. The detailed problem statement and scientific approaches are discussed in Gap III (see section 3.4), research question III in section 4.1.3, and article III [92] respectively. In addition, research dataset I [93] has been published as supporting documents. The contributions from the article are discussed in the following subsections. They also highlight how the contributions lead to achieving sub-objectives I and II.

6.3.1 Proposal of the integration concept classification

We performed a systematic literature review on integrating every component in the control loop to achieve sub-objective I. We concluded that the integration type could be classified as follow (with example in Figure 6.4):

- *Complete independence*, see Figure 6.4a. The component separation limits the space, resources, functionality, and interaction exchanges between the control and safety systems.
- *Conditional independence*, see Figure 6.4b. The control and safety system may have limited communication on the components. Each control loop needs to ensure that no interruption can occur due to the communications to achieve conditional independence.

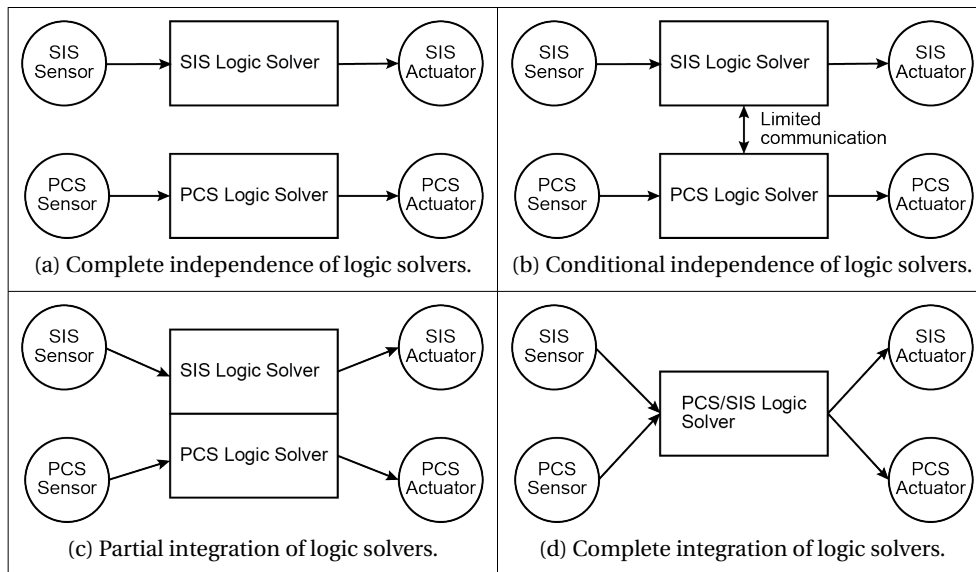


Figure 6.4: Generic architecture for control and safety logic solvers in for horizontal separation.

- *Partial integration*, see Figure 6.4c. The components are integrated partially, e.g., in space due to shared hardware. Separations are required logically to avoid unwanted interaction
- *Complete integration*, see Figure 6.4d. The component has full integration of hardware and software. No clear distinction between process control and safety elements to perform their respective functions.

This integration concept classification may assist the analyst during the investigation of alternative concepts for systems involving integration.

6.3.2 Proposal of hierarchical control structure modelling approach considering the integration

STPA is performed four times to a generic subsea gas compression system, considering different control and safety systems integration types. During the process, it is not easy to distinguish the resulting control structure model. Hence, the article proposes a new modelling approach to include the integration aspect in the hierarchical control structure. Please find the illustration in Figure 6.5. This model can assist the analyst in identifying hazards and loss scenarios during the subsequent process of STPA.

6.3.3 Challenges for implementing different types of integration concepts

From our study case, the integration does not necessarily produce new hazards. Nevertheless, it may result in new scenarios leading to these hazards, as shown in Figure 6.6. These results also validate the RQ II, where the differently identified scenarios can be developed

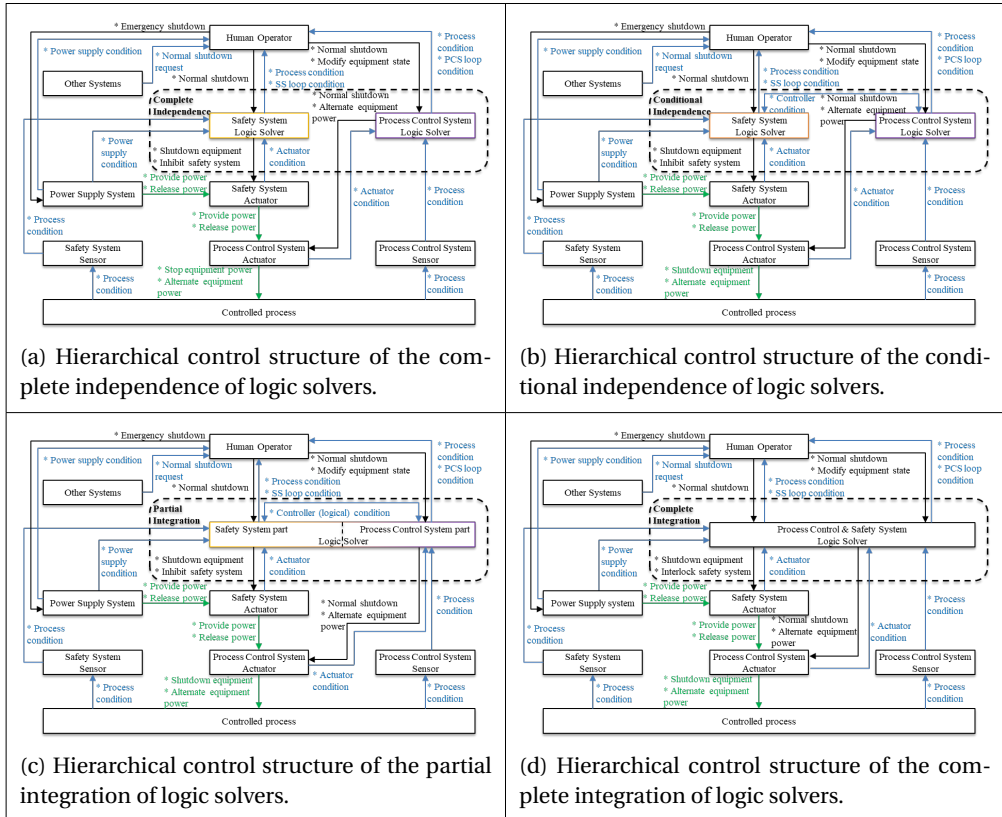


Figure 6.5: Proposed hierarchical control structure model considering different integration concept for STPA.

further as the safety requirements for the system when implementing a particular integration type.

One aspect that could be investigated further is whether these scenarios represent different risk magnitudes. Currently, the difference in number is trivial as arguments cannot be made. A higher number of scenarios with lower risk may be better than a lower number of scenarios with higher risk. It is crucial to understand how risk assessment can be performed as STPA does not include the process.

The discussions above highlight our contribution to fulfilling sub-objectives II. This contribution may help the stakeholders to have an efficient argument on selecting a particular integration type. Nevertheless, it would still be challenging to adopt a system beyond the conditional independence concept even if a solid argument has been presented. One of the main reasons is the fear of the unknown unknown that accompanies the complex system. Also, the oil and gas industry, specifically, has been scrutinized heavily in the past few years due to major safety and environmental accidents offshore. Therefore, it would be considered too risky to go beyond the known safe areas, resulting in a pragmatic approach for the designs and operations.

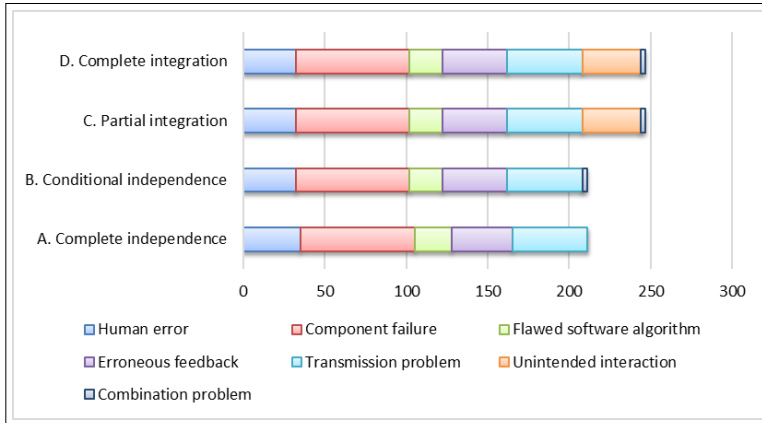


Figure 6.6: Comparisons of loss scenarios for system with different integration level.

6.4 Contribution IV – Effect on risk

The fourth contribution is about topic IV – effect on risk. The detailed problem statement and scientific approaches are discussed in Gap IV (see section 3.4), research question IV in section 4.1.4, and article IV [94] respectively. Research dataset II [95] has been published as supporting documents. The contributions from the article are discussed in the following subsections. They also highlight how the contributions lead to achieving sub-objective II.

6.4.1 Improvement of a modelling approach for STPA's loss scenarios based on finite-state automata modelling type

The model is called STPA-FSA (finite state automata) and is an improvement from early work by Zhang et al. [83]. We propose several modelling patterns that can standardize the modelling approach. The modelling patterns help the user on reducing the modelling uncertainty. The modelling patterns are based on combining a generic failure model for the safety system (see Figure 6.7a) and the hazardous event model based on the scenarios from STPA (see Figure 6.7b) as FSA models. The combination is theoretically supported by Jin et al. [96]. The model is simulated through Monte-Carlo simulations to obtain the frequency of loss scenarios. Validations of the model have been performed by comparing it with the Markov model.

It is expected that the contribution helps to characterize quantifiable aspects of the scenarios better. Therefore, this contribution fulfils the sub-objective II on the suitable methods to support the safety demonstration process.

6.4.2 Discussion on the model capability to address dependency

In the STPA-FSA model, the dependency captured by STPA is represented through logical operators (e.g., AND, OR, IMPLY, or NOT). This would allow the modelling of combinations of dependent causal factors that may present in the loss scenarios. This approach ensures that the benefits of STPA for identifying complex loss scenarios are not lost during the quantification process.

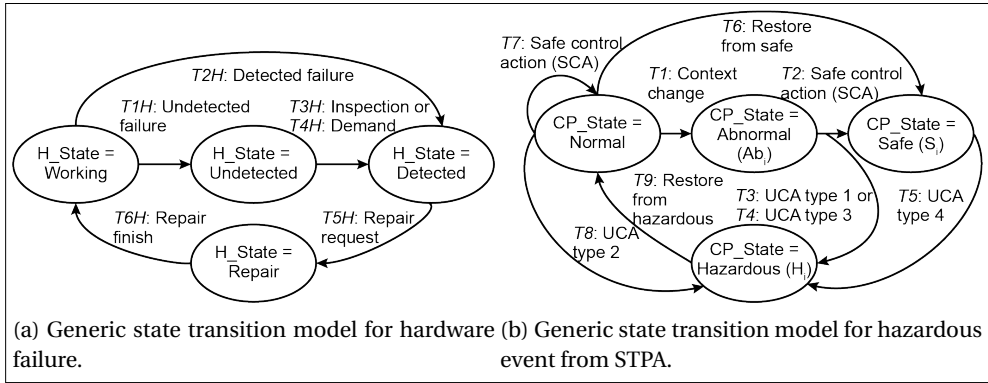


Figure 6.7: Example of generic model in STPA-FSA.

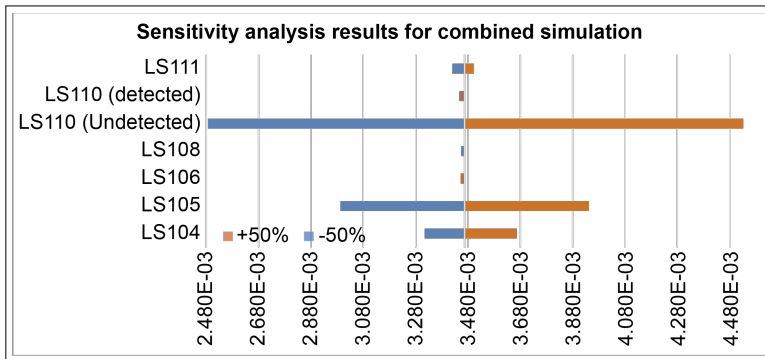


Figure 6.8: Example of sensitivity analysis results.

6.4.3 Model's capability to identify unnecessary requirements

STPA-FSA framework could be improved further. Simulations managed to reveal unnecessary requirements represented by zero frequency. This finding indicates that the STPA's loss scenario analysis process is disjointed. It should be noted that improvement of the STPA process is necessary to increase the efficiency of analysis.

6.4.4 Identification of the proposed model limitations when compared to the available modelling approaches

One main challenge on the quantification is the unavailability of field data, especially for novel technology. A technology qualification program would be necessary to gather more information. This challenge, however, is common to all other quantification processes. The remaining uncertainties should be managed to improve our confidence in the results. For example, sensitivity analyses (see result example in Figure 6.8) can be performed to identify the magnitude of deviations from these uncertainties. Another disadvantage of STPA-FSA is that non-quantifiable scenarios not captured by the process will require different treatment methods beyond the scope of our article, e.g., by following the IEC 61508 systematic approach.

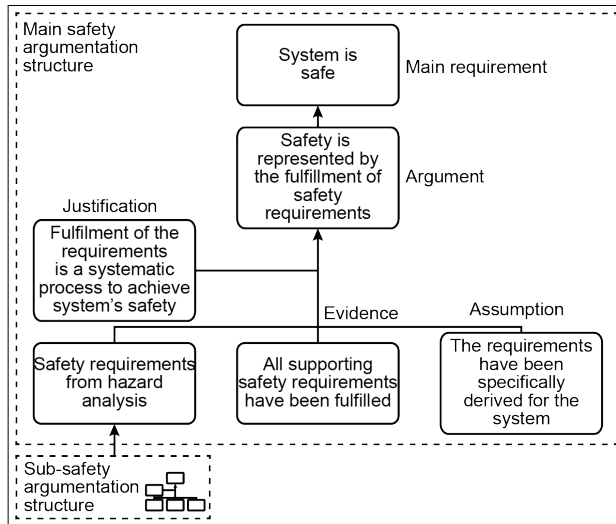


Figure 6.9: Example of safety argument.

6.5 Contribution V – Safety argumentation

The fifth contribution is about topic V – safety argumentation. The detailed problem statement are discussed in Gap V (see section 3.4) and research question V in section 4.1.5 respectively. The contributions are obtained through a systematic literature review and discussion based on logical reasoning. They are aimed to answer sub-objective III and are described in the following subsections.

6.5.1 Clarification for the safety argumentation concept

The ultimate goal of a safety demonstration is to prove that a system is safe. As discussed earlier in section 3.3, the safety argument concept could support the safety demonstration. Figure 6.2 shows the updated argumentation concept based on the PhD project’s result. Figure 6.9 shows an example implementation of the concept through the use of safety requirements. The argumentation shall have a hierarchical structure, with multiple sub-safety argumentation structures to support the main requirement. When the technology is novel with limited information, the sub-structures are developed for every requirement obtained from the hazard analysis process. Comparatively, when the technology is mature, the standardized prescriptive requirements may be used as an alternative. This requirement can be proven by presenting evidence of compliance, e.g., the safety demonstration report.

STPA is the recommended hazard analysis method for identifying complex scenarios that may be present in the UC 2 or the technology involving software-intensive systems in general. The link between the proposed method and the relevant safety lifecycle phases is presented in Figure 6.10. It is shown that the proposed methods are mostly STPA-based methods. The reason is due to the results of contribution II, which affect the development direction of the safety demonstration framework. Nevertheless, this recommendation does not make the other hazard analysis methods obsolete. On the contrary, the intention is to fit the other methods depending on the need for the novel technology. For example, if the novel technology changes how the process works, HAZOP may be the most suitable method to de-

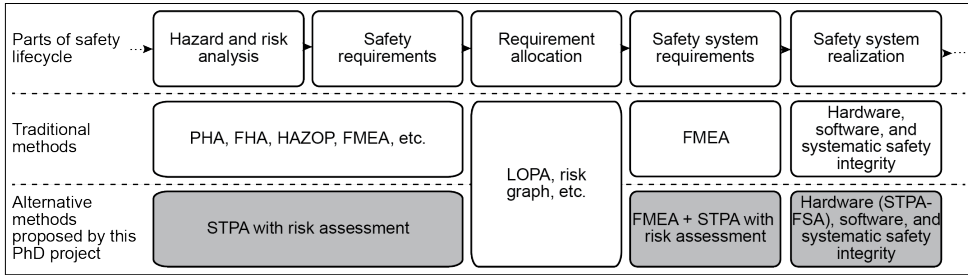


Figure 6.10: Linking the proposed methods with the safety lifecycle.

rive requirements. This consideration should always be made when developing the plan for a safety demonstration.

The solution-specific safety requirements obtained from the hazard analysis process should be categorized based on the safe design principles. This categorization ensures that the new or different design solutions that affect one or more principles could be compared with the current technical solutions. For example, the integration concept in UC 2 affects the safety vs complexity and segregation principles. As discussed earlier, integration is not compliant with the prescriptive requirements from the local regulations. However, since the international standards may allow integration, it is the job of the developer to demonstrate that the integrated system could still fulfil the segregation principle through alternative methods, e.g., logical separation in software. Successful safety demonstration of the integration concept may increase the regulator’s confidence with the technology the prescriptive requirements may be updated.

The type of evidence for a complex system would still follow the classification by [68]. However, the difference is in how the evidence is generated. For example, STPA-FSA has been developed as an alternative modelling approach to proving the hardware safety integrity during the realization of the safety system phase. STPA-FSA could supplement the current modelling approach such as the Boolean and finite state automata models. Though, similar to hazard analysis, selecting a modelling approach should still depend on the system’s characteristics to avoid overcomplicating the process.

Assumption management is vital since it set the boundary of the evidence relevance with the requirement. When parts of the technology are revised, the generated evidence may become irrelevant, and re-assessment is necessary to maintain the system’s safety. It is important since the experience, e.g., with the Bhopal disaster [97], shows the catastrophic consequence of the engineer’s negligence on the safety management process. This can be forewarned, for example, by performing sensitivity analysis to understand the magnitude of deviation in the assumptions. Hence, prioritization could be made when performing the re-assessment of the safety arguments.

6.5.2 Relevance assessment for the available methods in the framework

STPA, in specific, is still lacking additional processes to fit with the safety argumentation concept. STPA focuses on having a comprehensive perspective of the systems. However, it does not have any risk assessment process yet, making objective comparison and assessment more difficult. Research [60, 61] has been done to include the risk assessment process for STPA. The results are promising, indicating that STPA with risk assessment is now more comparable to the traditional hazard and risk analysis methods and can be recommended as

part of the safety lifecycle process. Due to this reason, it would be recommended to include the risk assessment process of STPA to increase its suitability to the safety lifecycle.

6.5.3 Remaining aspects of safety argumentation concept

Due to limited information on the technological implementation for the integration concept, it was not possible to conduct in-depth research on the software and systematic safety integrity aspects in this PhD project. For a complex system, both software and systematic safety integrity would be more critical since complexity affects how interactions occur in the system.

This PhD research did not go deeper into the investigation of the argument and the justification of the argument. It is because arguments could only be made according to the requirement and the available evidence. Since the research focuses on high-level requirements without any information on the implemented system technological solutions, precise details are unavailable. However, it is still valid to conclude that including new methods would still align with the safety life cycle since they are developed with the process in mind. Hence, the requirements, evidence and assumptions could still be linked through logical deduction and reasoning.

Chapter 7

Summary and Recommendations for Further Work

This chapter starts by summarizing the PhD project contributions and the conclusion of this PhD thesis in section 7.1. Further considerations are presented for the future directions from this research in section 7.2.

7.1 Summary and Conclusion

This PhD has carried out research, which recommends methods and approaches to support the safety demonstration process. This has been done by focusing on the novel technology implementing integration between the control and safety systems within the scope of functional safety. This PhD project is part of the Safety 4.0 project and is associated with SUBPRO, allowing close collaborations with the industry. This collaboration ensures high project's quality and significance. The research direction is established based on the industry's standards and best practices while focusing on the remaining industrial challenges, especially safety demonstrations.

We first clarified the safe design principles used to derive the industrial requirements for technology. These principles are checked against IEC 61508 as the recommended standard used in the industry. Then, we identified STPA as the recommended method for hazard analysis of complex systems. This conclusion is based on various factors, including discussing the method's approach, modelling coverage, and analysis capability. These factors are essential to ensure that the complex behaviour of the system can be captured and managed during the risk management process. Afterwards, we developed a quantitative modelling approach to assess the risk of the hazards. These are useful for assessing the importance of every scenario in more detail and allows an objective decision-making process. We finally compile all of the results as safety argumentation concepts and clarify how they fit within the framework for a safety demonstration.

Case studies have been performed on a case by case basis to exemplify how the developed methods could be used, mainly our study case, UC 2 integration of process control and safety. The integration is applied to a subsea processing system, specifically a subsea gas compression system that we developed based on discussion with industry experts and information from international standards. The results indicate the practicality of the proposed methods.

The overall implications of this PhD project results for the industry are that the engineers can utilize the recommended methods and approaches for performing safety demonstration

of novel technology. This is based on the safety argumentation concept with the support of the new STPA-based methods. However, this can still be a challenge for the industry partners since the methods are significantly different from the traditional methods currently used in practice. Thus, implementation of the proposed methods would require close accompaniment by experts to achieve the intended results. Furthermore, this process may require additional investments in personnel and time. As a result, the different practical application aspect of the proposed method, i.e., work process efficiency, is still in question. Therefore, further works are necessary to complement the framework.

7.2 Recommendation for Further Works

This thesis is complete only within the scope of the PhD project. However, the project has a limited duration, and the research topic in safety demonstration is vast and distinctive. Therefore, claiming that the work is finished is impossible. Instead, in this section, I guide the future work on the framework proposed in this PhD thesis. More detailed discussion on future works for each contribution has been discussed in the attached articles separately.

7.2.1 Generic application of the framework

The framework for the safety demonstration process is developed with the use cases from the subsea oil and gas industry. The key features of the novel technology are novelty and complexity. These features are common to many systems across the industries. Considering that the concept of functional safety in IEC 61508 is generic, it should be possible to use the framework in another industry. However, we had spent a significant effort focusing on the specific safe design principle of the use cases, i.e., integration. Theoretically, the proposed framework could still work for other safe design principles. However, this has not been proven. Therefore, for future work, a pilot study on other systems with different characteristics would be useful to verify and validate the proposed frameworks.

7.2.2 Uncertainty management of the results

All hazard analysis methods, including STPA, are performed based on the current contextual assumption. Hence, there is a risk for unknown unknowns due to uncertainty in the assumptions and the system's modification. Sensitivity analysis has been performed whenever possible to manage the uncertainty. Nevertheless, it is not exhaustive since it is not performed systematically. Therefore, there is a need to investigate a method for managing uncertainty in the hazard analysis results, e.g., identifying the risk of deviations in the assumptions.

7.2.3 Management of software and systematic safety integrity

As discussed earlier in Section 6.5.3, the software and systematic safety integrity aspects for the complex systems have not been covered during this PhD. The reasons are the limited technical details and the starting point of the research, which is from a high-level perspective. I have mentioned that complexity would lead to more software and systematic failures as causal factors. Hence, it would be necessary to clarify how it can be included in the safety demonstration process. Furthermore, it would be necessary to have the actual technology, not just a concept, to propose a tailored method for verifying the safety of the complex system.

7.2.4 Testing of the safety argumentation concept

Similar to the previous subsection, the arguments and the justification for the safety argumentation concept has also not been covered during this PhD. However, this is because of the different focus of this PhD, which focuses on developing the methods and proposal of the framework, not on the formulation of safety arguments. It is known that a safety argument typically ends in a large safety argumentation model with interdependencies in the evidence, assumptions, and sub-structures. When the systems become too large, the network in the structure becomes complex, leading to difficulty in checking the consistency of the structure. A pilot study is required to identify the new problems caused by this complex structure.

Bibliography

- [1] Rystad Energy. Technologies to improve NCS competitiveness. Report commissioned by og21, 2019.
- [2] Carsten Mahler, Markus Glaser, Simon Schoch, Stefan Marx, Stefan Schluenss, Tobias Winter, Julian Popp, and Sebastian Imle. Safety capability of an all-electric production system. In *Offshore Technology Conference*. OnePetro, 2019.
- [3] CCPS. *Guidelines for safe automation of chemical processes*. 2017.
- [4] Andreas Hafver, Kenneth Kvinnesland, Meine J. P. van der Meulen, Odd Ivar Haugen, Simen Eldevik, Tore Myhrvold, Frank Børre Pedersen, and Mary Ann Lundteigen. Gaps, challenges, opportunities, and needs. Safety 4.0 report, DNV, 2018.
- [5] DNV. Safety 4.0, n.d. Retrieved 2021-03-10. URL: <https://www.dnv.com/research/oil-gas/safety40/index.html>.
- [6] IEC 60050-192. International Electrotechnical Vocabulary (IEV) part 192: Dependability. Standard, International Electrotechnical Commission, 2015.
- [7] IEC 61508. Functional safety of electrical/electronic/programmable electronic safety-related systems – part 1-7. Standard, International Electrotechnical Commission, 2010.
- [8] IEEE 1471. IEEE recommended practice for architectural description for software-intensive systems. Standard, Institute of Electrical and Electronics Engineers, 2000.
- [9] ISO/IEC Guide 51. Safety aspects. Standard, International Electrotechnical Commission, 2014.
- [10] ISO/TR 12489. Petroleum, petrochemical and natural gas industries – Reliability modelling and calculation of safety systems. Standard, International Organization for Standardization, 2013.
- [11] ISO 9000. Quality management systems – Fundamentals and vocabulary. Standard, International Organization for Standardization, 2015.
- [12] OG21. *Strategy document*. Oil and Gas for the 21st Century, 2016.
- [13] OG21. Technologies for cost and energy efficiency. Report, Oil and Gas for the 21st Century, 2019.
- [14] T. Ruud, A. Idrac, L. J. McKenzie, and S. H. Høy. All subsea: A vision for the future of subsea processing. OTC Offshore Technology Conference, May 2015. doi: 10.4043/25735-MS.

- [15] Ole Økland and Rune Mode Ramberg. Subsea factory—standardization of the brownfield factory. OTC Offshore Technology Conference, May 2015. doi: 10.4043/25903-MS.
- [16] API RP 17A. Design and operation of subsea production systems-general requirements and recommendations. Recommended practice, American Petroleum Institute, 2017.
- [17] Yong Bai and Qiang Bai. *Subsea engineering handbook*. Gulf Professional Publishing, 2012.
- [18] Hyungju Kim, Mary Ann Lundteigen, and Christian Holden. A gap analysis for subsea control and safety philosophies on the norwegian continental shelf. In *13th International Conference on Probabilistic Safety Assessment and Management (PSAM 13)*, Seoul, Korea, 2016.
- [19] Alexander Haro. A lightning strike caused the “Eye of Fire” in the Gulf of Mexico, 2021. Newspaper article, retrieved 2021-09-15. URL: <https://www.theinertia.com/environment/a-lightning-strike-caused-the-eye-of-fire-in-the-gulf-of-mexico/>.
- [20] API RP 17V. Recommended practice for analysis, design, installation, and testing of safety systems for subsea applications. Standard, American Petroleum Institute, 2015.
- [21] Eric P. Steinhauser. Addressing the challenges of implementing safety instrumented systems in multi-product batch processes. *Journal of Loss Prevention in the Process Industries*, 57:164 – 173, 2019. ISSN 0950-4230. doi: 10.1016/j.jlp.2018.11.019.
- [22] NORSOK S-001. Technical safety. Standard, NORSOK, 2008.
- [23] Paul Gruhn and Harry Cheddie. *Safety instrumented systems: design, analysis, and justification*. The Instrumentation, Systems, and Automation Society, 2006.
- [24] Petroleum Safety Authority Norway. *Regulations relating to health, safety and the environment in the petroleum activities and at certain onshore facilities (the Framework regulations)*. 2019.
- [25] Petroleum Safety Authority Norway. *Regulations relating to conducting petroleum activities (the Activities regulations)*. 2019.
- [26] Petroleum Safety Authority Norway. *Regulations relating to design and outfitting of facilities, etc. in the petroleum activities (the Facilities regulations)*. 2019.
- [27] Petroleum Safety Authority Norway. *Regulations relating to management and the duty to provide information in the petroleum activities and at certain onshore facilities (the Management regulations)*. 2019.
- [28] NORSOK P-002. Process system design. Standard, NORSOK, 2014.
- [29] NORSOK I-001. Field instrumentation. Standard, NORSOK, 2019.
- [30] NORSOK I-002. Safety and automation system (SAS). Standard, NORSOK, 2001.
- [31] ISO 13628. Petroleum and natural gas industries - Design and operation of subsea production systems. Standard, International Organization for Standardization, 2005.

- [32] NORSOK Z-013. Risk and emergency preparedness assessment. Standard, NORSOK, 2008.
- [33] ISO 31000. Risk management. Standard, International Organization for Standardization, 2018.
- [34] API RP 17N. Subsea production system reliability, technical risk, and integrity management. Recommended practice, American Petroleum Institute, 2017.
- [35] DNVGL-RP-A203. Technology qualification. Recommended practice, DNV, 2013.
- [36] NOG 070. Guidelines for the application of IEC 61508 and IEC 61511 in the petroleum activities on the continental shelf (Recommended SIL requirements). Guideline, Norwegian Oil & Gas, 2018.
- [37] IEC 61511. Functional safety – Safety instrumented systems for the process industry sector. Standard, International Electrotechnical Commission, 2010.
- [38] ISO 14224. Petroleum, petrochemical and natural gas industries – Collection and exchange of reliability and maintenance data for equipment. Standard, International Organization for Standardization, 2016.
- [39] Department of Defense. Technology readiness assessment (TRA) deskbook. Technical report, Director, research directorate (DRD) and office of the director, defense research, and engineering (DDR&E), 2009.
- [40] Cambridge Dictionary. Online dictionary, n.d. Retrieved 2021–08–01. URL: <https://dictionary.cambridge.org/dictionary/english/>.
- [41] EN 50126. Railway applications – The specification and demonstration of reliability, availability, maintainability and safety (RAMS). Standard, European Committee for Electrotechnical Standardization (CENELEC), 2017.
- [42] IEC 61513. Nuclear power plants – Instrumentation and control important to safety – General requirements for systems. Standard, International Electrotechnical Commission, 2011.
- [43] ISO 26262. Road vehicles – Functional safety. Standard, International Organization for Standardization, 2018.
- [44] David Liddle. Subsea technology: A reflection on global challenges and solutions. OTC Offshore Technology Conference, April 2012. doi: 10.4043/23092-MS.
- [45] S. Furtado, H. Hamedifar, K. Mateen, L. Huyse, L. M. Rivero, and G. Kusinski. Development of Novel Integrity Assurance Approach for Technology Qualification of New Subsea Technologies by Deepstar®. OTC Offshore Technology Conference, May 2016. doi: 10.4043/27296-MS.
- [46] Nancy G. Leveson and Kathryn Anne Weiss. Software system safety. In *Safety Design for Space Systems*, pages 475–505. Elsevier, 2009.
- [47] Inger Lise Johansen and Marvin Rausand. Defining complexity for risk assessment of sociotechnical systems: A conceptual framework. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 228(3):272–290, 2014.

- [48] Ludwig von Bertalanffy and Edgar von Taschdjian. *Perspectives on general system theory*. 1976.
- [49] Andreas Hafver, Carla Ferreira, Christian Agrell, Dag McGeorge, Erik Andreas Hektor, Frank Børre Pedersen, Meine van der Meulen, Odd Ivar Haugen, Simen Eldevik, and Tore Myhrvold. On the meaning of assurance. In *Proceedings of the 31st European Safety and Reliability Conference (ESREL)*, pages 3288–3295. Research Publishing, 2021. ISBN 978-981-18-2016-8.
- [50] Luciano Floridi. The method of levels of abstraction. *Minds and machines*, 18(3):303–329, 2008.
- [51] Mario Bunge. *Emergence and convergence*. University of Toronto Press, 2016.
- [52] Odd Ivar Haugen. Framework for safety analysis of complex systems. In *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference (ESREL 2020 PSAM15)*, pages 2095–2102. Research Publishing Services, 2020. ISBN 978-981-14-8593-0.
- [53] Stanley Kaplan and B. John Garrick. On the quantitative definition of risk. *Risk analysis*, 1(1):11–27, 1981.
- [54] Terje Aven and Genserik Reniers. How to define and interpret a probability in a risk and safety setting. *Safety science*, 51(1):223–231, 2013.
- [55] Anders Jensen and Terje Aven. A new definition of complexity in a risk analysis setting. *Reliability Engineering & System Safety*, 171:169–173, 2018.
- [56] Marvin Rausand and Stein Haugen. *Risk assessment: Theory, methods, and applications*. John Wiley & Sons, 2020. ISBN 978-1-119-37728-3.
- [57] Clifton A. Ericson. *Hazard analysis techniques for system safety*. Wiley, Hoboken, New Jersey, 2nd edition, 2016. ISBN 1-119-10170-0.
- [58] Nancy G. Leveson. *Engineering a safer world: Systems thinking applied to safety*. MIT press, 2011.
- [59] Erik Hollnagel. *FRAM: The functional resonance analysis method: Modelling complex socio-technical systems*. Ashgate, 2012.
- [60] Hyungju Kim, Mary Ann Lundteigen, Andreas Hafver, and Frank Børre Pedersen. Utilization of risk priority number to systems-theoretic process analysis: A practical solution to manage a large number of unsafe control actions and loss scenarios. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 235: 92–107, 2021.
- [61] N. A. Zikrullah. *Prioritization approach for Systems-Theoretic Process Analysis (PA-STPA): Applied for subsea systems*, 2018.
- [62] NUREG 1855. *Guidance on the treatment of uncertainties associated with PRAs in risk-informed decision making*. Standard, United States Nuclear Regulatory Commission, 2009.

- [63] Snorre Sklet. Safety barriers: Definition, classification, and performance. *Journal of Loss Prevention in the Process Industries*, 19(5):494 – 506, 2006. ISSN 0950-4230. doi: 10.1016/j.jlp.2005.12.004.
- [64] Urban Kjellén. *Prevention of accidents through experience feedback*. CRC Press, 2000.
- [65] Tim Kelly and Rob Weaver. The goal structuring notation – A safety argument notation. In *Proceedings of the dependable systems and networks 2004 workshop on assurance cases*, page 6. Citeseer, 2004.
- [66] Jon Holt, Simon A. Perry, and Mike Brownsword. *Model-Based Requirements Engineering*. The Institution of Engineering and Technology, 2011. ISBN 978-1-849-19487-7.
- [67] Odd Ivar Haugen. Properties of evidence and evidence generation. In *Dynamic positioning conference*. Marine Technology Society, 2018.
- [68] S. Nair, J. L. de la Vara, M. Sabetzadeh, and L. Briand. An extended systematic literature review on provision of evidence for safety certification. *Information and Software Technology*, 56(7):689–717, 2014. ISSN 0950-5849. doi: 10.1016/j.infsof.2014.03.001.
- [69] Rajwinder Kaur Panesar-Walawege, Mehrdad Sabetzadeh, Lionel Briand, and Thierry Coq. Characterizing the chain of evidence for software safety cases: A conceptual model based on the iec 61508 standard. In *2010 Third International Conference on Software Testing, Verification and Validation*, pages 335–344. IEEE, 2010.
- [70] Sunil Nair, Jose Luis de la Vara, Mehrdad Sabetzadeh, and Davide Falessi. Evidence management for compliance of critical systems with safety standards: A survey on the state of practice. *Information and Software Technology*, 60:1–15, 2015.
- [71] Timothy Patrick Kelly. *Arguing safety: A systematic approach to managing safety cases*. PhD thesis, University of York York, UK, 1999.
- [72] F. Drogoul, S. Kinnersly, A. Roelen, and B. Kirwan. Safety in design – Can one industry learn from another? *Safety Science*, 45(1):129–153, 2007. doi: 10.1016/j.ssci.2006.08.004.
- [73] Mary Ann Lundteigen, Marvin Rausand, and Ingrid Bouwer Utne. Integrating RAMS engineering and management with the safety life cycle of IEC 61508. *Reliability Engineering & System Safety*, 94(12):1894–1903, 2009.
- [74] Mehrdad Sabetzadeh, Davide Falessi, Lionel Briand, and Stefano Di Alesio. A goal-based approach for qualification of new technologies: Foundations, tool support, and industrial validation. *Reliability Engineering & System Safety*, 119:52–66, 2013.
- [75] Victor Bolbot, Gerasimos Theotokatos, Luminita Manuela Bujorianu, Evangelos Boulougouris, and Dracos Vassalos. Vulnerabilities and safety assurance methods in Cyber-Physical Systems: A comprehensive review. *Reliability Engineering & System Safety*, 182:179–193, 2019.
- [76] Ian F. Alexander and Neil Maiden. *Scenarios, stories, use cases: Through the systems development life-cycle*. John Wiley & Sons, 2005.
- [77] SAE 3016. Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Standards, Society of Automotive Engineers, 2014.

- [78] NI641 R01. Guidelines for autonomous shipping. Guidelines, Bureau Veritas, 2019.
- [79] Antoine Rauzy. Notes on computational uncertainties in probabilistic risk/safety assessment. *Entropy*, 20(3):162, 2018.
- [80] Tatiana Prosvirnova. *AltaRica 3.0: A model-based approach for safety analyses*. PhD thesis, École Polytechnique, 2014.
- [81] Steven F Railsback and Volker Grimm. *Agent-based and individual-based modeling: A practical introduction*. Princeton university press, 2019.
- [82] Ondřej Nývlt, Stein Haugen, and Lukáš Ferkl. Complex accident scenarios modelled and analysed by Stochastic Petri Nets. *Reliability Engineering & System Safety*, 142:539–555, 2015.
- [83] Juntao Zhang, Hyungju Kim, Yiliu Liu, and Mary Ann Lundteigen. Combining system-theoretic process analysis and availability assessment: A subsea case study. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 233(4): 520–536, 2019.
- [84] Daniele Fanelli. Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90(3):891–904, 2012.
- [85] OECD. *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*. 2015. doi: 10.1787/9789264239012-en.
- [86] John W. Creswell and J. David Creswell. *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications, 2018.
- [87] Nanda A. Zikrullah, Meine J. P. van der Meulen, Hyungju Kim, and Mary A. Lundteigen. Clarifying implementation of safe design principles in IEC 61508: Challenges of novel subsea technology development. In *Proceedings of the 29th European Safety and Reliability Conference (ESREL)*, pages 2928–2936. Research Publishing, 2019. ISBN 978-9-811-12724-3.
- [88] Nanda A. Zikrullah, Hyungju Kim, Meine J. P. van der Meulen, Gunleiv Skofteland, and Mary Ann Lundteigen. A comparison of hazard analysis methods capability for safety requirements generation. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 235:1132–1153, 2021.
- [89] ISO/PAS 21448. Road vehicles–Safety of the intended functionality. Standard, International Organization for Standardization, 2019.
- [90] Nancy G. Leveson, Chris Wilkinson, Cody Fleming, John Thomas, and Ian Tracy. A comparison of STPA and the ARP 4761 safety assessment process. Technical report, MIT PSAS, 2014.
- [91] Sardar Muhammad Sulaman, Armin Beer, Michael Felderer, and Martin Höst. Comparison of the FMEA and STPA safety analysis methods—A case study. 27(1):349–387, 2019. ISSN 0963-9314.

- [92] Nanda A. Zikrullah, Meine J. P. van der Meulen, and Mary A. Lundteigen. A comparison of hazardous scenarios in architectures with different integration types. In *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference (ESREL 2020 PSAM15)*, pages 4001–4008. Research Publishing Services, 2020. ISBN 978-981-14-8593-0.
- [93] Nanda Anugrah Zikrullah, Meine J. P. van der Meulen, Gunleiv Skofteland, and Mary Ann Lundteigen. Systems-theoretic process analysis results for system with different integration types, 2021. Dataset V1.
- [94] Nanda A. Zikrullah, Meine J. P. van der Meulen, and Mary Ann Lundteigen. Finite-state automata modeling pattern of systems-theoretic process analysis results. *Reliability Engineering and System Safety*, page Under review, n.d..
- [95] Nanda Anugrah Zikrullah, Meine J. P. van der Meulen, and Mary Ann Lundteigen. Systems-theoretic process analysis – Finite state automata (STPA-FSA) modeling approach source code, n.d.. Dataset V1.
- [96] Hui Jin, Mary Ann Lundteigen, and Marvin Rausand. Reliability performance of safety instrumented systems: A common approach for both low-and high-demand mode of operation. *Reliability Engineering & System Safety*, 96(3):365–373, 2011.
- [97] Edward Broughton. The Bhopal disaster and its aftermath: A review. *Environmental Health*, 4(1):1–6, 2005.

Part II

Articles

Article I

N. A. Zikrullah, H. Kim, M. J. P. van der Meulen, M. A. Lundteigen, Clarifying implementation of safe design principles in IEC61508: Challenges of novel subsea technology development, in: Proceedings of the 29th European Safety and Reliability Conference (ESREL), Research Publishing, 2019, pp. 2928–2936.

Clarifying Implementation of Safe Design Principles in IEC 61508: Challenges of Novel Subsea Technology Development

N.A. Zikrullah, H. Kim, M.A. Lundteigen

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Norway. E-mail: nanda.a.zikrullah@ntnu.no

M.J.P. van der Meulen

Group Technology and Research, DNV GL, Norway. E-mail: meine.van.der.meulen@dnvgl.com

When new technologies are introduced to safety systems, they may raise some new discussions and clarifications about established practices. IEC 61508 represents a general framework which may apply to all electrical/electronic/programmable electronic (E/E/PE) novel technologies aiming for safety-critical applications. At the same time, it is important to handle issues of inconsistency between the new concepts and sector-specific standards and guidelines that give more specific requirements to how the technical solutions shall be realized. An important starting point can be to clarify the governing principles of safe design philosophies, beyond the detailed clauses, in order to guide the discussion when new technologies require new design philosophies (e.g. on how to achieve the fail-safe function). When technical solutions are novel, it is also important to acknowledge the importance of a safe design process for building confidence to the solution. So, it can be of interest to discuss the role of the safe design process to reduce uncertainty associated with the performance of the new technical solution (e.g. battery instead of spring). This paper is intended to contribute to the foundation theory for safe design of novel subsea technology in the Safety 4.0 project, a research-based joint industry project which aims to develop a safety demonstration framework for the novel subsea technology.

Keywords: IEC 61508, safe design, design principle, safety philosophy, novel technology, subsea technology.

1. Introduction

Many industries rely on the adoption of new technologies to improve performance of existing systems, reduce costs, or accommodate new ways of operating. At the same time, these technologies represent a risk, both financially and potentially from a safety point of view, as there is no or limited experience about the performance. A technology qualification programme (TQP, e.g. in DNVGL-RP-A203 (2017) and API RP 17N (2017)) is therefore normally initiated to ensure that the risk is reduced to an acceptable level.

In the qualification of new technologies, it is important not to consider the system in isolation, but also the interaction with other systems. For example, a new safety system may need to interact adequately with other risk reducing measures. When technologies involved rely on electrical/electronic/programmable electronic (E/E/PE) components, the target is to ensure functional safety of the system. The term functional safety has been introduced in standards on E/E/PE safety-related systems, like IEC 61508 (2010) and its sector-specific standards, IEC 61511 (2016) for process industry and IEC 61513 (2011) for nuclear power plants. Ensuring

functional safety becomes of paramount importance before putting the technology into operation and must be an integral part of the TQP.

When introducing new technology, trade-offs need to be made between the TQP, generic standards and sector-specific standards before qualifying the new technology for operation. However, a dilemma arises when the new technology clearly uses different technical solutions than the one specified in the sector-specific standards (e.g. battery instead of spring for fail-safe function of a shutdown valve). It is important to handle issues of inconsistency between the new concept and the affected standards and guidelines to ensure compliance of the new technology.

There are two key questions to be asked during a process of demonstrating functional safety: 1) What are the functional safety and safety integrity requirements in light of the hazards and risks that the systems need to control? 2) What approach(es) are suitable for providing sufficient evidence of compliance (safety demonstration)? IEC 61508 specifies goal-oriented requirements where the users shall define their own acceptance criteria and approach to demonstrate the safety integrity of the system. It does not prescribe

Proceedings of the 29th European Safety and Reliability Conference.

Edited by Michael Beer and Enrico Zio

Copyright © 2019 European Safety and Reliability Association.

Published by Research Publishing, Singapore.

ISBN: 978-981-11-2724-3; doi:10.3850/978-981-11-2724-3_0112-cd

2928

specific technical solutions for compliance but only mentions the required safe attributes of the system. While these goal-oriented requirements allow creativity for new technology design, it affects the required time and resources for the safety demonstration process if we do not have relevant tools at hand.

Attempts to handle both the issues of inconsistency and compliance have been ongoing research for quite some years. Some researchers have developed approaches for qualification of new technologies and standards focusing more on the process (Lundteigen et al. 2009, Sabetzadeh et al. 2013), while some others focus more on the safety demonstration aspect of new technology (Nair et al. 2015, Bolbot et al. 2019).

In this paper, we want to start from a different perspective by clarifying the principle of safe design, beyond the detailed clauses in the standards, in order to guide the discussion when new technologies require new design philosophies (e.g. on how to achieve a fail-safe solution). Several researchers have proposed their own classification for the safe design principle. Most of them define the terms differently and end up with definitions that actually explain either (one or a combination of) the aim, context, principle, and/or attribute of safe design (Hussey and Atchison 2000, Sharma 2017, van de Poel and Robaey 2017). It is only Drogoul et al. (2007), to the best of our knowledge, that started by defining the meaning of each aspect clearly and classify the findings associated with them. The paper has identified seven high-level safe design principles from five industry domains: railway, air traffic management, aircraft, process industries and automobile. These issues are elaborated further in the latter section.

Although, the principles may be regarded as general, it is interesting to see how they are adopted by IEC 61508. The main objective of this paper is: (1) to analyse the implementation of safe design principles by Drogoul et al. (2007) in IEC 61508 and (2) to identify the practical implication to the design of novel subsea technology as a study case.

The paper scope is limited to the design phase of system according to IEC 61508. However, functional safety also includes aspects of system development not covered in this paper, including risk analysis, implementation, operation and maintenance, and functional safety management.

This paper has been developed as contribution to the foundation theory in the Safety 4.0 project. Safety 4.0 is a research-based joint-industry project that aims to develop a framework (work processes, methods and tools) for standardized demonstration of safety for novel subsea technologies, applicable to and tested on a number of industry-relevant use cases (DNV GL 2019).

The remainder of this paper is organized as follows: literature study for safe design and the identified safe design principles are presented in Section 2. Section 3 starts with an explanation of our methodology which is followed by a discussion of the analysis result of the implementation of safe design principle identified in IEC 61508. Section 4 presents a study case on safe design of novel subsea technology. Finally, the paper is summarized with a conclusion and proposes a way forward in Section 5.

2. Safe Design

2.1 Aim and definition

The problem with defining the term “safe design” is that both the identified terms and the definitions in the literature are so diverse that it makes it difficult to generalize its meaning. Some of the identified terms include: safe architectural design (Hussey and Atchison 2000), safety in design (Drogoul et al. 2007), safety system design (Sharma 2017), safe by design (van de Poel and Robaey 2017), system resilience (Zhu et al. 2016), and inherent safety design (CCPS 2012). Some of the definitions are:

- Design of the implemented system to minimize the occurrence of system failures (Hussey and Atchison 2000).
- Safety methods employed to protect against or mitigate harm or damage to personnel, plants, and their environment to reduce risk in a system (Sharma 2017).
- The process of designing attributes and features into a design that enables its implementation to achieve the required level of safety (Drogoul et al. 2007).

A basis for suggesting a classification for these definitions is to investigate the aim of safe design.

In general, safe design employs a risk-based approach. Risk is used as a measure due to the definition of safety as “freedom from unacceptable risk of harm” (IEC 61511 2016). For example, inherent safety design aims specifically

to remove the risk of hazards by using a systematic approach to eliminate identified hazards (e.g. changing the chemical property) (CCPS 2012). The safe by design approach covers a more comprehensive risk management process in an attempt to avoid the expected hazards (van de Poel and Robaey 2017). The ineradicable risk, due to the required process and/or technology, needs to be reduced by reducing likelihood and consequence of the hazardous event.

System resilience is another concept that has some overlaps with the aim of safe design. System resilience focuses on the need to handle all hazards, including identified, omitted, and/or unforeseen hazards (that are not identified by the risk analysis process). The focus of resilience is the system ability to recover to normal operation in a short time (Zhu et al. 2016). Resilience may be built into the system by different measures, for example error recovery in software. This measure, however, has the limitation that it is always predetermined within the context of design (Hussey and Atchison 2000). For example, during power shutdown, we may program the software to move the system into a predetermined safe state. However, the software cannot confirm whether the initial premise of safe state is true (or not) and whether the safety purpose has been achieved since it does not have the ability for creative thinking that a human has.

To conclude, the aim of safe design is (i) to eliminate/reduce the risk of hazards during the design phase to a tolerable level and (ii) for omitted or unforeseen hazards, the system should have the capability to recover to normal operation.

The definition by Drogoul et al. (2007) is deemed to be the most suitable in light of our discussions above. The authors have introduced three additional terms to support their definition (Drogoul et al. 2007): (i) context: the boundary of the intended operations as assumed during the design, (ii) principle: the general intention of implementing the required attributes, and (iii) attribute: interpretation of requirement as a feature of system. It includes various aspects of man, technology and organization.

The output of design activity is a product that satisfies the required context (function, performance, environment, etc.). The product of safe design will have one (or several) safe

attribute(s) that work within the context of design and satisfy the safe design principle.

2.2 Safe design principle

According to Drogoul et al. (2007), there are seven safe design principles: (i) two levels of design, (ii) strategy against failures and errors, (iii) safety vs. complexity, (iv) modularity, (v) segregation, (vi) documentation, and (vii) demonstration of safety. Fundamentally, Drogoul et al.'s safe design principles can be classified into three categories: (1) how the system is built safely (by implementing safe attributes), (2) how to select "safer" technical solutions (given the recommended safe attributes), and (3) how safety of the technical solution can be justified.

Principles (i) and (ii) relate with the first category which is to specify safe attribute requirements of a system. Principle (i) "two levels of design" means that the selected design should have a high-level architecture and a detailed architecture of the system. High-level architecture includes the minimum required information (e.g. hardware, software, high-level function and expected performance) for communication between stakeholders. The detailed architecture specifies the functional realization by every technical solutions and interactions between system and/or subsystem. Principle (ii) "strategy against failures and errors" means that the designer acknowledges the possibility that the system may fail. It is necessary to prepare for countermeasures against failure by incorporating safe attributes such as fault detection, fault tolerance, fault response and fault recovery during different operational modes. Fault detection means that system needs to have capability to detect failure. Fault tolerance means the system has ability to tolerate failure up to a certain level. Fault response means that when the system fails, the appropriate response to ensure safety of the overall system needs to be defined. Fault recovery relates to the procedure to recover from a fault state to a working state.

Principles (iii), (iv), and (v) relate to the second category which ascribes the premise of technical solution choice for every safe attribute. Principle (iii) "safety vs. complexity" means that there is always a trade-off for every design solution regarding the safety performance and the complexity of the design that need to be made.

Principle (iv) “modularity” encourages simplicity and minimizes complexity to facilitate integration and testing of subsystems. Principle (v) “segregation” attempts to increase fault tolerance of the system and reduce complex interaction between system and subsystem.

Principles (vi) and (vii) relate with the third category which describes the need to justify every safe attribute. Principle (vi) “documentation” means that all decisions made regarding the solution during the design process, including the arguments and the assumptions used should be recorded. This helps the organization in two ways. First, it reinforces the trust to the analyst that the procedure for safe design has been performed and second, it can be used as reference when there is need for change (e.g. due to an accident or modification request). Finally, principle (vii) “safety demonstration” specifies the requirement to prove that the designed system can achieve an acceptable level of safety integrity. Acceptance criteria should be formulated as the target and evidence should be provided as support to the argument that the system is safe.

2.3 IEC 61508 and safe design

Functional safety of a system is realized through a step-by-step approach called the safety lifecycle (IEC 61508 2010). A system, designed according to IEC 61508, is called a safety-related system (SRS) and is used as safety barrier to achieve the required risk reduction.

IEC 61508 distinguishes between necessary safe attributes (using the word “shall”) and complementary safe attribute (using the word “should”) in the clauses. For necessary safe attributes, the users shall demonstrate evidence of how they implement the attribute in the system. If the evidence is missing or deemed inadequate by the assessor, the system is automatically considered non-compliant to the standard. However, for complementary safe attributes, it is up to user discretion whether to implement the safe attributes or not. Often several parameters are considered (e.g. complexity and cost).

A comparison of the aim between IEC 61508 and safe design shows similarity in their risk-based approach for safety. The only difference is the aspect that is considered. IEC 61508 distinguishes two types of risks to be eliminated/reduced: external and internal risks. A SRS is designed to either prevent or mitigate

external risk which is its function as a safety barrier. However, a SRS also has the possibility to behave in a harmful way (failure, unwanted interaction, etc.). In this case, SRS introduces a new (internal) risk to the overall system. Internal risks of the SRS should be managed to ensure that the intended risk reduction is not compromised. Comparatively, safe design does not distinguish between both risk types. It is because the term refers to a more general design process where the function of the system is not only for safety.

3. Analysis of the Implementation of Safe Design Principle in IEC 61508

3.1 Analysis scope and approach

A hypothesis could be made that IEC 61508 also employs safe design principles during the formulation of each clause. A comparison analysis is performed to check the hypothesis. The comparison scope is to section 7.2.3 and 7.4 of IEC 61508 (2010) part 2 where the clauses refer to the design requirement specification and design and development of SRS.

The approach is summarized into three steps: (1) identification of safe attributes, (2) classification of safe attribute under safe design principles, and (3) analysis of the findings. Every identified safe attribute is labelled by either necessary (N), complementary (C) or necessary or complementary (N || C) to denote the necessity of each attribute. Afterwards, they are compared and classified under the safe design principles by Drogoul et al. (2007). Analyses are performed separately to each of the safe design principles.

3.2 Results

The analysis is performed to 72 clauses from IEC 61508 part 2.

Principle (i) “two levels of design”. The SRS design is formulated based on the safety requirements specification. It includes the functional and the safety integrity (for architectural constraints) requirements. These requirements become the basis of the high-level architecture of the SRS. Detailed architecture information clarifies several things, such as: type of required hardware (e.g. E/E/PE) (N), realization of function by each element in the SRS (N), integration of main and subsystem (N), interaction between SRS, operator, and equipment under control (N), and system behaviour during different operational modes (N).

Principle (ii) “strategy against failures and errors”. IEC 61508 recommends several safe attributes as a strategy to handle internal risk. The safe attributes can be classified according to their function: fault tolerance, fault detection, fault response, and fault recovery.

For example, there are three safe attributes for fault tolerance: hardware specification (for environmental conditions and electromagnetic immunity) (N), derating of hardware performance (C) and redundancy of hardware (N || C). Hardware specification is necessary due to the context of design. Derating is a complementary safe attribute that should be considered as a measure to reduce the wear of hardware. Redundancy can be either necessary or complementary. It is necessary if required from the architectural constraint. However, it can be complementary if not required for safety but is required to ensure the availability of production.

Safe attributes of fault detection are proof test (N) and diagnostics (N). Safe attributes of fault response are automatic response to failure (N), safe state of both SRS and equipment (N), *and* fault isolation (N). Safe attributes of failure recovery are procedures during each operational mode (N) and repair (N).

Principle (iii) “safety vs. complexity”. While IEC 61508 does not advocate against designing a complex system, it warns about the possibility of having unknown faults or mistakes *and* the difficulty to demonstrate safety of the SRS. Therefore, the complexity (C) of SRS should be in a “manageable” level (although interpretation of the term may vary). Requirements for the avoidance and control of systematic faults are the recommended methods to manage complexity.

Principle (iv) “modularity”. Modularity in IEC 61508 refers to two different purposes: (1) to define every subsystem and their interface and (2) for approval of the design solution. The first type (N) is used as an attempt to control complexity and prevent the introduction of systematic faults in the design process. The second type refers to independent (or modular) approvals (C) of the SRS (including hardware, software, test, programming tools and appropriate language for software) that are intended to reduce the complexity of system application engineering.

Principle (v) “segregation”. IEC 61508 uses the term independence to cover requirements for

segregation. We may consider two levels of independence: physical (e.g. redundancy (N || C), diverse technology (C) and no common parts (C)) and functional (e.g. functional diversity (C), no common procedures, and application (C)). They can be used to achieve fault tolerance of the SRS. Independence of the system shall be justified by performing common cause failure analysis. Independence (C) of the SRS (or element in the SRS) is intended to reduce the complexity in carrying out system safety lifecycle activities (e.g. design, validation, assessment and maintenance).

Principle (vi) “documentation”. IEC 61508 demands proper documentation (N) for every section (or subsection) in the standard. It is required as evidence that the design process has been performed “correctly” and checked when the management of functional safety is performed.

Principle (vii) “demonstration of safety”. The analysts are required to derive their own acceptance criteria for every element of the SRS. Examples of criteria are: high-level performance (e.g. according to Sklet (2006): effectiveness (N), reliability/availability (N), response time (N), robustness (N), and condition (N)) and the safety integrity (e.g. safety integrity level (N)). Safety demonstration (e.g. test (N) and analysis (N)) is performed to collect evidence (e.g. documentation (N), test results (N), analysis (N), and certificates (N)) to support the safety claim.

3.3 Discussion on IEC 61508

In general, IEC 61508 is aligned with the safe design principles for formulation of its clauses.

Although not explicitly mentioned, IEC 61508 also regulates the management of external risk through the same safe design principles. However, it differs slightly on the perspective with the principle (ii) “strategies against failures and errors”. The strategies mentioned here are to manage the internal risks of the SRS to ensure that the target (external, from SRS perspective) risk reduction is achieved. In a sense, SRS is one of the strategies against failures and errors (or internal risk) of a process system.

During the analysis process, one of the main challenges is to determine the safe attribute from the clauses. The reason is that the definition of attribute is very broad, from human (e.g. repair), technical (e.g. redundancy), and organizational (e.g. procedures). We acknowledge the risk of

omission during all parts of the analysis approach, and we have tried to ensure completeness of the results by performing the approach several times.

4. Study Case

Safe design is already a practice in several industry domains, such as: space (Sgobba et al. 2009), process (CCPS 2012), and oil and gas (Kjellén 2007). For subsea, the safe design practice refers to the oil and gas domain which is developed with a topside context in mind (Kim et al. 2016). Kim et al. (2016) pointed out that, at least in Norway, there are gaps in the current safety philosophies between topside and subsea oil and gas system which may result in overly complex and costly design solutions for subsea. A study is required to identify the applicability of safe design principles for subsea systems.

One example is a concept involving integration of process control and safety (IPC&S). The available information collected through literature research shows that the IPC&S concept is new or unproven (it was proposed in process industry domain but is not widely used in practice (Gruhn and Cheddie 2006, CCPS 2016)). The technology would then be applied for a new application area (subsea). Our high-level analysis concludes that IPC&S system possibly demands new technical challenges (refer to DNVGL-RP-A203 (2017) table 7-1 “technology categorization”).

In practice, complete separation for subsea systems is difficult to achieve due to the presence of subtle dependencies at the lower component level (e.g. dependencies of physical, logical, and location) (DNV GL 2019). However, they have been generally accepted due to two reasons: (1) The cost required to implement a completely separated system is higher than the benefit gained and (2) the consequence of failure is system shutdown which is considered as a safe failure. The IPC&S concept provides opportunities to reduce the equipment cost. However, for failure consequence, IPC&S attempts to integrate components that may introduce new types of dangerous failure (e.g. harmful interaction). Based on these reasons, an exhaustive analysis is required to justify the use of the IPC&S concept.

4.1 System description

Fundamentally, there are three possible types of integration/separation for subsea systems:

- Complete separation. Two separate components are required to perform each process control and safety function (e.g.: NOG GL 070 (2018)).
- Physical integration with logical separation. Shared hardware is used together with both temporal and spatial separation in the software (e.g.: IEC 61508 (2010), part 3, appendix F).
- Complete integration. Both process control and safety functions are performed in one hardware and software with no logical separation (no known standard and regulation in the process industry advocates this integration type).

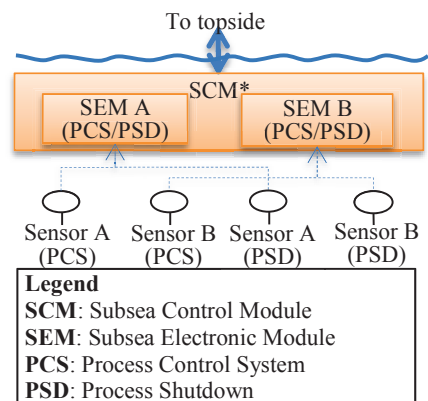


Fig. 1. Shared controller concept for subsea system (*: Physical integration of hardware with functional separation of software).

The study case is a subsea system that implements second integration/separation type as depicted in Fig. 1 (latter called “IPC&S system”). Both process control and the safety system are used to protect the same equipment. Integration occurs at the subsea electronic modules (SEMs) where they implement functional separation on the software for both process control and safety. The system that is used as reference for (relative) comparison is the “traditional system” that employs the complete separation solution.

4.2 Findings on the safe design of IPC&S

Introducing integration to subsea system have implications to the implementation of safe design principles. The findings from high-level analysis of IPC&S system are discussed below.

Principle (i) “two levels of design”. The high-level architecture of the IPC&S system can be obtained using the current approach in IEC 61508. However, problems arise when formulating requirements for the detailed architecture.

Integration introduces new hazard types due to the possibility of harmful interaction. The current approach, stemmed from chain-based accident models (e.g. FMECA and HAZOP) is not suitable to discover this problem (Kim et al. 2018). An approach that takes both systemic perspective and the effect of interaction is necessary.

Principle (ii) “strategy against failures and errors”. Due to hardware integration, failure of one SEM will reduce the hardware fault tolerance (HFT) to zero (failure of the other SEM will result into direct shutdown of both process and safety system). From safety point of view, HFT of 0 is acceptable for the SEM up to a certain SIL level (depend on safe failure fraction). However, the redundancy is sometimes required for availability. One alternative to compensate for this is by increasing reliability and/or retrievability of the SEM. However, the cost incurred can be high. Additional analysis should be performed to identify the cost-effectiveness of the strategy.

According to IEC 61508, functional separation within the software part shall be tested since there is a possibility for either conflicting requirements or uncontrolled interactions between the process control and the safety function. It could affect both fault detection and fault response of the IPC&S system.

When the IPC&S system needs to be restarted (e.g. due to failure, power loss, or after executing the safety function), a procedure should be prepared to ensure that failure recovery process is able to return both the IPC&S system and the process to a working condition.

Principle (iii) “safety vs. complexity”. There is a trade-off between the complexity of physical architecture, which decreased due to reduced components number and the complexity of the realization of the function, which increased due to functional separation in software. Software itself has inherent complexity (Leveson and Weiss 2009). It is difficult to justify whether the traditional system or the IPC&S system has better safety than the other. It is reasonable to identify the potential unsafe behaviour and focus the effort to manage this problem.

Principle (iv) “modularity”. It is required to clarify how the subsystems interact with each other, especially between integrated components and their directly related components. However, modular test of each subsystem may be

insufficient. Integrated test should be emphasized to check the possible behaviour of the system (although it should also be acknowledged that test cannot reveal all possible system behaviours).

Principle (v) “segregation”. The IPC&S system employs the functional separation concept. As discussed earlier, analysis and test are required to ensure that the isolation between both functions can be achieved.

Principle (vi) “documentation”. The documentation process is similar as for traditional systems. The increased number of measures and requirements of the IPC&S system requires more exhaustive documentations.

Principle (vii) “demonstration of safety”. The general premise of the recommended methods for allocation of performance target is that, if the IPC&S system has multiple protection layers (e.g. process control and safety), they should be independent of each other. However, dependency within the IPC&S system nullifies the premise of independent protection. Two things need to be done: (i) development of a new allocation approach that considers dependency between all protection layers, (ii) propose a method to quantify the dependency between both systems.

One of the required evidence for safety demonstration of IPC&S systems is the dependency effect to overall system risk. Most of the evidences are obtained via qualitative analysis which is difficult to justify (by measures for avoidance and control of systematic failure). Hauge and Hokstad (2006) proposed to use expert judgement to approximate the contribution of systematic failure to the overall system risk. However, it is not known to be widely used in the industry, presumably due to the possibility of bias (e.g. due to subjectivity and experience). Quantitative analysis is only limited to the dependent failure analysis (e.g. common cause failure). Further work in both qualitative and quantitative aspects are required to justify the safety claim of an IPC&S system.

5. Conclusion

In this paper, a clarification has been made to the term safe design and on how IEC 61508 implements the safe design principle in its clauses. However, based on high-level analysis, several practical implications arise when trying to implement the safe design principle for the study

case IPC&S system. The discussions in section 4.2 become our inspiration for future research.

Acknowledgement

This paper has been written under the Safety 4.0 project. The author would like to thank the Research Council of Norway, the industrial, and universities partners involved in this project for the support.

References

API RP 17N. 2017. Recommended Practice on Subsea Production System Reliability, Technical Risk, and Integrity Management. American Petroleum Institute.

Bolbot, Victor, Gerasimos Theotokatos, Luminita Manuela Bujorianu, Evangelos Boulougouris, and Dracos Vassalos. 2019. "Vulnerabilities and safety assurance methods in Cyber-Physical Systems: A comprehensive review." *Reliability Engineering & System Safety* 182:179-193.

CCPS. 2012. *Guidelines for Engineering Design for Process Safety*. 2nd ed. USA: John Wiley & Sons, Inc.

CCPS. 2016. *Guidelines for Safe Automation of Chemical Processes*: John Wiley & Sons.

DNV GL. 2019. "Safety 4.0 - Demonstrating safety of novel subsea technologies." DNV GL, accessed 24 January. <https://www.dnvgl.com/technology-innovation/oil-gas/safety40/project-description.html>.

DNVGL-RP-A203. 2017. "Qualification of new technology." *Recommended Practice*.

Drogoul, F., S. Kinnersly, A. Roelen, and B. Kirwan. 2007. "Safety in design – Can one industry learn from another?" *Safety Science* 45 (1):129-153.

Gruhn, P., and H. Cheddie. 2006. *Safety instrumented systems : design, analysis, and justification*. 2nd ed. Research Triangle Park, N.C.: ISA-The Instrumentation, Systems, and Automation Society.

Hauge, S., and P. Hokstad. 2006. *Reliability prediction method for safety instrumented systems : PDS method handbook*. Vol. A13503. Trondheim: SINTEF.

Hussey, A., and B. Atchison. 2000. Safe architectural design principles. In *Technical Report*. Queensland, Australia: Software Verification Research Centre.

IEC 61508. 2010. Functional safety of electrical/electronic/programmable electronic safety-related systems. In *Part 1-7*: international electrotechnical commission.

IEC 61511. 2016. Functional safety - Safety instrumented systems for the process industry sector. In *Part 1-3*: international electrotechnical commission.

IEC 61513. 2011. Nuclear power plants - Instrumentation and control important to safety -

General requirements for systems. international electrotechnical commission.

Kim, H., M.A. Lundteigen, A. Hafver, F.B. Pedersen, G. Skofteland, C. Holden, and S.J. Ohrem. 2018. "Application of systems-theoretic process analysis to a subsea gas compression system." In *Safety and Reliability—Safe Societies in a Changing World*, 1467-1475. CRC Press.

Kim, H., M.A. Lundteigen, and C. Holden. 2016. "A gap analysis for subsea control and safety philosophies on the Norwegian continental shelf." 13th International Conference on Probabilistic Safety Assessment and Management (PSAM 13), Seoul, Korea.

Kjellén, U. 2007. "Safety in the design of offshore platforms: Integrated safety versus safety as an add-on characteristic." *Safety Science* 45 (1):107-127.

Leveson, N.G., and K.A. Weiss. 2009. "Chapter 15 - Software System Safety." In *Safety Design for Space Systems*, edited by Gary Eugene Musgrave, Axel M. Larsen and Tommaso Sgobba, 475-505. Burlington: Butterworth-Heinemann.

Lundteigen, M.A., M. Rausand, and I.B. Utne. 2009. "Integrating RAMS engineering and management with the safety life cycle of IEC 61508." *Reliability Engineering & System Safety* 94 (12):1894-1903.

Nair, S., J.L. de la Vara, M. Sabetzadeh, and D. Falessi. 2015. "Evidence management for compliance of critical systems with safety standards: A survey on the state of practice." *Information and Software Technology* 60:1-15.

NOG GL 070. 2018. Norwegian oil and gas application of IEC 61508 and IEC 61511 in the Norwegian petroleum industry. Norsk olje & gass.

Sabetzadeh, M., D. Falessi, L. Briand, and S. Di Alesio. 2013. "A goal-based approach for qualification of new technologies: Foundations, tool support, and industrial validation." *Reliability Engineering & System Safety* 119:52-66.

Sgobba, T., A.M. Larsen, and G.E. Musgrave. 2009. "Chapter 4 - Basic Principles of Space Safety." In *Safety Design for Space Systems*, edited by G.E. Musgrave, A.M. Larsen and T. Sgobba, 163-174. Burlington: Butterworth-Heinemann.

Sharma, K.L.S. 2017. "18 - Safety Systems." In *Overview of Industrial Process Automation (Second Edition)*, edited by K. L. S. Sharma, 293-319. Elsevier.


van de Poel, I., and Z. Robaey. 2017. "Safe-by-Design: from Safety to Responsibility." *NanoEthics* 11 (3):297-306.

Zhu, Q., D. Wei, and K. Ji. 2016. "Hierarchical Architectures of Resilient Control Systems: Concepts, Metrics, and Design Principles." In *Cyber Security for Industrial Control Systems From the Viewpoint of Close-Loop*, 151-182. CRC.

Article II

N. A. Zikrullah, H. Kim, M. J. P. van der Meulen, G. Skofteland, M. A. Lundteigen, A comparison of hazard analysis methods capability for safety requirements generation, Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability. 2021, 1748006X211003463.

A comparison of hazard analysis methods capability for safety requirements generation

Proc IMechE Part O:
J Risk and Reliability
1–22
© IMechE 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1748006X211003463
journals.sagepub.com/home/pio


Nanda Anugrah Zikrullah¹, Hyungju Kim², Meine JP van der Meulen³,
Gunleiv Skofteland^{1,4} and Mary Ann Lundteigen¹

Abstract

A safety-critical system comprising several interacting and software-intensive systems must be carefully analyzed to detect whether new functional requirements are needed to ensure safety. This involves an analysis of the systemic properties of the system, which addresses the effect of the interaction between systems and system parts. The paper compares two hazard analysis methods, which are often considered well-suited for such software-intensive systems: the Functional Hazard Analysis (FHA) and Systems-Theoretic Process Analysis (STPA). The focus is on the selection and improvement of the best methods, based on the lesson learned from the comparison of FHA and STPA. The analyses cover the hazard analysis processes, systemic properties, and the criteria of requirements. The paper concludes that STPA is the better choice over FHA. Insights are obtained to align both STPA and FHA methods with the broader topic on risk management, that is, hazard analysis method improvement, cautionary thinking, uncertainty management, and resilience management.

Keywords

Hazard analysis, functional hazard analysis, systems-theoretic process analysis, functional requirement, software-intensive system

Date received: 10 February 2021; accepted: 18 February 2021

Introduction

When novel technologies involving more electronics and programmable systems are developed to increase the efficiency and safety of a system in the industry, it may lead to more complex interactions of hardware and software, with failure modes that are difficult to foresee. Failures may not only stem from component failures, but can also be systemic due to unintended interaction of component and functions.^{1,2} Hence, it is important to select suitable analysis tools to identify possible ways in which the system might fail, including systemic failures. Many sectors rely on IEC 61508³ to qualify novel Electrical/Electronic/Programmable Electronic technology for systems that are critical for ensuring industrial facilities' safety. According to the standard, a hazard analysis process is necessary before the system can be qualified for operation.^{3–5}

A good starting point before selecting a hazard analysis method is to define the relevant terms. Hazard is defined as a *source of danger that may cause harm to an asset*.⁶ A hazardous event is *the point at which control*

of the hazard is lost.⁶ The event involves interaction between the hazards and the contextual conditions (e.g. environmental state or human activity). Hazard analysis is a process to identify hazards, hazard consequences, and the causal scenarios (or factors) leading to the hazards.⁵ Management of such hazards (e.g. by prevention or mitigation) may result in additional system requirements that might affect its design, operation, and maintenance activities.^{3,7}

If the hazard analysis methods are to be applied to novel technologies, they must have several

¹Norwegian University of Science and Technology (NTNU), Trondheim, Norway

²University of South-Eastern Norway (USN), Borre, Norway

³DNV GL, Høvik, Norway

⁴Equinor, Trondheim, Norway

Corresponding author:

Nanda Anugrah Zikrullah, Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), Høgskoleringen 1, Trondheim 7491, Norway.
Email: nanda.a.zikrullah@ntnu.no

characteristics. For example, the methods should be suitable for analyzing functions, rather than their realization. This means that the analysis should consider the expected (or specified) behavior that may harm the system, rather than the actual behavior since many of the realization details are abstract.^{4-6,8} Also, the method should facilitate a systemic approach,^{1,9} whereby the system elements and the implication of their interactions are revealed at the system level. Last, the methods should allow for a structured approach to producing new design and operation requirements based on hazardous scenarios.¹⁰ The purpose is to integrate the hazard analysis results in the system development process. Based on the above-described characteristics, we identified several alternative methods of hazard analysis: Preliminary Hazard Analysis (PHA), Functional Hazard Analysis (FHA), Software System Failure Mode and Effect Analysis (SSFMEA), Hazard and Operability study (HAZOP), Systems-Theoretic Process Analysis (STPA), and Functional Resonance Analysis Method (FRAM).^{1,2,5,11} Some of these methods have been advocated as part of the sector-specific standards, including aerospace industry⁴ (FHA), automotive industry¹² (HAZOP), and process industry¹³ (PHA and HAZOP). STPA and FRAM are relatively recent hazard analysis methods that have attracted wide attention.¹⁴⁻¹⁶ STPA has recently been recommended in ISO/PAS 21448¹⁷ to ensure the safety of the intended functionality of autonomous vehicles. Variants of the above-described methods are not explored further in this paper (e.g. control-HAZOP is considered HAZOP). The only exception is in SSFMEA, which is a system-based analysis, whereas the original Failure Mode and Effect Analysis (FMEA) is a component-based analysis.

The long list of hazard analysis methods makes the selection for the most suitable method a challenge. The main objective of this paper is to analyze and compare the hazard analysis methods based on the characteristics mentioned above. The goal is to select and, where needed, improve the best method for hazard analysis of novel technology. The objective comprises the following three research questions:

- RQ1. How do the selected hazard analysis methods identify the same or different functional hazards?
- RQ2. How do the selected hazard analysis methods provide a systemic perspective on the system for analysis?
- RQ3. What are the main differences between the derived safety requirements?

The remaining part of this paper is organized as follows. The next section provides a review of the list of hazard analysis methods and the preliminary selection made to limit the comparison process into two methods based on the derived characteristics. The *methodology* section describes the approaches to answer the research

questions and includes the procedures for hazard analysis. The *case study* section describes the example from the oil and gas industry to demonstrate the two methods' capability. This is followed by a presentation of the *results* of the analysis and discussions on the findings. Section *overall implication* contains our recommendations and the implications for other subject areas. The final section concludes the finding in the paper.

Review of the hazard analysis methods

We reviewed the hazard analysis methods to limit the number of methods to be considered for further analysis into a maximum of two. We identified two attributes that capture the methods' functional and systemic characteristics: the ability to capture the undesired functional behavior and the linearity of the utilized accident model. The requirement generation characteristic requires an in-depth understanding of the methods' results. Hence, it was not considered suitable for inclusion as part of the preliminary review.

Ability to capture the undesired functional behaviors

During operation, the actual behavior of functions may deviate from expectations. Examples of the functional behavior are the realization of function (e.g. activated, not activated, when needed, not needed, as required, too short, or too much) and the function timing (e.g. correct, early, or late). The undesired functional behavior needs to be assessed according to the context (e.g. where and when it may occur) to be classified as a functional hazard.

All methods have different procedures to identify hazards (e.g. the required inputs, the process, and the outputs¹⁰). Some methods might have influenced each other during decades of development, resulting in substantially similar hazard identification procedures. For example, PHA was designed to analyze broader types of hazards, including energy source, functional, operational, component, material, lesson learned from other systems, undesired mishaps, and failure modes.⁵ These hazards are captured through the use of a checklist. PHA is designed to be a preliminary analysis and has extensive coverage. The results of the analysis performed using the method suffer from the lack of depth, and therefore additional methods are needed to supplement the process.

Ericson⁵ recommends using FHA for analysis of functional hazards because the method utilizes a list of functional hazard types (e.g. functional failure, operates incorrectly, and function timing). A variant of FHA called Functional Failure Analysis (FFA) focuses on how the function can fail.¹⁸ Both of them are deemed the same method because they utilize a similar functional hazard type list. Many authors also consider FFA a variant of FMEA known as *predictive FMEA*, due to the utilization of the FMEA method.¹⁸ The FMEA method involves systematic checking for

possible combinations of functions, failure mode types, and operational mode. In this paper, the term FHA is used to represent FHA and FFA.

According to Pumfrey,¹⁸ both SSFMEA and (software) FHA utilize the same procedures to identify undesired functional behavior. SSFMEA is tailored to analyze the software's functional behavior. By contrast, HAZOP was initially developed to analyze hazard and operational problems in system design⁶. HAZOP analyzes combinations of parameters (e.g. flow or pressure) and guide words (e.g. more, less, no) to check the possible deviation from the design intent. STPA regards hazards as all unsafe control actions (UCAs) performed by controllers to the system (or controlled processes) that occur in a specific context.¹ Finally, FRAM checks whether the aggregation (or coupling) of the variability of all functions in the system may result in an increased, unchanged, or dampened variability at the system level.²

Linearity of the accident model

Causal analysis processes for the hazards are developed based on an accident model. Hollnagel² states that the accident models can be classified into three types, based on differences in their principles of causality: simple linear models (e.g. the Domino model), complex linear models (e.g. the Swiss Cheese model), systemic model (e.g. the Systems-Theoretic Accident Model and Process (STAMP), and the Functional Resonance Accident Model). In a simple linear model, the accident is caused by a linear sequence of causes (e.g. failures, errors, or organizational problems). Here, the focus is to provide recommendations to eliminate one cause in the sequence. In a complex linear model, dependencies between events may affect the event sequence that results in accidents. To manage this dependency, the focus is shifted by strengthening the barriers and defenses. In a systemic model, the dependencies are not only due to a combination of events but also due to complex couplings between interacting components. An accident can be prevented by controlling the system state to prevent transition into an uncontrolled (unsafe) state.^{1,2}

Both the simple linear model and the complex linear model have been utilized in the causal analysis process of the hazard analysis methods such as PHA, FHA, SSFMEA, and HAZOP. Initially, the causal analysis focuses only on finding the direct root cause of a hazardous event (simple linear model). This approach works due to the simplistic type of system utilized at the time (i.e. mechanical or hydraulic system). When the number and complexity of the system's components increase (i.e. electronic system), the interaction or dependency may become a significant contributor to a hazardous event. Hence, a complex linear model is then

adopted to the traditional method to increase the analysis coverage. The shift from utilizing a simple linear accident model to a complex linear accident model shows how the methods' causal analysis process is evolving depending on the system to be analyzed (i.e. simple or complex).

Recently, the systemic accident model has been developed to include the different complexity characteristic of the system. Leveson's¹ and Hollnagel's² criticize of the limited perspective of the linear accident models. According to them, while dependencies are considered already in the complex linear model, they still occur due to combinations of failures. A systemic model allows for identifying possible harmful interactions without failure in the system. STPA and FRAM are the hazard analysis methods that utilize the systemic accident model

Several comparison analysis results support their critics. For example, Leveson et al.¹⁹ perform a comparison between STPA and the ARP 4761 safety assessment process and claim that the former is better for safety assessment. However, they did not indicate whether this difference in result is due to the accident model used or due to the flaws of the methods utilized in ARP4761. For example, to claim that FHA (part of ARP 4761) considers only failures during the analysis does not mean that it is limited to consider component failure as a cause. It is possible to expand the perspective to the systemic level and find that functional failure can also be caused by an interaction problem between two or more components (without any failure). This argument shows that the limitation in ARP4761 is not because of the method but by the accident model's limitation.

Yousefi et al.¹⁵ compare AcciMap, STAMP, and FRAM. This comparison focuses on the systemic model and does not discuss the contrast with the linear model. In another research, Sulaman et al.²⁰ have a different claim. They perform a comparison between Software System FMEA (SSFMEA) and STPA for a collision-avoidance system and conclude that neither method is superior. Some hazards are unique to both SSFMEA and STPA. They claim that both methods complement each other. The SSFMEA method that they utilized focuses more on component failures and does not have a systemic perspective of the system due to the bottom-up approach.

The examples above show the systemic accident model's advantages over the linear accident model for causal analysis. This does not mean that the traditional hazard analysis methods (e.g. FHA) are not as good as the new hazard analysis methods (e.g. STPA and FRAM). The shift from a simple linear model to a complex linear model in the traditional methods indicates that they can apply a new accident model for improvement. If the systemic model is as better as it is claimed, research on its application need to be

Table 1. Review of hazard analysis method attributes.

Methods	Ability to capture undesired functional behavior	Linearity of the accident model
PHA	Type of functional hazards (can be expanded to other type of hazards)	Complex linear model
FHA	Type of functional hazards	Complex linear model
SSFMEA	Type of functional failures	Complex linear model
HAZOP	Combination of guidewords and parameters for process condition	Complex linear model
STPA	Type of unsafe control actions	Systemic model
FRAM	Aggregation of variability in the function	Systemic model

performed with the traditional methods. This would provide users with options to develop the traditional methods (if possible) or to utilize the new methods.

Method selection

Table 1 summaries the attributes of the reviewed hazard analysis methods. The varying abilities to capture the undesired functional behavior make it difficult to distinguish between each method. Therefore, the method selection is mainly based on the linearity of the accident model, with one method for each model. This is also to verify the claim for the systematic accident model advantages over the linear accident model. Logical reasoning and reviews of relevant literature are performed to support the decision.

For the complex linear model, PHA, FHA, and SSFMEA have similar procedures in capturing the undesired functional behavior, with FHA as the recommended method for analyzing functional hazards. Comparatively, HAZOP may not be suitable for analyzing novel technology due to a lack of detailed system design. Therefore, FHA is selected for the method with a complex linear model.

For the method with a systemic model, we refer to the comparison analysis by Yousefi et al.¹⁵ He finds that STPA is more capable of finding hazards systematically as compared to FRAM. We use this finding as the basis for the selection of STPA in the paper.

Methodology

The research methodology is as follows. First, FHA and STPA are performed separately on a case study. The functional list's input to both methods is controlled to be the same to accentuate the differences between both methods' results. It is validated by associating each function in the FHA to the function in the STPA. Both hazard analyses are performed by the same person (first author). This may introduce subjectivity in the assessment process. Verification is performed by all the authors on the presented results to reduce the subjectivity. The case study focus is on both method's ability to identify hazards and produce requirements. Therefore we decided not to do a risk assessment for both methods.

Then, a comparison analysis is performed to answer the RQ1. A mapping between FHA and STPA procedures is required for the comparison process, which is described later. The analysis focuses on analyzing the cause of the similarity or difference of the results from every step of the hazard analysis methods.

RQ2 is answered by comparing the properties of the causal scenarios with the system properties. We utilize the Composition, Environment, Structural, and Mechanism (CESM) model²¹ as the reference system properties. Composition refers to every component that built the system (e.g. controller, sensor). Environment refers to the boundary condition in which the system may influence or be influenced by (e.g. water depth or temperature). Structure refers to the (physical or abstract) relation between the components or the components and the environment in the system (e.g. communication between components). Mechanism is a process that describes the behavior of a given component, structure, or environment (e.g. interaction in the software function). According to Wan,²² the CESM model can aid in investigating systemic behavior (i.e. emergence). Thus, we can evaluate whether these four properties in the hazard analysis method can lead to the identification of systemic causal scenarios.

RQ3 is answered by evaluating the requirements against the criteria for a requirement. While there is no consensus on what makes a good requirement, Holt et al.²³ state that these eight criteria should be considered: (1) identifiable, (2) clear, (3) solution-specific, (4) have ownership, (5) have origin, (6) verifiable, (7) able to be validated, and (8) have priority. (1) *Identifiable* refers to the ability of the requirements to be traced back to their cause. (2) *Clear* refers to the need to have unambiguous meaning for every requirement. (3) *Solution-specific* refers to the application of the requirements to a specific system. (4) *Have ownership* refers to the stakeholders that need to satisfy the requirements. (5) *Have origin* refers to the targeted subjects that need to follow the requirements. (6) *Verifiable* refers to the ability of the requirements to be checked for correctness by the designer. (7) *Able to be validated* refers to the ability of the requirement to be demonstrated for compliance. (8) *Have priority* refers to the relative level of importance of one requirement to the other. We assumed that the above criteria are necessary to form a requirement that can be utilized immediately for

decision making. Thus, we can utilize them to evaluate whether the hazard analysis methods can provide such requirements.

Finally, all the research questions' analyses results are discussed at a higher level to conclude the selection of the better method for hazard analysis of novel technology. The research's implication is analyzed according to the risk management topic in general, to indicate the required next step for integration of the method with the safety assessment process.

The following subsections describe the FHA and STPA procedures. Modifications are applied based on the identified literature. Afterward, a mapping of FHA and STPA procedures is provided for the comparison analysis process.

Functional hazard analysis (FHA) procedure

FHA procedures have evolved over the years. It seems that there is no consensus on how exactly FHA should be performed.^{4,5,8,24} While there are different wordings and number of steps in FHA from different sources, essentially, the procedure includes the following seven steps:

- (1) Describe the system. The system description may be obtained from the conceptual design and operation of the system and functional list.⁵
- (2) Model the interactions of the functions. The model may be constructed based on the functional list. While this step is not recognized as a separate step in the referred documents, Ericson⁵ recommends using a model to aid the analysis. Examples of the modeling methods are the functional flow diagram and Functional Analysis Structure method (FAST) diagram.
- (3) Identify hazards. Hazards may be identified systematically by checking the combinations between functions, operational modes, and functional failure modes.⁴ The operational modes are obtained based on the conceptual operational procedures for each function. The functional failure mode is a generic list that is defined early before the hazard identification starts. Examples of functional failure modes: *functional loss*, *unintended activation*, and *incorrect operation*.²⁴
- (4) Identify consequences. Each consequence may be identified by checking the possible propagation effects from the functional hazard to the system level (e.g. using an inductive method⁵).
- (5) Analyze causal factors (or scenarios). A single (or a combination of) causal factor(s) may form a scenario that caused hazards. The causal factors are based on conceptual design and operation, the function model, and historical experiences. ARP 4761⁴ focuses on causal factors due to failure. As argued in the previous discussion, it may be

possible to expand the causal factors' perspective into possible scenarios involving multiple causal factors with no failure. No failure means that the system has been implemented according to the specification, but the specification lacks the ability to handle the scenarios.

- (6) Assess risk. The risks for every hazardous event are assessed from the magnitude of the consequences and the likelihood of every causal scenarios.⁴ According to Rausand,⁶ the risk analysis process for the hazard analysis method may be qualitative (e.g. utilizing qualitative scale) or semi-quantitative (e.g. utilizing risk priority number).
- (7) Provide recommendations or generate functional requirements. Depending on the analysis purpose, it is possible to either directly recommend solution(s) to prevent/mitigate the hazard or to generate a functional requirement²⁴ as guidance during the detailed design process. The first option is preferable for mature technology with historical experience. For the conceptual design of new technology, functional requirements are better as they do not limit the possible solutions. The functional requirements can be coupled with other methods (e.g. FTA, FMEA, and common cause analysis) to derive the non-functional requirements (e.g. reliability and safety performance requirements) as performed in ARP 4761.⁴

Several researchers^{8,24} has demonstrated the FHA for hazard analysis at the system level and find several weaknesses of FHA. Allenby and Kelly²⁴ argue that the generic functional failure mode list in step 3 still has a limitation due to the overuse of incorrect operation hazard type as the complementary keyword to capture abstract functional failures. They propose to utilize HAZOP guide words to obtain more comprehensive safety requirements.²⁴ Besides, the processes of causal and consequence analysis are still based on a brainstorming process that does not guarantee the completeness of the results.^{5,8} Wilkinson and Kelly⁸ claim that it is challenging to discover coupling or dependent failure causal scenarios using the brainstorming process.

Based on the identified weaknesses above, we made several considerations for FHA's application in our study. First, system modeling was supported by using a FAST diagram. A FAST diagram depicts the model sequence and dependency between functions (e.g. main, supporting, and continuous).²⁵ Each function is modeled as a box with connections to the other functions and may have different roles in the system (e.g. main function or supporting function). In the FAST diagram, the right function is the precursor of the function to the left (a sequence).

Next, HAZOP guidewords (i.e. omission, commission, late, early, and value) were utilized for a functional failure mode list as recommended by Allenby

Table 2. Transformation rule from keywords into FHA functional requirement and STPA controller constraints.

Keywords	FHA functional requirements	STPA controller constraints
Omission/not provided (when needed)	... Must be provided Must provide ...
Commission/provided (when not needed)	... Must not be provided Must not provide ...
Provided too late	... Must work within required time Must provide within required time ...
Provided too early	... Must not start working too early Must not provide too early ...
Stopped too soon	—	... Must provide continuously as required ...
Applied too long	—	... Must stop providing after the condition changes ...
Provided wrong value	... Must be provided correctly ...	—

and Kelly²⁴ to have a comprehensive scope for the analysis.

For causal scenario analysis, we utilized the FAST diagram and the system conceptual design and operation. The possible causal scenario was obtained by identifying the potential agent (or component) performing the function and its dependency on the next function. Information from the conceptual design and operation is used to infer the agent's (e.g. temperature or pressure) possible external effect on the system. We decided not to go too deep into detail to maintain simplicity (e.g. rotor, stator, or motor shaft failure would be assumed as one pump motor hardware failure).

We developed a rule for safety requirement generation to transform the functional failure mode keywords into functional requirement keywords. The transformation rules are listed in Table 2.

Systems-theoretic process analysis (STPA) procedure

STPA utilizes system theory and system thinking based on STAMP. The STPA procedure consists of four steps:²⁶

1. Define the purpose of the analysis
 - (a) Describe the system. The system description is based on the conceptual design and operation of the system and functional list.
 - (b) Identify System-level Loss, System-level Hazards, and System-level Safety Constraints. They may be obtained through a brainstorming process based on system description and experience from similar systems.
2. Model the control structure
 - (a) Identify controller responsibility and process model. They may be developed based on a system description. They describe how the controller responds to new/updated information.
 - (b) Build the Hierarchical Control Structure (HCS) model. The model is constructed based on the functional list, controller responsibility, and process model. Every agent in the system (e.g.













controller, controlled process, or supporting system) is modeled as a box. Each box may have connections (modeled as arrows) with other boxes based on the functions (e.g. control actions or feedbacks). In the HCS, the controller is an agent responsible for controlling agents at the lower hierarchy level.

3. Identify Unsafe Control Actions (UCA)
 - (a) Identify UCAs. Each UCA may be identified by checking the combination between control actions, environmental conditions/system states, and UCA types. Control actions are obtained from the controller responsibilities. Environmental conditions are obtained from the process model. There are six types of UCA: control action not provided when needed, provided when not needed, provided too late, provided too early, stopped too soon, and applied too long.
 - (b) Generate Controller Constraints (CC). Each CC may be generated by transforming the UCA type keywords into constraint keywords (e.g. not provided is transformed into must provide).²⁷
4. Identify Loss Scenarios (LSc). Each scenario may be identified based on every aspect in the control loop (e.g. controller, sensor, actuator, controlled process, communication, and environmental influence).

Several researchers have demonstrated STPA for analysis of complex systems^{28–30} and found several weaknesses of STPA. Due to the attempt to increase the hazard coverage, STPA suffers from a state explosion of the number of UCAs to be analyzed.³¹ Prioritization is required as follow up to focus the available resource. Also, the use of STPA is not straightforward since it requires the analyst to develop an HCS. This may not be a familiar task for the common practitioner of hazard analysis.²⁸ Finally, Kim et al.³⁰ also question the absence of stop criteria preventing the analyst from going too deep into the details.

Based on the identified weaknesses above, we have made considerations for applying STPA in our study.

Table 3. Mapping of FHA and STPA procedures.

FHA term	FHA	(Generic) hazard analysis procedures	STPA	STPA term
–		System description		–
Functional analysis structure technique diagram		System modeling		Hierarchical control structure
Hazardous events		Hazard identification		Unsafe control action
Consequence		Consequence identification		System-level loss and system-level hazard
Causal scenario		Causal scenario analysis		Loss scenario
Safety requirement		Safety requirement generation		Controller constraints
				

First, we did not perform a prioritization for STPA since it does not conform with the original intent of the STPA method by Leveson.¹ She argues that the main strength of STPA is to derive a comprehensive list of safety constraints. Those interested in the risk analysis process for STPA may refer to the paper by Kim et al.³¹

Next, we utilized a recommendation by Kim et al.³² when modeling the system. They propose to include the power supply as part of the control action and include it for UCA identification. This may avoid the omission of essential hazards from the analysis. The power supply was modeled as a supply function with a green arrow in the HCS model.

Like FHA, we developed a rule to transform the type of UCA keywords into controller constraints keywords for controller constraint generation. The transformation rules are listed in Table 2.

Comparison analysis procedure

The descriptions of FHA and STPA procedures show that they have different methods and perspectives on analyzing hazards. However, the core objectives of each step are similar. For example, step 3 *identify hazards* of FHA and step 3 *identify UCAs* of STPA are processes to identify hazardous events (or hazards in STPA). Table 3 shows the mapping of both FHA and STPA procedures based on each step's core objectives. The listed terms for each step of FHA and STPA denote the different terms used by each method during the specific phase of the analysis. Table 3 also shows how the process of FHA (2a–6a) and STPA (2b–6b) are different.

The mapping of both method procedures allows comparing the case studies' results in each analysis step. The analysis is performed at a higher level to avoid the influence of technical discussion that may blur both method's characteristics and presented in separate discussions. Specific to the comparison of causal scenarios and safety requirements, we utilized the previously mentioned approaches to answer the RQ2 and RQ3.

Case study

The Åsgard subsea compression system in Norway³³ inspires the case study, where two protection systems (process control and safety) exist independently of each other. The integration of process control and safety concept is a novel technology applied as an alternative solution to reduce the complexity of the physical architecture.^{9,34} This concept is part of the use case in the Safety 4.0 project, where the goal is to develop a standardized safety demonstration approach for novel subsea technologies.³⁴ This concept may increase software complexity, thus decreasing the confidence in its functional capability. This case study is deemed as sufficiently complex and relevant for use in our study.

System description

The system process flow diagram is illustrated in Figure 1. Redundant equipment and utility systems (e.g. network switches) are not illustrated in the Figure 1 for simplification. The subsea compression system consists of a scrubber, a compressor, and a pump. The system's goal is to ensure high gas flows and recovery rates from the well. The liquid mixture is recovered from the well and goes to the scrubber for separation. The dry gas is then compressed in a compressor, while a pump pumps the separated liquid. Both the dry gas and the liquid are then delivered to the topside facility for further processing. The study focuses specifically on the control and safety mechanism in the pump. A high voltage electronic power unit is used to power the pump operation. Here, the Process Control System (PCS) is utilized to maintain the level of liquid inside the scrubber by changing the pump's speed. If the liquid level gets too low, the gas can go through the pump (gas blow-by) and cause overpressure downstream.³⁵ The Process Shut Down system (PSD) is implemented to increase the pump protection system's integrity by shutting

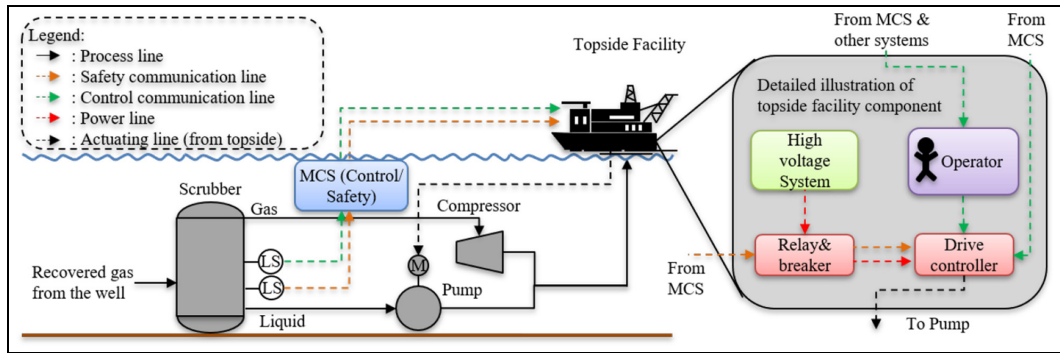


Figure 1. Simplified process flow diagram of subsea gas separation and compression to topside facility with the communication lines for PCS and PSD.

down the pump in case of the low-low (a technical term to describe the low limit for PSD) level detection in the scrubber.

The PCS loop consists of level sensors, Master Control Station (MCS), operator, PCS node, driver controller, and other systems. The level sensor detects the deviation of process condition and sends the signal to MCS for automatic logic solver response. Information from the MCS is also provided to the operator to see whether manual intervention is required. Depending on the control loop mode (automatic or manual), the PCS node needs to select the prioritized response (from either the MCS or the operator command) to the driver controller for regulating (increase or decrease) the pump speed.

The PSD loop consists of level sensors, MCS, PSD node, relay and breaker, operator, and other systems. The level sensor detects whether abnormal condition occurs in the system and informs the MCS for automatic logic solving response. During an abnormal condition, MCS needs to automatically shut down the equipment by passing information through the PSD node to relay and breaker to stop the pump's power supply. It is also possible to receive shutdown command from other systems in case of emergency. In this case, the operator is responsible for shutting down the power supply directly.

In this system, a physical integration with a logical separation concept⁹ is implemented at the Master Control Station. It means that the PCS and PSD share the same hardware while separated logically in the software architecture. They are designed to work parallel to each other, with the safety system has higher priority over the process control system when utilizing the same hardware resources.

Results

FHA results

The functions of the described system were modeled in the FAST diagram, as illustrated in Figure 2. The top

path describes the pathway for activation of safety function while the bottom path describes the pathway for activation of process control function. Each function's operational mode was specified based on the output of the targeted function's preceding function and condition. For example, the operational mode of *aut. command pump shutdown* function was the output of *detect abnormal level* (i.e. normal, low, or low-low) and the condition of *detect pump status* (i.e. running, unknown, or stopped). The complete functional list and operational mode are listed in Table 4.

Examples of the FHA results for step 3a–5a are presented in Table 5. The hazard identification process identified 64 hazardous events from 168 possible combinations (between functions, operational modes, and the failure mode list). Identification of the consequences showed that 21 HEs might result in Con1 *equipment damage*, 40 HEs might result in Con2 *unnecessary loss of production*, and three HEs might result in both types of losses. The causal analysis process identified 206 possible Causal Scenarios (CaS) associated with the 64 hazardous events (HE).

Safety Requirements (SR) are generated for the functions based on the identified HEs and CaSs. Sixty-four SRs corresponded one to one to the identified HEs. The identified CaSs were included in the SRs as guidance during the formulation of prevention/mitigation solutions. Examples of the SRs based on the HEs listed in Table 5 are (the SR format is *SRId. SR [CaSid]*. SRId and CaSid refer to the numbering of the SR and the related CaS):

- SR001. Stop pump function must be provided within the required time when there is shutdown command, and the pump status is running/unknown [CaS001].
- SR015. Aut. command pump shutdown function must not be provided when scrubber level status is normal, and the pump status is running/unknown [CaS050–056].

Table 4. Functional list of the pump protection system for FAST diagram & HCS.

FAST ID	Function	Operational mode/process model (condition)	HCS ID	Function type	Agent	Target
Fun02	Stop pump	Shutdown command (yes/no)	C01	Control	Pump motor	Pump
Fun03	Stop motor power supply	Pump status (run/stop/unknown) Shutdown command (yes/no)	C02	Control	Driver controller	Pump motor
Fun04	Aut. open circuit	Pump status (run/stop/unknown) Shutdown command (yes/no)	C03	Control	Relay and breaker	High voltage system
Fun05	Proceed shutdown command	Pump status (run/stop/unknown) Shutdown command (yes/no)	C04	Control	PSD node	Relay and breaker
Fun06	Aut. command pump shutdown	Scrubber level status (PSD) (normal/low/low-low)	C05	Control	MCS	PSD node
Fun07, Fun20	Provide scrubber level status (PSD)	Pump status (run/stop/unknown)	F06	Feedback	Sensor PSD	MCS
Fun20	Provide process information	Process condition (normal/abnormal)	F07	Feedback	MCS	PSD node
Fun20	Provide process information	Process condition (normal/abnormal)	F08	Feedback	PSD node	Operator
Fun08	Man. open circuit	Shutdown command (yes/no)	C09	Control	High voltage system	Relay and breaker
Fun09	Man. command pump shutdown	Pump status (run/stop/unknown) Shutdown command (yes/no)	C10	Control	Operator	High voltage system
Fun10, Fun20	ESD/PSD command	Pump status (run/stop/unknown)	F11	Feedback	Other systems	Operator
Fun11	Regulate pump power output	Alternate power command (yes/no)	C12	Control	Pump motor	Pump
Fun12	Alternate motor power supply	Pump status (run/stop/unknown)	C13	Control	Driver controller	Pump motor
Fun13	Command change pump output	Alternate power command (yes/no) Command priority result (act/no act)	C14	Control	PCS node	Driver controller
Fun14	Check command priority	Pump status (run/stop/unknown) Human command (yes/no)	C15	Control	PCS node	PCS node
Fun15	Aut. pump output change command	MCS command (yes/no) Priority status (MCS/human) Scrubber level status (PCS) (normal/low/low-low)	C16	Control	MCS	PCS node
Fun16	Man. pump output change command	Pump status (Run/Stop) Scrubber level status (PCS) (normal/low/low-low)	C17	Control	Operator	PCS node
Fun17, Fun20	Provide scrubber level status (PCS)	Pump status (run/stop)	F18	Feedback	Sensor PCS	MCS
Fun20	Provide process information	Process condition (normal/abnormal)	F19	Feedback	MCS	PCS node
Fun20	Provide process information	Process condition (normal/abnormal)	F20	Feedback	PCS node	Operator
Fun18, Fun20	Provide pump motor status	Process condition (normal/abnormal)	F21	Feedback	Pump motor	MCS
Fun18, Fun20	Provide relay status	Process condition (normal/abnormal)	F22	Feedback	Relay & breaker	PSD node
Fun18, Fun20	Provide driver controller status	Process condition (normal/abnormal)	F23	Feedback	Driver controller	PCS node
Fun19	Supply Power	Process condition (normal/abnormal)	S24	Supply	High voltage system	Relay and breaker
Fun19	Maintain power supply	Process condition (normal/abnormal)	S25	Supply	Relay and breaker	Driver controller

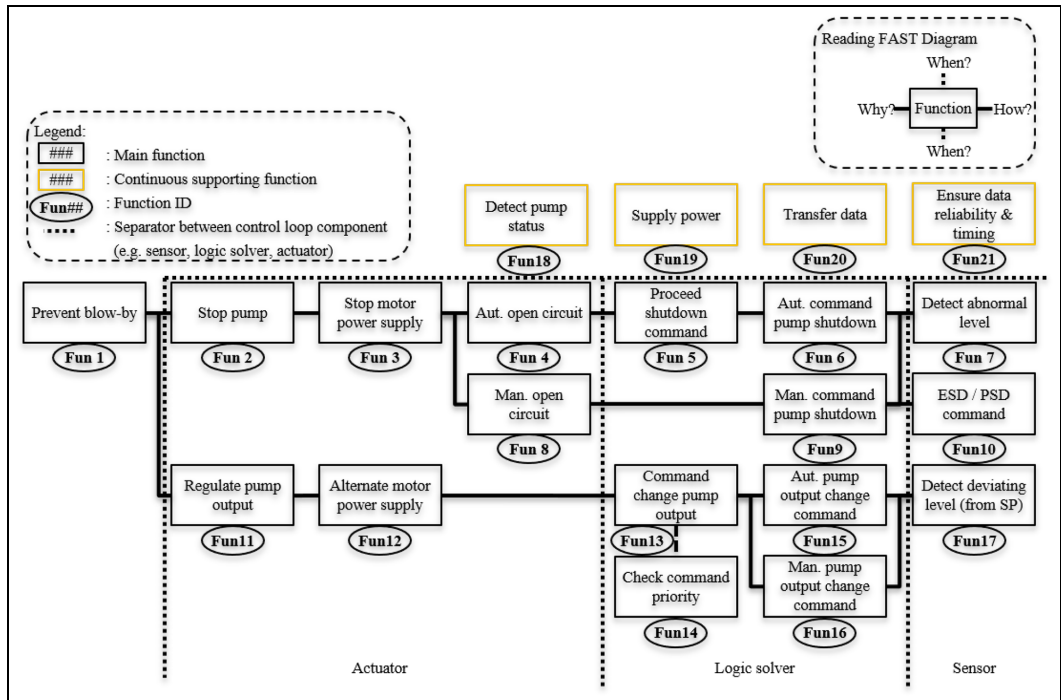


Figure 2. FAST diagram of pump protection system.

Table 6. System-level losses, hazards, and safety constraints identified on STPA.

L tag	System-level loss (L)	H tag	System-level hazard (H)	SC tag	System-level safety constraint (SC)
L1	Equipment damage	H1	Equipment operates outside normal operating condition	SC1	Equipment must be protected from extreme operating conditions that can result into damage
L2	Unnecessary loss of production	H2	Equipment operates outside optimal operating condition	SC2	Equipment must be operated within optimal operating conditions
		H3	Unintended stop of equipment when needed	SC3	equipment must be available to work when needed

- SR043. Command change pump output function must be provided correctly when the priority check result is to change pump output, and the pump status is running [CaS125–126].
- SR048. Aut. pump output change command function must be provided when scrubber level status is low, and the pump status is running/unknown [CaS142–148].

UCAs were identified from the combination of control actions, process models, and UCA types. In total, out of 134 identified combinations, 56 were classified as UCAs. Fifteen UCAs might result in H1, 32 UCAs in H2, and 9 UCAs in H3. Table 6 shows that H1 corresponds to L1 (15 UCAs), while both H2 and H3 correspond to L2 (41 UCAs combined). Examples of identified UCAs are (the UCA format is *UCAId. UCA [Hid]*). UCAId and Hid refer to the numbering of UCA and H):

STPA results

The boundaries of STPA analysis were the System-level losses, System-level hazards (H), and System-level safety constraints, as listed in Table 6. The equipment protection system was modeled as an HCS in Figure 3. The complete list of functions, associated agents, function types, and process models are listed in Table 4.

- UCA001. Pump motor provides stop pump command to the pump too late when there is a shutdown command, and the pump status is running/unknown [H1].
- UCA015. MCS provides Aut. command pump shutdown to the PSD node when Scrubber level

status is normal and the pump status is running/unknown [H3].

- UCA026. Pump motor stops providing regulate pump output to the pump too soon before the condition there is a command to change pump output, and the pump is running changes [H2].
- UCA027. Pump motor provides regulate pump output to the pump too long after the condition, there is a command to change pump output, and the pump is running changes [H2].
- UCA044. MCS does not provide Aut. pump output change command to the PCS node when scrubber level status is low, and the pump status is running/unknown [H2].

The control loops associated with every UCA were analyzed further to identify the Loss Scenario (LSc). There are 346 identified LScs. Examples of the LScs are (The format of LSc is *UCAId.LScId.LSc*. UCAId and LScId refer to the numbering of UCA and LSc. UCAId.LScId shows the link between every LSc to the associated UCA):

- UCA001.LSc001. Local battery as spare power prevents an automatic shutdown of the pump.
- UCA015.LSc093. Problem in the control path caused by unreliable data from topside communication.
- UCA015.LSc094. Problem in the control path information caused by topside communication failure.
- UCA015.LSc095. Problem in the received information caused by unreliable data from subsea communication.
- UCA015.LSc096. Problem in the received information caused by subsea communication failure.
- UCA015.LSc097. Problem in the controlled process due to PSD node hardware failure.
- UCA015.LSc098. Problem in the controlled process due to PSD node software error.
- UCA015.LSc099. Problem in the controller due to MCS hardware failure.
- UCA015.LSc100. Problem in the controller due to MCS (safety) software error.
- UCA015.LSc101. Problem in the controller due to unintended interaction between PCS and SIS that cause software error.
- UCA015.LSc102. Problem in the received information due to level sensor (safety) hardware failure.
- UCA015.LSc103. Problem in the received information due to level sensor (safety) software error.

CCs are generated based on the transformation rule to the identified UCAs. Fifty-six CCs correspond one to one to the identified UCAs. The identified LScs are listed to show the possible scenarios, possibly affecting the fulfillment of the constraint. Examples of the generated CCs are (the CC format is *CCId.CC [LScId]*,

CCId and LScId refer to the numbering of CC and the related LSc):

- CC001. Pump motor must provide stop pump to the pump within the required time when there is shutdown command, and the pump status is running/unknown [LSc001].
- CC015. MCS must not provide aut. command pump shutdown to the PSD node when scrubber level status is normal, and the pump status is running/unknown [LSc093–103].
- CC026. Pump motor must provide regulate pump output to the pump continuously as required when there is a command to change pump output, and the pump status is running [LSc149–152].
- CC027. Pump motor must not stop providing regulate pump output to the pump before the condition there is a command to change pump output, and the pump status is running changes [LSc153–157].
- CC044. MCS must provide aut. pump output change command to the PCS node when scrubber level status is low, and the pump status is running/unknown [LSc252–262].

Discussion

The following sections contain discussions of the comparison results from the case study.

Comparison of the modeling techniques

Both analyses utilized a model to assist hazard identification, consequence identification, and causal scenario analysis processes. FHA and STPA utilized different models, the FAST diagram for the former and HCS for the later. Three properties distinguish the two models: model type, function type, and process flow.

The FAST diagram is a model of sequential functions, while HCS is a control structure model. In the FAST diagram, as seen in Figure 2, the focus is to depict how each function interacts with other functions in a structured and sequential manner to achieve the desired function. It is unknown which agent (system or subsystem) performs each function. Also, the interactions between the system with the environment are not modeled. Comparatively, HCS modeled the conceptual system operation as a structure of control loops. Every function (e.g. control action, feedback, or supply) has a subject (performing the function) and an object (the target of the function). For example, Figure 3 shows that *C03. aut. open circuit control action* is performed by *relay and breaker* (subject) to the *high voltage system* (object). Due to the association of function to subject and object, it is possible to have several agents performing the same function. For example, *high voltage system* and *relay and breaker* has responsibility to maintain power supply function (represented as *S24 supply power* and *S25 maintain power supply*). These comparable functions are only modeled as a single function *Fun18*

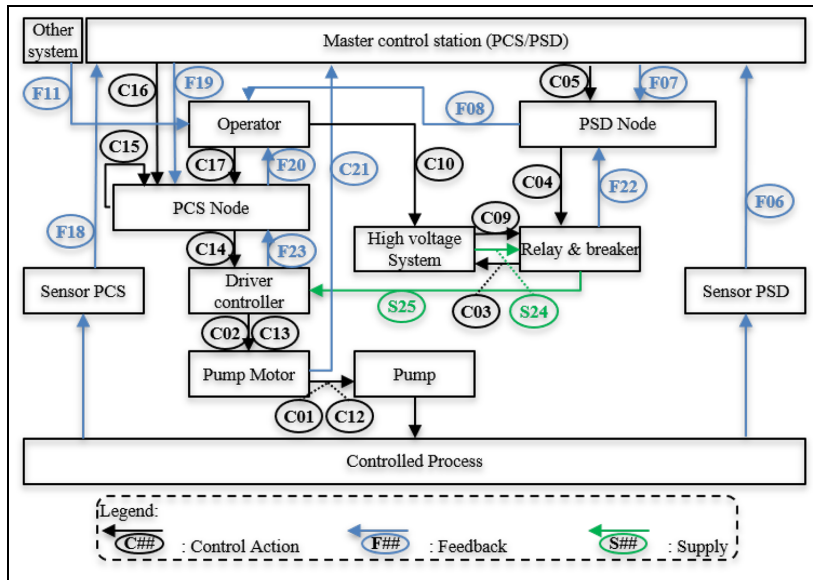


Figure 3. HCS of pump protection system.

Table 7. Differences in function classification between FAST diagram and HCS.

Modeling differences	FAST function type	FAST ID	HCS function type	HCS ID
Type 1	Main	Fun02 Fun06, Fun08, Fun09, Fun11 Fun13, Fun15, Fun16	Control action	C01 C05, C09, C10,
Type 2	Main	Fun07, Fun10, Fun17	Feedback	F06, F11, F18
Type 3	Supporting	Fun14	Control action	C15
Type 4	Continuous	Fun19	Supply	S24, S25
Type 5	Continuous	Fun18, Fun20, Fun21	Feedback	F07, F08, F19, F20 F23

supply power in the FAST diagram. In HCS, it is possible to model the influence from the environment (anything outside the system boundary) to the system in the HCS by modeling it as a box performing a function to the agent.

The FAST diagram and the HCS classified the functions into different types. In the FAST diagram, each function is classified either as a main, a supporting, or a continuous function. In the HCS, each function is classified either as a control action, a feedback, or a supply function. Since the analyzed system is the same, it is possible to map every function’s classification between the FAST diagram and the HCS. The summary of the mapping is listed in Table 7. For example, the function *stop pump* is classified as the main function *Fun02* in the FAST diagram and as control action *C01* in the HCS (type 1). In another example, the function *detect abnormal level* is a main function *Fun07* in the FAST diagram and is a feedback *F06* in the HCS (type 2). This mapping is unique for this equipment protection system and may be different depending on the investigated system.

For the process flow, it is clear how every function’s sequential process is modeled in the FAST diagram. The horizontal (left-right) sequence shows how the function to the right of the selected function is the causative function, while the function to the left is the reactive function. In contrast, HCS models the hierarchy in a vertical (top-down) relation. It depicts how one controller has higher authority than the agent (e.g. another controller or controlled process) at the lower hierarchy level. This vertical hierarchy does not show the system operational process (i.e. the starting, the preceding, the following, and the finishing point).

To understand the HCS (i.e. in Figure 3), it is necessary to read the controller responsibility and process model at any given point (e.g. in Table 4). For example, *MCS’s* responsibility as the controller is to provide *C05 aut. command pump shutdown* to the *PSD node*. From the HCS, *PSD node* has two output pathways, *C04 proceed shutdown command*, or *F08 provide process information*. From the Table 4, *C04* has a process model *shutdown command status* that indicates *PSD*

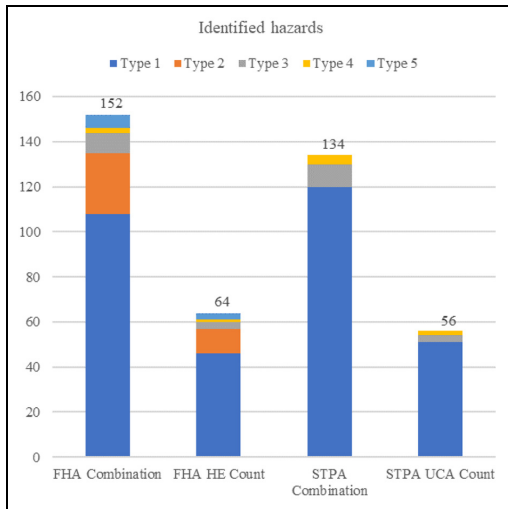


Figure 4. Comparison between the number of the assessed combination from FHA and STPA process and the identified HEs and UCAs (types refer to Table 7).

node responsibility to pass the shutdown command to the *relay and breaker*. In contrast, *F08* shows *PSD node* responsibility to provide feedback information to the *operator*. *F08* is not consistent with the control action *C05*. It is more logical to have *C04* as the following operational sequence after *C05*. This way of reasoning is necessary to gain an understanding of the system process from the HCS. Arguably, for a more complex controller (with a higher number of input/output functions), it would be more difficult to understand the step-by-step sequence of the function for people who never looked into the system before the analysis.

These three differences between the modeling of the FAST diagram and HCS may affect the latter hazard analysis process that will be discussed in the later section.

Comparison of the hazardous events and unsafe control actions

Figure 4 shows statistics of the identified HEs and UCAs from the pump protection system. It appears that FHA captured a higher number of HEs than STPA did with UCAs. It is due to three reasons: the use of keywords for hazard identification, the function type classification in the selected model, and the modeling approach.

The keywords comparison can be seen in Table 8. Overlapping keywords result in the identification of the same type of HEs and UCAs. For example, *HE001* (in Table 5) and *UCA001* (in section *STPA result*) are inherently the same type of hazardous events. However, for keywords such as stopped too soon, applied too

Table 8. Comparison between failure mode and UCA types.

Type of failure mode	Type of UCA
Omission error	Not provided (when needed)
Commission error	Provided (when not needed)
Late	Provided too late
Early	Provided too early
–	Stopped too soon
–	Applied too long
Value	–

long, and wrong value with no comparable guidewords in the other methods (the two former keywords for STPA and the following keywords for FHA), the hazards identified by utilizing these keywords were unique to the particular method. For example, STPA did not identify UCA similar to *HE043*, while FHA did not identify HE similar to *UCA026* and *UCA027*.

Second, as mentioned during modeling technique comparison, some functions are classified differently between the FAST diagram and HCS. While it does not affect FHA's hazard analysis, the classification affected the identification process of STPA. STPA considers hazardous events as unsafe control actions. It results in a limitation of the hazard identification process only to include the control action and the supply functions. Therefore, functions classified as type 2 and type 5 from Table 7 are analyzed for possible HEs in FHA, while not analyzed for possible UCAs in STPA. Figure 4 shows that there is no orange-colored box (type 2) and light blue-colored box (type 5) in both the STPA combination and the identified UCAs.

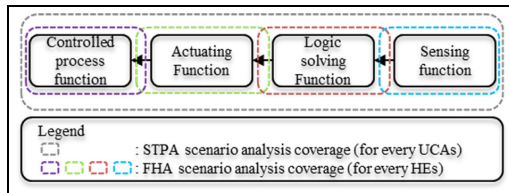
Finally, how the FAST diagram and HCS approached to model the system also contributed to the number of identified hazards. Some functions can be performed by several agents (i.e. one function in the FAST diagram can be two or more functions in HCS). This modeling approach increased the number of identified hazards in the STPA. For example, analysis of *S24 supply power* and *S25 maintain power supply* in STPA resulted in two UCAs, while analysis of *Fun18 supply power* in FHA resulted in one HE.

Comparison of the consequences and system-level losses

The identified consequences and System-level Losses for FHA and STPA are the same: equipment damage and unnecessary loss of production. However, they were derived differently. In FHA, consequences are assessed as a possible effect of HEs (inductive technique). Comparatively, the loss results in STPA were identified at the start as unwanted loss caused by a system-level hazard that needs to be avoided (obtained either from past experiences with a similar system or from standards and regulations). These Ls became the starting point for deductive analysis in STPA to

Table 9. Comparison between the number of analyzed scenarios.

Type	FHA count	STPA count
Hazard	64	56
Causal scenario	206	346
Caused by composition	108	247
Caused by environment	0	0
Caused by structure	87	80
Caused by mechanism	11	19

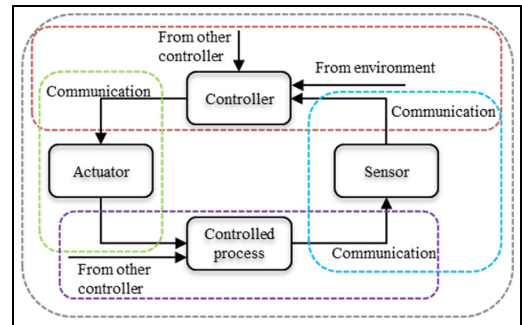
**Figure 5.** Example of loss scenario analysis perspective comparison based on the FAST diagram.

identify UCAs. This different perspective affects the analysis boundary.

The identified Losses in STPA limit the boundary of the analysis to the pre-described system-level hazards and losses. The focus of STPA was then to determine what type of possible UCAs can result in the pre-described Losses. That culminated in comprehensive top-down traceability from the Loss – Hazards – UCAs – LScs (shown in the inclusion of various IDs resulting from UCAs). Arguably, it is tough to have a complete list of unwanted Ls and Hs from the start of the analysis, especially if it is performed for novel technology. When encountering this problems, Leveson²⁶ recommended to start the analysis at a higher level of abstraction. This, however, caused the resulting list of Ls and Hs to be too generic and necessitates an iteration process to ensure completeness. The top-down method of STPA shows its limitation when there are omitted system-level losses or system-level hazards. In this condition, it is necessary to redo the analysis from the start to check whether there are any omitted UCAs or LScs from the analysis. This problem soon becomes unmanageable for a larger and complex system. In comparison, The FHA process is not limited by the identified consequences. When there is a change in the system, what needs to be done is to check whether the identified hazard's implication results in the same/different consequence.

Comparison of the causal scenarios and loss scenarios

Table 9 presents a statistic of the analyzed CaS and LSc by FHA and STPA for the pump protection system. It shows that in contrast to the higher number of

**Figure 6.** Example of loss scenario analysis perspective comparison based on the HCS.

identified HEs than UCAs, the number of identified LScs is significantly higher than the CaSs. This is caused by how the utilized model aids the causal scenario analysis process and the availability of other relevant information.

In FHA, all HEs are associated with a function. The analysis scope of the CaS from the FAST diagram is limited to the respective function and its immediate connection, as shown in Figure 5. Comparatively, the LScs are analyzed from their associated control loops that provide the UCAs. A control loop includes all necessary functions to perform the control action function (e.g. actuating, logic solving, and sensing function). Therefore, the analyzed control loop may include several agents that perform different functions, as shown in Figure 6. This results in a higher number of identified loss scenarios. For example, *HE015* and *UCA015* are the same type of functional hazard for a logic solver *MCS*. Causal scenario analysis of *HE015* identified seven distinct CaSs, while loss scenario analysis of *UCA015* identified eleven distinct LScs. FHA's causal scenario analysis process was only able to find scenarios related to the *MCS* (that perform logic solving function) and *topside* and *subsea communication* (that transfer the function from the previous function and to the following function). STPA's loss scenario analysis process managed to find additional unique scenarios related to the *PSD node* (that perform actuating function) and the (*safety*) *level sensor* (that performs sensing function).

As described in the procedure of FHA and STPA, both the FAST diagram and HCS models are utilized as the aid for the causal analysis process. The FAST diagram is a model that depicts how every mechanism in the system is connected structurally. It does not specify any components and how they interact with the environment. The analyst must identify the C and E properties (of the CESH model) from other information sources. First, each function is associated with the agent (or composition) performing it. Then conceptual design and operation are utilized to check whether there will be a process condition (or environmental effect) that may cause a hazard.

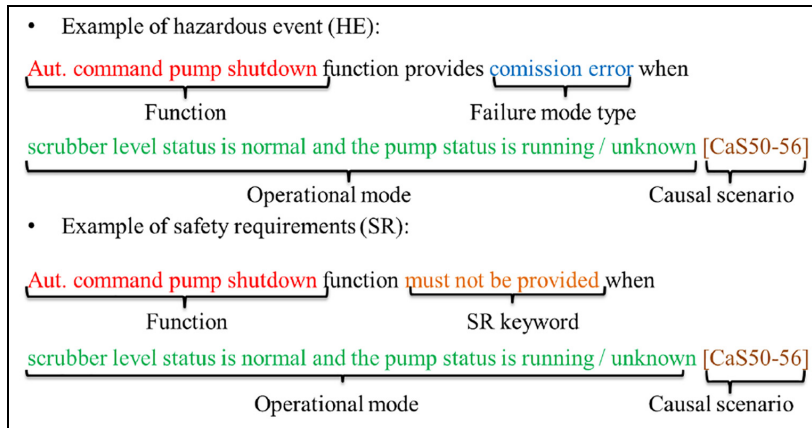


Figure 7. Example of key attributes tagging on the HEs and SRs (colors are used to distinguish between the key attributes).

In contrast, all aspects of CESM are modeled into the HCS (see Figure 6). An HCS depicts how the Agent (or composition) is connected structurally to each other by performing functions (or mechanism). Influence from the environment can be added to the model to consider the possible implications to the UCA. For example, Table 9 shows the classification of the identified causal scenarios based on the CESM properties. While both techniques cover all the CESM properties for the causal scenario analysis, the HCS provides more help due to the inclusion of all the model properties. It reduces the omission possibility when checking the causal scenarios from several information sources.

Comparison of the safety requirements and controller constraints

FHA derived 64 SRs, while STPA derived 56 CCs. These requirements/constraints are obtained solely from the identified hazards. The SRs and CCs are evaluated based on the eight criteria for requirement:²³ (1) identifiable, (2) clear, (3) solution-specific, (4) have ownership, (5) have origin, (6) verifiable, (7) able to be validated, and (8) have priority.

- (1) *Identifiable.* Both SRs and CCs are derived to ensure the safety of the system. They can be accounted for the hazard analysis process of FHA and STPA. If any changes arise in the system, the listed requirement may not be applicable anymore, depending on the implication of change to the analyzed system.
- (2) *Clear.* The derived SRs and CCs achieved this by utilizing the key attributes from the HEs and UCAs to word the requirements/constraints. The examples of the tagging from the key attributes to the generated SRs and CCs are shown in Figures 7 and 8. For example, an SR is identifiable by its composition of function, SR keyword, operational mode, and

causal scenario. Similarly, a CC is recognizable by its composition of the controller, CC keyword, control action, controlled process, process model, and loss scenario. The transformation rule from HEs to SRs and UCAs to CCs in Table 2 makes the derived SRs and CCs more evident.

- (3) *Solution-specific.* Both FHA and STPA are performed to analyze a specific system. The derived SRs and CCs are only applicable, given the context and scope of the analysis initially defined. Like the first criterion, the SRs and CCs may not be applicable anymore if any changes occur.
- (4) *Have ownership.* In the context of systemic hazard analysis of functions, it is best performed during the early design phase. Both SRs and CCs need to be followed by the designer to develop the system's detailed design. The stakeholder may change if the hazard analysis is performed at the different phases of design.
- (5) *Have origin.* In FHA, the SRs subject is the function itself. It does not specify which component (system or subsystem) needs to follow the requirement. It allows the decision-maker to assign any agent that needs to carry out the functions. In STPA, the CCs subject is a specific controller (see Figure 8) that needs to be constrained. If the control action is assigned to a different agent, the initial requirement does not apply anymore. Another STPA process needs to be done to check whether additional CCs are required for the new agent.
- (6) *Verifiable.* If the requirement is too detailed and technical, the requirements may not be satisfied by the new technology and limit the options for solutions. Both SRs and CCs are functional requirements. They do not limit the possible solution as long as it is possible to achieve the required functionality. The causal/loss scenarios can be used as guidance to satisfy the requirements (e.g. by identifying the barrier to eliminate,

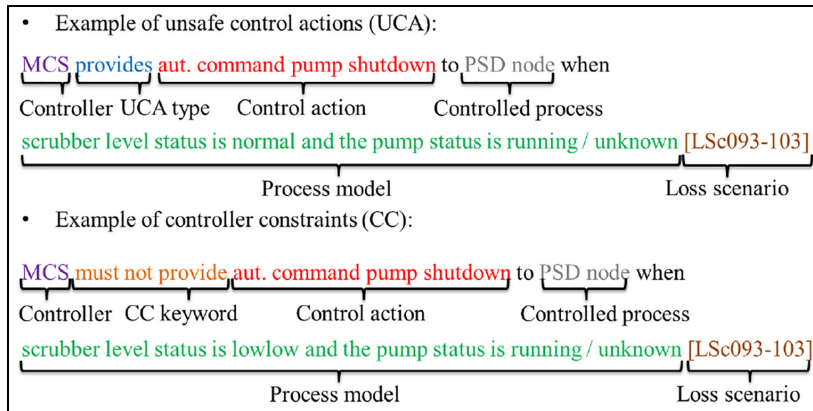


Figure 8. Example of key attributes tagging on the UCAs and CCs (colors are used to distinguish between the key attributes).

prevent, or mitigate the scenarios). Both SRs and CCs also need to be checked against the system's original functional requirement. There may be conflicting requirements due to the different perspectives in the initial requirement (e.g. between achieving safety of the system or availability of the production).

- (7) *Able to be validated.* In this paper, the SRs and CCs are qualitative requirements. It is difficult to justify whether the derived requirements can be achieved or not given the current form of the SRs and CCs (without measurable criteria). FHA is originally a semi-quantitative hazard analysis tool. Typically, a risk assessment process is integrated into the FHA process (see section *FHA procedure*) to validate the requirements (e.g. by quantifying the effects of risk reductions and checking them against the risk criteria). In contrast, STPA is originally not supported by quantitative measures due to Leveson's skepticism with the individual number assignment (e.g. for likelihood assessment)¹. Only recently that Kim et al.³¹ proposed a semi-quantitative approach for risk analysis with STPA.
- (8) *Have priority.* In this paper, we do not perform any prioritization of the safety requirement. As discussed previously, the semi-quantitative measures are also used to prioritize essential requirements in both FHA and STPA (although still need further research for the latter method).

simultaneously since they capture similar types of functional hazards and scenarios. From the analysis to answer the RQ2 and RQ3, STPA has two advantages over FHA: the modeling technique captures all four systemic properties, and the safety requirements structure complies with more criteria of a requirement.

If looking into the system model, the STPA's modeling technique captures all four systemic properties of a system, while FHA's modeling technique can only capture two systemic properties. This makes the causal analysis process of STPA easier than FHA due to the latter's need to refer to other documents/models for support. From the criteria of a requirement, we identify that every safety requirement in STPA has been assigned to an agent (e.g. physical component or human). This makes the safety requirement of STPA ready for use, while an additional process is required in FHA to identify the agent.

Based on the reasons above, we conclude that STPA is more suitable than FHA to analyze novel technology. Due to their focus on functionality, rather than the realization, both methods are theoretically general enough to be used across different application areas. Our recommendation is valid in the process industry, as demonstrated in this study, and in the aerospace industry,¹⁹ where FHA is the recommended methods.⁴ STPA demonstration in other industrial applications, for example, medical,^{29,36} and maritime,³⁷ indicates its versatility across different subject areas and implies that our recommendation can be relevant as well.

Conclusion of the comparison analysis

Table 10 provides a summary of the comparison results. To assess the implication, we need to bring the results one step higher and reflect on the analysis to answer the RQs and our initial objective. Based on the analysis to answer the RQ1, the comparison indicates that both methods are similarly suitable for analyzing novel technology. It is unnecessary to utilize FHA and STPA

Overall implications

This section discusses the implications of the findings with several topics in the risk management area.

Insights into hazard analysis methods

The comparison analysis highlights the differences between FHA and STPA procedures that can be used

Table 10. Summary of the comparison analysis.

FHA vs. STPA steps	FHA	STPA
FAST diagram vs. HCS	<ul style="list-style-type: none"> Model system as structured functional diagram No information on the subject that performs the function (system or subsystem) No information on interaction with the environment Function is classified as main, support and continuous Function sequence is clear 	<ul style="list-style-type: none"> Model system as a structured control loop Every function has a subject (that perform the function) and object (target of the function) Possible to model effect on the environment Function is classified as control, feedback, and supply Control sequence is ambiguous One function can be performed by several agents (e.g. for pass-through of function)
Hazardous Events (HE) vs. Unsafe Control Actions (UCA)	<ul style="list-style-type: none"> Type of failure mode: omission error; commission error; late, early, value Analyze every function as possible HEs When analyzing comparable functions and utilizing similar keywords, both methods find the same type of hazards Identified as result of HEs (bottom-up approach) 	<ul style="list-style-type: none"> Type of UCA: not provided (when needed) provided (when not needed), provided too late, provided too early, stopped too soon, applied too long Does not analyze feedback function as possible UCAs (not a control action). When analyzing comparable functions and utilizing similar keywords, both methods find the same type of hazards Need to be defined at first (top-down approach) Boundary of the analysis
Consequence vs. system-level loss Causal Scenario (CaS) vs. Loss Scenario (LSc)	<ul style="list-style-type: none"> The analyzed scenarios are considered from the hazardous function and its interaction with connecting function Need additional information sources to assess the compositional (e.g. component failure) and environmental problem (e.g. pressure influence) 	<ul style="list-style-type: none"> The analyzed scenarios are considered from every function in the control loop The compositional, environment, structure and mechanism properties of a system is included into single HCS model
Safety Requirement (SR) vs. Controller Constraint (CC)	<ul style="list-style-type: none"> Does not have an agent as the subject for a requirement Has established procedure to apply criteria for validation and prioritization 	<ul style="list-style-type: none"> Have an agent as the subject for a requirement Originally, do not include criteria for validation and prioritization. Recent works indicate the attempt for prioritization.

as lessons learned to improve both methods. For example, we found several unique hazards to FHA and STPA due to the different keywords used by each method. FHA and STPA may increase hazards coverage by borrowing the missing keywords (refer to Table 8) and used them for the identification of hazardous events or UCAs.

In STPA, the feedback functions are not considered for UCAs' identification, which results in a lower number of the hazards. Error in the feedback functions (e.g. detection error) are later identified as possible scenarios that lead to the UCA (see discussion on causal scenario comparison). Therefore, there is a lower risk of omission by not considering the feedback functions as UCA. It is not necessary to modify the STPA procedure based on this issue.

For the modeling technique, HCS captures more systemic properties of CSM than the FAST diagram. In FHA, this is complemented by analyzing the remaining properties from other information sources. Utilizing the HCS model (or similar model that captures CSM properties in a single model) during the FHA's causal analysis process would reduce the omission possibility of relevant scenarios.

The analysis using STPA in our study case produces a significantly higher number of causal scenarios than FHA. Most of the scenarios found by STPA are caused by similar causal factors that are redundant with FHA causal factors. When analyzing the type of causal factors in scenarios, we found that both methods still suffer the same limitation for identifying either scenario due to a single point causal factor (e.g. component failure or software error) or known scenario (due to simple interaction). This is contrary to Leveson et al.¹⁹ findings that analysis using STPA could find causal scenarios that could not be found by analysis using FHA. Currently, both hazard analysis methods still rely heavily on expert judgment and historical experiences. For novel technology involving complex software-intensive systems, experts and experiences' advantages are lower (due to limited information). Having a systemic perspective when analyzing the causal scenarios does not imply capturing systemic causal scenarios. We conclude that the systemic model used by STPA does not have an advantage over the complex linear model used by FHA. Therefore, a procedure to analyze systemic scenarios caused by multiple point problems and unknown scenarios is required.

Insights into the cautionary principle

Both FHA and STPA generate qualitative requirements with equal weight for all the identified requirements. This is in line with the cautionary principle that *if the consequences of an activity could be serious or subject to uncertainties, then cautionary measures should be taken and, or the activity should not be carried out.*^{38,39} However, according to Kim et al.,³¹ equal weight does not provide decision-making support. For example,

our case studies with FHA and STPA identified many scenarios for a small system with only 12 components. Without prioritization, the decision-maker would not be able to select the most critical requirements as resources and time for implementation are limited.

When the application is safety-critical, risk and reliability requirements are applied to the functions (not treated in the paper). The requirements (expressed as safety integrity level requirements) stem from as low as reasonably practicable principle. Some functions may not be implemented with the same integrity (or having different priority levels). If the requirements have low priorities (e.g. having minimal consequences, or less uncertainty), they can be assumed to have an insignificant impact on the system. Thus, not conflicting with the cautionary principle. While risk analysis is already an established practice in FHA, the experience with risk analysis on STPA is low. Kim et al.³¹ approach has a risk of screening out essential scenarios. Further research is vital to improve and validate the latter approach with other study cases.

Insights into the uncertainty management

Assumptions used during the hazard analysis process imply that the result's validity may have uncertainties. For example, during the modeling process, the analyst would have an initial preconception of the system behavior. Karanikas⁴⁰ presented at least ten types of assumptions during each step of the STPA procedure. Similarly, assumptions are used when performing FHA for our case study.

We found that Both FHA and STPA do not guide on communicating both assumptions and their uncertainties sufficiently. The results of FHA and STPA only covers what were considered as hazardous scenarios and did not record what was not found (e.g. safe scenarios). The omission of such results may be dangerous as the latter scenarios' assumptions may deviate (due to uncertainties) and result in the need to consider such scenarios as hazardous. Bjerga et al.⁴¹ recommended performing a separate assessment to analyze the implication of uncertainty in the assumption in their work for treating uncertainty in risk analysis of a complex system. The recommended methods, the assumption-deviation risk,⁴² may also be useful for the management of uncertainties in both FHA and STPA. Here, the strength of knowledge in every assumption is assessed for the risk of deviation. Suppose the deviations have a significant impact on the analysis results, an additional study may be performed on the assumptions to increase the strength of knowledge and minimize the deviation effect.

Insights into the resilience management

Resilience refers to the system's ability to react and recover from disturbances.⁴³ In the context of safety, we focus only on the disturbances that may cause losses. The safety requirements of FHA and STPA are

limited on the prevention of hazardous events (i.e. react part of resilience). We found that there is a lack of attention on managing the resilience if the hazardous events (leading into losses) (i.e. react part of resilience) and the consequences do occur (i.e. recover part of resilience). Even if the safety requirements of FHA or STPA are fulfilled, there is no guarantee that this would result in a perfectly safe system. Thus, management on the missing part mentioned above would be necessary to increase the system's safety.

For the hazardous events leading to the losses part, we may learn from the process to generate the safety requirements in FHA and STPA. We suggest two-step procedures. The first is to add safety requirements to prevent or mitigate the losses. For example, in FHA, the requirements may be formed as follows, *hazardous event #xx should not lead to consequence #xx*. Similar requirements can be formed for STPA by changing the hazardous event with UCA. Then, FHA or STPA may be coupled with consequence assessment methods (e.g. cause-consequence diagram⁵ or event tree analysis^{5,6}). The later step is similar to the causal analysis procedure, where the results are attached to the requirements as guidance for scenarios that need to be prevented.

For the consequences part, we suggest forming the requirements to recover from the consequences. For example, *system should be able to recover from consequence #xx*. The requirement allows the decision-maker to formulate the recovery approach specific for each system condition.

Conclusion

This paper has carried out a systematic comparison of FHA and STPA with a case study from an equipment protection system in the oil and gas industry. We compared each step of the analysis individually (i.e. during system modeling, hazard identification, consequence identification, causal scenario analysis, and safety requirement generation). We have analyzed how each process is beneficial to identify functional hazards, provide a systemic perspective of the system, and generate safety requirements. This study found that STPA is more suitable than FHA to analyze the investigated system due to the advantages of the modeling technique used and the format of safety requirements that it generated. Recommendations are provided to improve FHA and STPA based on the lessons learned from each method. If the recommendations are implemented, FHA and STPA may have a similar level of capability and may replace each other. Obvious that this finding differs with other research claim where STPA is shown to be significantly better than the compared methods¹⁹ or is required as supplementary methods.²⁰ Further works by investigating other system with different functionalities and complexities are required to verify our claim.

The selection of STPA (or FHA as an alternative) requires further works to be aligned with the functional safety standard, that is, IEC 61508. We have discussed our insights related to the cautionary principle, uncertainty management, and resilience management as guidance for further works. First, the risk assessment process for prioritization of STPA results needs to be validated. Second, there is a need to investigate the assumption-deviation risk method to manage uncertainty in the FHA or STPA results. Another future work is to test and validate the suggested procedures for improving the resilience management in FHA and STPA.

Acknowledgements

This paper has been written under the Safety 4.0 project. The author(s) would like to thank the Research Council of Norway, industrial and university partners involved in this project for the support. For more information about the project, see <https://www.dnvgl.com/research/oil-gas/safety40/index.html>. The author(s) are grateful for the two anonymous reviewers' comments and insights into the earlier version of the paper.

Declaration of conflicting interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Research Council of Norway has partly funded this work as part of the Petromaks 2 program [grant number 281877/E30].

ORCID iDs

Nanda Anugrah Zikrullah  <https://orcid.org/0000-0002-9922-9045>

Hyungju Kim  <https://orcid.org/0000-0003-1829-3369>

Mary Ann Lundteigen  <https://orcid.org/0000-0002-9045-6815>

References

1. Leveson N. *Engineering a safer world: systems thinking applied to safety*. Cambridge, MA: The MIT Press, 2011.
2. Hollnagel E. *FRAM, the functional resonance analysis method: modelling complex socio-technical systems*. Surrey: Ashgate, Publishing, Ltd., 2012.
3. IEC 61508. *Functional safety of electrical/electronic/programmable electronic safety-related systems – part 1–7*. Standard, International Electrotechnical Commission, 2010.

4. ARP 4761. *Guidelines and methods for conducting the safety assessment process on airborne systems and equipments*. Standard, SAE. Warrendale: The Engineering Society for Advancing Mobility Land Sea Air and Space, 1996.
5. Ericson CA. *Hazard analysis techniques for system safety*. 2nd ed. Hoboken, NJ: Wiley, 2016.
6. Rausand M. *Risk assessment: theory, methods, and applications*. *Statistics in practice*. Hoboken, NJ: Wiley, 2011.
7. Alexander IF and Maiden N. *Scenarios, stories, use cases: through the systems development life-cycle*. Chichester: John Wiley & Sons, 2005.
8. Wilkinson P and Kelly T. Functional hazard analysis for highly integrated aerospace systems. In: *IET conference proceedings*, London, 17 February 1998. IEEE.
9. Zikrullah NA, van der Meulen MJP, Kim H, et al. Clarifying implementation of safe design principles in IEC 61508: Challenges of novel subsea technology development. In: *Proceedings of the 29th European safety and reliability conference (ESREL)*, pp.2928–2936. Singapore: Research Publishing.
10. Raspotnig C and Opdahl A. Comparing risk identification techniques for safety and security requirements. *J Syst Softw* 2013; 86(4): 1124–1151.
11. Raheja D. Software FMEA: A missing link in design for robustness. Technical Report, SAE, The Engineering Society for Advancing Mobility Land Sea Air and Space, 2005.
12. ISO 26262, Road vehicles — Functional safety – part 1–12. Standard, International Organization for Standardization, 2016.
13. IEC 61511, Functional safety — Safety instrumented systems for the process industry sector – part 1–4. Standard, International Electrotechnical Commission, 2016.
14. Ishimatsu T, Leveson NG, Thomas JP, et al. Hazard analysis of complex spacecraft using systems-theoretic process analysis. *J Spacecraft Rockets* 2014; 51(2): 509–522.
15. Yousefi A, Rodriguez Hernandez M and Lopez Peña V. Systemic accident analysis models: a comparison study between AcciMap, FRAM, and STAMP. *Process Safety Progr* 2019; 38(2): e12002.
16. de Carvalho EA, Gomes JO, Jatobá A, et al. Employing resilience engineering in eliciting software requirements for complex systems: experiments with the functional resonance analysis method (fram). *Cogn Technol Work* 2020: 1–19.
17. ISO/PAS 21448, Road vehicles—safety of the intended functionality. Standard, International Organization for Standardization, 2019.
18. Pumfrey DJ. *The principled design of computer system safety analyses*. PhD Thesis, University of York, 1999.
19. Leveson N, Wilkinson C, Fleming C, et al. A comparison of STPA and the ARP 4761 safety assessment process. Technical Report, MIT PSAS, 2014.
20. Sulaman SM, Beer A, Felderer M, et al. Comparison of the FMEA and STPA safety analysis methods—a case study. *Softw Qual J* 2019; 27(1): 349–387.
21. Bunge M. *Emergence and convergence: qualitative novelty and the unity of knowledge*. Toronto: University of Toronto Press, 2003.
22. Wan PYZ. Emergence à la systems theory: epistemological totalausschluss or ontological novelty? *Philos Social Sci* 2011; 41(2): 178–210.
23. Holt J, Perry SA and Brownsword M. *Model-based requirements engineering*. London: The Institution of Engineering and Technology, 2011.
24. Allenby K and Kelly T. Deriving safety requirements using scenarios. In: *Proceedings fifth IEEE international symposium on requirements engineering*, Totonto, ON, 27–31 August 2001, pp.228–235. IEEE.
25. ASTM E2013-12, Standard practice for constructing FAST diagrams and performing function analysis during value analysis study. Standard, ASTM International, 2012.
26. Leveson N and Thomas J. *STPA handbook*. 2018.
27. Thomas IV JP. *Extending and automating a systems-theoretic hazard analysis for requirements generation and analysis*. PhD Thesis, Massachusetts Institute of Technology, 2013.
28. Rachman A and Ratnayake RC. Implementation of system-based hazard analysis on physical safety barrier: a case study in subsea HIPPS. In: *2015 IEEE international conference on industrial engineering and engineering management (IEEM)*, Singapore, 6–9 December 2015, pp.11–15. IEEE.
29. Masci P, Zhang Y, Jones P, et al. A hazard analysis method for systematic identification of safety requirements for user interface software in medical devices. In: *Lecture notes in computer science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10469, pp.284–299. Springer Verlag.
30. Kim H, Lundteigen MA, Hafver A, et al. Application of systems-theoretic process analysis to a subsea gas compression system. In: *Safety and reliability – safe societies in a changing world – proceedings of the 28th international European safety and reliability conference, ESREL 2018*. pp.1467–1476. CRC Press/Balkema.
31. Kim H, Lundteigen MA, Hafver A, et al. Utilization of risk priority number to systems-theoretic process analysis: a practical solution to manage a large number of unsafe control actions and loss scenarios. *Proc IMechE, Part O: J Risk and Reliability* 2020: 1748006X2093971.
32. Kim H, Lundteigen MA, Hafver A, et al. Application of system-theoretic process analysis to the isolation of subsea wells: opportunities and challenges of applying STPA to subsea operations. In: *Offshore technology conference*.
33. Vinterstø T, Birkeland B, Ramberg RM, et al. Subsea compression – project overview. In: *Offshore technology conference*.
34. DNV GL. Safety 4.0, <https://www.dnvgl.com/research/oil-gas/safety40/index.html> (2018, accessed on 20 August 2020)
35. API RP 17V, Recommended practice for analysis, design, installation, and testing of safety systems for subsea applications. Standard, American Petroleum Institute, 2015.
36. Antoine B. *Systems Theoretic Hazard Analysis (STPA) applied to the risk review of complex systems: an example from the medical device industry*. PhD Thesis, Massachusetts Institute of Technology, 2013.
37. Rokseth B, Utne IB and Vinnem JE. Deriving verification objectives and scenarios for maritime systems using

- the systems-theoretic process analysis. *Reliab Eng Syst Saf* 2018; 169: 18–31.
38. Aven T and Renn O. Improving government policy on risk: eight key principles. *Reliab Eng Syst Saf* 2018; 176: 230–241.
 39. Aven T. The cautionary principle in risk management: foundation and practical use. *Reliab Eng Syst Saf* 2019; 191: 106585.
 40. Karanikas N. Documentation of assumptions and system vulnerability monitoring: the case of system theoretic process analysis (stpa). *Int J Saf Sci* 2018; 2(1): 84–93.
 41. Bjerga T, Aven T and Zio E. Uncertainty treatment in risk analysis of complex systems: the cases of stamp and fram. *Reliab Eng Syst Saf* 2016; 156: 203–209.
 42. Aven T. Practical implications of the new risk perspectives. *Reliab Eng Syst Saf* 2013; 115: 136–145.
 43. Hollnagel E, Woods DD and Leveson N. *Resilience engineering: concepts and precepts*. Hampshire: Ashgate Publishing, Ltd., 2006.

Article III

N. A. Zikrullah, M. J. P. van der Meulen, G. Skofteland, M. A. Lundteigen, A comparison of hazardous scenarios in architectures with different integration types, in: Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference (ESREL 2020 PSAM15), Research Publishing Services, 2020, pp. 4001–4008.

An error has been found in one figure after the article has been published. After contacting the publisher, unfortunately, the mistake was unable to be redacted. Therefore, we provide the correct figure here in the thesis for clarifications.

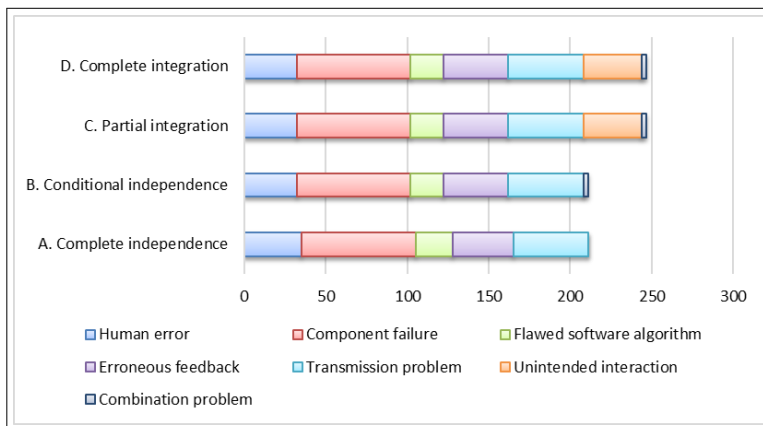


Figure A.1: Revision of 'Fig. 7. Number of loss scenarios for system with different integration types on Article III [92].

A Comparison of Hazardous Scenarios in Architectures with Different Integration Types

Nanda Anugrah Zikrullah

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Norway. E-mail: nanda.a.zikrullah@ntnu.no

Meine J.P. van der Meulen

Group Technology and Research, DNV GL, Norway. E-mail: meine.van.der.meulen@dnvgl.com

Gunleiv Skofteland

Process Technology Automation, Equinor, Norway. E-mail: gusk@equinor.com

Mary Ann Lundteigen

Department of Engineering Cybernetics, Norwegian University of Science and Technology, Norway. E-mail: mary.a.lundteigen@ntnu.no

Whether or not to allow some integration between process control and safety systems has been an ongoing debate amongst safety researchers and practitioners. The principle of keeping it simple and the principle of having segregation between the two systems are often considered as equal. The current trend is that traditional hardware implemented functions are, to an increasing extent, replaced by programmed functions and that control and safety systems rely on standard communication technologies and devices. Despite the goal of having physical segregation, the systems are no longer simple and without dependencies. Some programmable controllers have inbuilt solutions that can logically separate safety and non-safety (software and hardware) functions inside a single programmable system. It is, therefore, of interest to explore if some of these technological advances can have a positive effect on safety compared to the complexity from duplication of hardware required with segregation. Before such alternative design concepts are selected, it is necessary to evaluate if they are as safe as with physical segregation. The main objective of this paper is to identify and compare the hazards and hazardous scenarios for some selected hardware architectures ranging from complete segregation of process control and safety systems to full integration. This analysis applies the Systems-Theoretic Process Analysis (STPA) method, which has been developed to analyze complex and software-intensive systems. The result from the analysis of the selected architectures indicates that having integration will increase the number of possible scenarios leading to hazards. These scenarios may cause both safety and availability losses. This research is part of Safety 4.0, a joint industry project on research-based innovation that aims to develop a framework for safety demonstration of novel subsea technologies.

Keywords: Integration of process control and safety, subsea system, oil & gas industry, hazard analysis, systems-theoretic process analysis, STPA.

1. Introduction

The design of control systems for the subsea oil and gas industry has evolved throughout the years, starting from direct hydraulic systems in the 1960s, until the most recent all-electric systems (Bai and Bai, 2018). This evolution includes partial replacement of mechanical equipment with programmable controllers. Utilization of the latter components may allow various configurations of architectures between the process control and safety (PC&S) systems, i.e., by having different integration types (CCPS, 2016; Zikrullah et al., 2019).

Each architecture presents different kinds of scenarios leading to hazard. Hazard is defined as *a system state or set of conditions that, together with a particular set of worst-case environmental conditions, will lead to an accident (loss)* (Leveson, 2011). According to Leveson (2011), hazard analysis can be described as *investigating an accident before it occurs*.

Systems-Theoretic Process Analysis (STPA) is a new hazard analysis method that has been developed for the analysis of complex and software-intensive systems. While there are other hazard analysis methods for such systems, STPA pro-

vides several advantages. First, STPA considers hazard as a control problem. It allows the inclusion of scenarios where no failure occurs in the system (e.g., where multiple controllers provide conflicting commands) (Thomas et al., 2012). Also, it can identify interaction problems (e.g., caused by complex dependencies in the systems) (Aps et al., 2017). Theoretically, STPA can be used anytime during the design lifecycle. Leveson (2020) proposes to integrate the model used in STPA during conceptual architecture development where detailed information is not available. For more details on the comparison between STPA and other hazard analysis tools, see a technical report by Teikari (2014).

In the subsea oil & gas industry domain, STPA has been utilized to analyze different control procedures, such as the integrity pressure protection system (Rachman and Ratnayake, 2015), the isolation of subsea wells (Kim et al., 2018), and the operation of subsea gas compression system (Kim et al., 2018). Zhang et al. (2019) also used STPA for availability assessment in subsea production. However, nobody has applied STPA for systems with different integration types, as presented in this paper. The objective of this paper is to identify and compare the hazards and hazardous scenarios for some selected hardware architectures ranging from complete segregation of PC&S systems to full integration. The goal is to provide a more unobstructed view of the effect of integration on safety.

The remainder of this paper is organized as follows. Section 2 introduces the alternative concept design for PC&S systems, considering the integration type. Section 3 explains the theoretical foundation of STPA. Section 4 presents the study case of this paper. Section 5 presents the analysis results and discusses the findings. The final section identifies essential areas for further work.

2. Design of Process Control and Safety Systems for Subsea Oil & Gas Industry

2.1. Independence vs. integration

According to Drogoul et al. (2007), segregation (or independence) should be considered as one of the safe design principles to enhance safety. However, the decision to have independence or integration between process control and safety systems has been an ongoing debate among safety researchers and practitioners (Gruhn and Cheddie, 2006). IEC 61508 (2010) allows sharing the safety and non-safety elements as long as the requirement in part 1, clause 7.4.2.3 is followed. There is a limitation on the maximum safety integrity level that the system can achieve.

The aim of independence is mainly to have freedom from interference when performing the

intended function (IEC 61508, 2010). In contrast, applying integration introduces new interactions to the system that can affect its functionality. While there are measures to avoid/prevent unwanted interactions between integrated components, the system may need additional tests, analyses, and operational burdens to achieve the required functional reliability (CCPS, 2016).

In the process industries, both process control and safety systems are typically considered as a separate protection layer to achieve safety. However, integration may remove the contribution of the process control system as a protection layer (CCPS, 2014).

According to CCPS (2016), there are several issues to be addressed before claiming safety for the integrated PC&S system, as follows:

- (1.) The functional capabilities to perform the intended functions.
- (2.) The integrity of the functional performance.
- (3.) The protection against writes.
- (4.) The accessibility to control and change the safety functions.
- (5.) The barrier against cyber-threats.
- (6.) The protection against environmental issues (e.g., temperature or chemical corrosion).

In practice, for some systems (e.g., subsea oil & gas system), the option of complete independence may increase the complexity of the resulting hardware architecture. For a subsea environment with limited accessibility, performing maintenance on this complex hardware architecture may be another operational burden. The development of Commercial of The Shelf (COTS) programmable controllers that provide logical separation between process control and safety may be an alternative to solve this issue. To allow for acceptance of the integration concept for practical applications, there is a need to ensure that the integrated system can achieve the requirement in IEC 61508 (2010) associated with the safety integrity level requirements that have been derived.

2.2. Integration concept for process control and safety (PC&S) systems

Generally, the design of PC&S systems considering integration is common to any process industry. A paper by Steinhäuser (2019) discusses the integration of PC&S systems in the batch process industry.

The integration concept may be applied to any element in a control loop (e.g., sensor, logic solver, actuator, and communication). It is impossible to include all the solution spaces for integration if one would consider all the possible combinations. CCPS (2016) provides a classification type for PC&S systems, focusing on the logic solver and the communication network. A

Table 1. Breakdown of the proposed integration type classification

CCPS classification	Integration type	Physical components	Programmable logic	Network component	PC&S interaction
Air-gapped systems	A. Complete independence	Separated	Separated	None	None
Interfaced systems	B. Conditional independence	Separated	Separated	Interfaced	Limited by firewall
Integrated systems with isolated networks	B. Conditional independence	Separated	Separated	Separated	Limited by firewall
Integrated systems with shared networks	B. Conditional independence	Separated	Separated	Shared	Limited by firewall
Combined system with strong dependency	C. Partial integration	Shared	Separated	Shared	Limited by logical separation
	D. Complete integration	Shared	Shared	Shared	Depends on the configuration

new classification has been derived to simplify the CCPS classification, as follow:

- *A. Complete independence.* Each system has its control loop without any exchange between each other.
- *B. Conditional independence.* Communication between the PC&S systems (vary depending on the configuration) allows the exchange of information between systems. Independence between the two systems can still be achieved by limiting the access from the process control system to the safety system with a firewall to avoid unintended interactions. In the opposite direction access is not restricted.
- *C. Partial integration.* The integration occurs only at the hardware components of the logic solver (vary depending on the configuration). Logical separation exists between the two systems (IEC 61508 (2010) Part 3 Annex F specifies how to achieve logical separation between software elements on a single computer).
- *D. Complete integration.* Both systems completely share the use of the logic solver. No clear distinction between process control and safety logic in the programmed software.

Table 1 lists the breakdown of each integration type according to the associated components and interaction between PC&S systems. There are three architectures by CCPS (2016) that are combined into the *conditional independence* type of integration due to similarities between the use of network components. The CCPS classification of *combined system with strong dependency* is split into two integration types to distinguish the different types of integration that may exist in the logic solver (e.g., logical separation or shared software space).

3. Systems-Theoretic Process Analysis (STPA)

The STPA processes identify hazards from the variation of the control action that can be unsafe during a particular set of conditions. This hazard may develop further to become (unwanted) losses. Typically, the application of STPA produces a set of safety constraints that may limit the possible system behavior from the unwanted state.

The STPA procedures are, as follows (Leveson and Thomas, 2018):

- (i) *Define purpose of the analysis.* The analysis boundary includes the system description, system-level losses, hazards, and safety constraints.
- (ii) *Model the control structure.* A hierarchical control structure (HCS) is built based on the system (expected) interactions and behaviors during the predetermined conditions (or available information).
- (iii) *Identify Unsafe Control Action (UCA).* Keywords are used to determine whether every possible control action in the system during a set of worst-case environmental conditions will lead to UCAs. These keywords are (1) control action is provided, (2) not provided, (3) provided too early/too late, and (4) stopped too soon/applied too long.
- (iv) *Identify loss scenarios.* These scenarios are developed based on assessing every aspect in a control loop (including, e.g., feedback error, control algorithm flaws, component failures, transmission problem, incorrect actuation, and the combination between them).

4. Study Case

The study case is obtained from a subsea dry gas compressor provided by API RP 17V (2015).

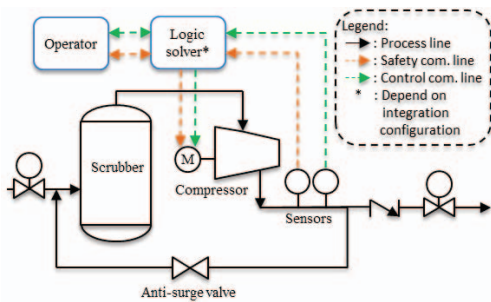


Fig. 1. Process flow diagram of a subsea compression system including process control and safety systems (developed based on API RP 17V (2015))

While it is similar to the study case used in the paper by Kim et al. (2018), this paper distinguished itself by including both process control and safety system in the loop. Also, the analyzed system is later developed at a higher level of abstraction and focuses solely on the analysis of the integrated component. This is because the objective of the paper is to compare the effect of integration, not to perform a full hazard analysis of the compressor.

Figure 1 shows the adapted process flow diagram for a subsea compression system. The system purpose is to compress the gas from the subsea flowline to the topside. Due to compression, there is a possibility to have a high temperature at the outlet of the compressors that needs to be detected by (a set of) sensors. A process control system (PCS) is used to control the compressor speed and maintain the temperature at standard conditions. The information from the sensor is processed by the PCS logic solver to provide an actuation command to the PCS actuator (in this case, a variable speed drive (VSD)). A safety system (SS) used similar information from other sensor(s) at the outlet of the compressor. At very high temperature, the SS logic solver provides a shutdown command to the SS actuator (in this case, a relay and switch) by releasing power to the system. Human operators can provide a command to the compressor through the PCS system. When there is a need for an emergency shutdown, the human operators need to shutdown the power supply system to stop the compressor operation. Human operators may also perform a normal shutdown to the system for maintenance by inhibiting the safety system (for a predefined time) and stopping the compressor manually through the PCS system.

5. Results and discussion

This section presents the selected results from STPA analyses of the subsea dry gas compression systems with different integration types. The discussions are based on thorough comparison between the obtainable results from each step of

STPA.

At the start of the analyses, it was required to define the system-level losses, hazards, and safety constraints as the boundary of the analysis. The results are captured in Table 2. The unwanted losses were related to safety issues (environment for SL1 and significant cost for SL2) and availability issues (minor cost for SL3). The boundaries of the analysis were identical for all integration types.

The analysis results for every integration types started to differ at step 2-4 of STPA. The differences are discussed in the following subsections.

5.1. Hierarchical control structure comparison

The system descriptions were modelled into HCSs for every integration type as shown in Figure 2-5. Some elements were common in the control loop, e.g., controller (human operator), sensor (PCS and SS sensor), actuator (PCS and SS actuator), and transmission line (the link between elements). Every controller had its process model based on the process information in the system description.

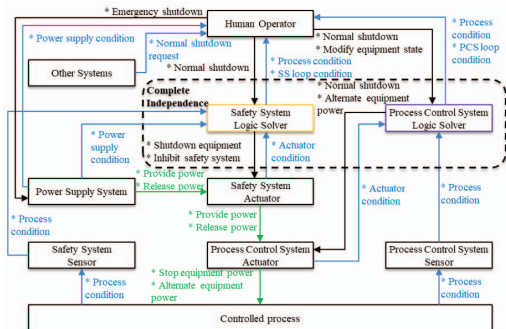


Fig. 2. Hierarchical control structure of type A. complete independence PC&S systems

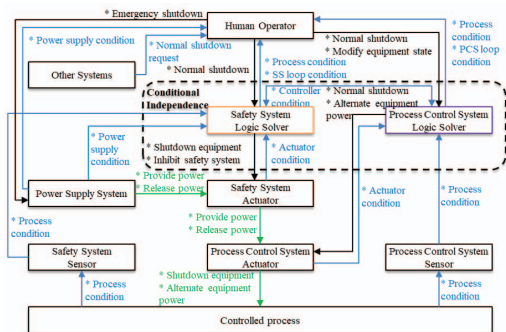


Fig. 3. Hierarchical control structure of type B. conditional independence PC&S systems

Table 2. System-level losses, hazards and safety constraints

System-level Losses (SL)	System-level Hazards (SH)	System-level Safety Constraints (SSC)
SL1. Hazardous material release to the sea	SH1. Loss of containment of hazardous material to the environment	SSC1. Equipment must be able to contain dangerous material from release to the environment
SL2. Damages to valuable equipments	SH2. Equipment operates outside normal operating condition	SSC2. Equipment must be protected from extreme operating conditions
SL3. Unnecessary interruption or reduction in hydrocarbon production	SH3. Equipment operates outside optimal operating condition	SSC3. Equipment must be operated within optimal operating conditions
	SH4. Unintended stop of equipment	SSC4. Equipment must be available to work as intended

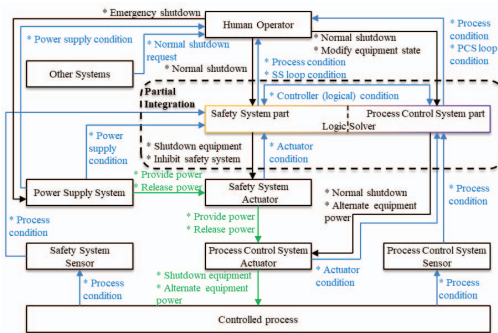


Fig. 4. Hierarchical control structure of type C. partial integration PC&S systems

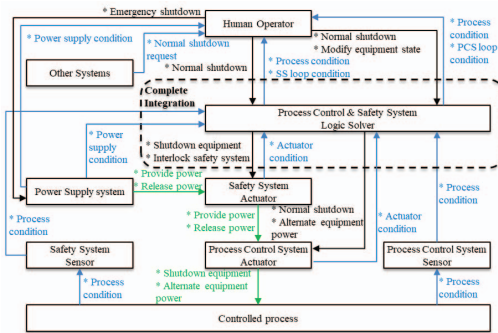


Fig. 5. Hierarchical control structure of type D. complete integration PC&S systems

Every element was connected by the black arrow, which represents the control command, blue arrow, which represents the feedback path, and green arrow, which represents the physical forces and electrical power.

Based on a discussion with experts from the industry, it was assumed that the integration does not change the required control action in the systems. However, information that can be provided

through the link between the PC&S systems may be used as a consideration when performing the specified control actions. According to Leveson (2020), the HCS should depict the control structure that does not necessarily reflect their physical architecture (especially during the conceptual operation phase where the architecture is not known yet). However, the HCS can be refined further to include (known) physical interactions in the system.

In this study case, we denote the differences between each HCS by modifying some elements in the HCS within the black dotted box to include the effect of integration. For example, in Figure 4, partial integration in the logic solver was modeled into one single box that was separated by an invisible layer (dotted lines) to show the logical separation between PCS and SS part in the logic solver. The links between elements (e.g., control action or feedback) were still located under the respective part. This model was proposed to show the distinction with the model of separated hardware, as in Figure 2 and 3.

5.2. Unsafe control action comparison

The STPA processes managed to identify 46 distinct UCAs. As shown in Figure 6, the number of identified UCAs were identical for all integration types. The possible reasoning is that the attempt to have integration in the system does not inherently

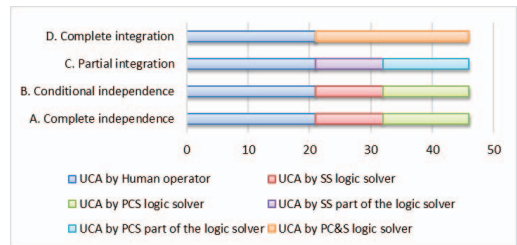


Fig. 6. Number of the UCAs for system with different integration types

change how the control action is performed (as per our initial assumption). Therefore, the development of a control action into unsafe control action during a particular set of conditions is similar for every configuration.

The only difference between the different integration types is the element performing the control action. While there were slight differences in the name of controller for integration type C and D, they were inherently the same controller as the one from integration type A and B (e.g., PCS logic solver, PCS part of the logic solver and PC&S logic solver are the hardware that has the same PCS controller responsibilities). To summarize, human operators contributed the highest number of UCAs (20) as compared to UCAs by PCS (16) and SS (11).

Some examples of UCAs are presented in Table 3. These UCAs correspond only to SH2 (10 UCAs), SH3 (23 UCAs), and SH4 (13 UCAs). More UCAs correspond to availability issues (36 UCAs from SH3 and SH4 that are linked to SL3) than safety issues (10 UCAs from SH2 that is linked to SL2).

5.3. Loss scenario comparison

Every UCA was analyzed further to identify the loss scenarios (LSc). Figure 7 shows the number of LScs for all the integration types.

A breakdown of the LScs shows that for all integration types, component failures (70 LScs) contribute the most to the number of scenarios. Erroneous feedback (37 LScs in type A, 40 LScs in type B-D) and human error (35 LScs in type A and 32 LScs in type B-D) provide a significant number of hazardous scenarios. For integration type C and D, unintended interaction (36 LScs) represents new hazardous scenarios that do not exist in the system with integration type A and B. This results in more scenarios in integration type C and D than type A and B.

Some examples of loss scenarios associated with their UCAs are presented in Table 3.

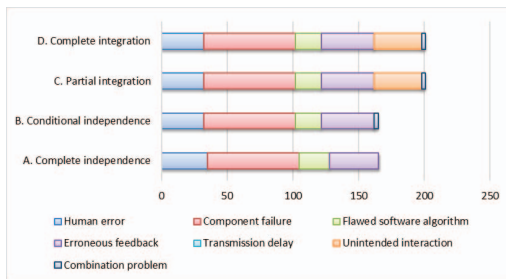


Fig. 7. Number of the loss scenarios for system with different integration types

5.4. Discussion

Identification of hazards and hazardous scenarios by STPA on the PC&S system with different integration types provides several useful insights. They are supported by selected examples that are presented in Table 3.

First, from the table, UCA 22 is one example of UCA that is identical for every integration type. It has been previously discussed that integration does not change how each controller should respond to a particular condition. However, having more integration in the system (e.g., in type C and D) may result in different scenarios that can cause hazards. Both A.LSc118.UCA22 and B.LSc120.UCA22 are scenarios that are identical for every integration type. However, C.LSc122.UCA22 and D.LSc121.UCA22 are unique scenarios that may occur only due to the integration of the logic solver hardware.

Second, for a system with integration type A, it has a higher reliance on the human operator for its decision making. In this configuration, there is no direct link between PCS and SS (see Figure 2). The human operator is an essential layer of protection in case of problems in the automated systems. The difference between scenario A.LSc118.UCA22 and B.LSc120.UCA22 shows that human error can be prevented by having an algorithm that can check whether it is possible to select the prohibited command (during a particular condition). For a system with integration type B-D, the algorithm can be implemented due to the ability to communicate directly between the two systems to allow PCS/SS condition check.

Third, there is no difference between the identified scenarios for a system having integration type C and D. One main reason is that having logical separation does not mean that the possibility of unintended interaction is removed. It just means that the engineers have, to the best of their ability, defined and limited the possible interaction paths. This issue will affect the safety demonstration process later (based on the produced safety constraints). Arguably, the system with integration type C has an easier safety demonstration process than type D due to clear separation between the PCS and SS logical architecture.

Fourth, the availability issues are more apparent when the system has more integration (type C and D). On scenarios C.LSc199.UCA35 and D.LSc199.UCA35, component failure of the shared hardware still represents a possible and major scenario leading to hazards. Countermeasures such as redundancy are vital to ensure that the availability of the systems is achieved. For the system with integration type A and B, the availability issues are shared by both PCS and SS logic solver (with possible redundancy on both controllers).

Table 3. Examples of UCAs and LScs for every integration type

Integration type	UCAs	Loss scenarios
A. Complete independence	A.UCA22 SS logic solver provides shutdown equipment command to SS actuator too late when the gas temperature is very high and the compressor is running [SH2]	A.LSc118.UCA22 Problem in the transmitted information (e.g., due to delay) prevents immediate response by the logic solver
	A.UCA25 SS logic solver does not provide shutdown equipment command to SS actuator when there is normal shutdown request, the compressor is running, and the PCS condition is not ok [SH3]	A.LSc131.UCA25 During this condition it is necessary for the human operator to provide shutdown command from the SS instead of SS inhibition command. However, human error (e.g., due to wrong procedure or no feedback information) prevents the provision of such command
B. Conditional independence	B.UCA22 SS logic solver provides shutdown equipment command to SS actuator too late when the gas temperature is very high and the compressor is running [SH2]	B.LSc120.UCA22 Algorithm flaw in the SS logic solver (e.g., due to timer or conditional algorithm) increases the processing time to provide the required response
	B.UCA25 SS logic solver does not provide shutdown equipment command to SS actuator when there is normal shutdown request, the compressor is running, and the PCS condition is not ok [SH3]	B.LSc135.UCA25 During this condition it is necessary for the human operator to provide shutdown command from the SS instead of SS inhibition command. A combination failure due to human error (that provide wrong response) and flaws in the software algorithm (that should have prevent the availability to choose inhibition command) may cause the UCA
C. Partial integration	C.UCA22 SS part of the logic solver provides shutdown equipment command to SS actuator too late when the gas temperature is very high and the compressor is running [SH2]	C.LSc122.UCA22 Resource sharing problem on the hardware delays the execution time of the command
	C.UCA35 PCS part of the logic solver provides normal shutdown command to PCS actuator when there is no normal shutdown request and the compressor is running [SH4]	C.LSc199.UCA35 Component failure of the shared logic solver may move the system to a safe state
D. Complete integration	D.UCA22 PC&S logic solver provides shutdown equipment command to SS actuator too late when the gas temperature is very high and the compressor is running [SH2]	D.LSc121.UCA22 Unintended overwrites from the PCS to SS part in the logic solver delays the proper command from SS part to the system
	D.UCA35 PC&S logic solver provides normal shutdown command to PCS actuator when there is no normal shutdown request and the compressor is running [SH4]	D.LSc199.UCA35 Component failure of the PC&S logic solver may move the system to a safe state

Finally, STPA provides help when addressing some issues (listed by CCPS) before claiming safety for the system with integration. For example, it provides us information on what scenarios that can hinder the functional capabilities of the system to perform its intended function. Other issues such as protection against writes, accessibility to control and change the safety function, barrier against cyber-threats, and the protection against environmental issues are covered by assessing whether these issues may lead to possible loss scenarios at a detail level (for every UCA). However, due to the qualitative nature of its anal-

ysis, STPA provides little help to identify whether the integrity of the functional performance has been achieved or not.

6. Conclusion and further work

This paper has discussed the application of STPA for analysis of various architectures of process control and safety system in the subsea oil & gas industry with different integration types. The discussion has been made on selected results and observable findings from the analysis.

One of the main takeaway from the analysis is that applying more integration to the PC&S

will change how the system behaves and results in more scenarios that can lead to both safety and availability losses. The designer of such a system needs to prepare countermeasures to avoid or prevent the occurrence of the scenarios.

Furthermore, the analysis of loss scenarios in STPA still has a heavy reliance on the knowledge about the possible system behavior and its assumptions. This problem, however, is similar to other hazard analysis methods. For the knowledge of the system, it is recommended that the system follows a technology qualification plan to let the analyst have more experience with the system. For the assumptions, a procedure to check and update assumptions (including the affected analysis results) during the later development of the system is needed.

Finally, the produced safety constraints from STPA should be used as guidance for the safety demonstration process. Demonstration of safety against complex scenarios in a software-intensive system still proves to be a major problem even if the possible scenarios are known. Focus on the hardware in the loop tests or digital twin procedures to provide evidence is needed as an essential research area for development.

Acknowledgement

This work has been written under the Safety 4.0 project that has been partly funded by the Research Council of Norway as part of the Petromaks 2 programme [grant number 281877/E30]. The author would like to thank the Research Council of Norway, the industrial, and university partners involved in this project for the support.

References

- API RP 17V (2015). Recommended practice for analysis, design, installation, and testing of safety systems for subsea applications. Standard, American Petroleum Institute.
- Aps, R., M. Fetissov, F. Goerlandt, P. Kujala, and A. Piel (2017). Systems-theoretic process analysis of maritime traffic safety management in the gulf of finland (Baltic Sea). *Procedia Engineering* 179, 2–12.
- Bai, Y. and Q. Bai (2018). *Subsea engineering handbook*. Gulf Professional Publishing.
- CCPS (2014). *Guidelines for Initiating Events and Independent Protection Layers in Layers of Protection Analysis*. John Wiley & Sons.
- CCPS (2016). *Guidelines for Safe Automation of Chemical Processes*. John Wiley & Sons.
- Drogoul, F., S. Kinnerly, A. Roelen, and B. Kirwan (2007). Safety in design—can one industry learn from another? *Safety Science* 45(1-2), 129–153.
- Gruhn, P. E. and H. Cheddie (2006). Safety instrumented systems: design, analysis, and justification.
- IEC 61508 (2010). Functional safety of electrical/electronic/programmable electronic safety-related systems – part 1-7. Standard, International Electrotechnical Commission.
- Kim, H., M. A. Lundteigen, A. Hafver, F. B. Pedersen, G. Skoffeland, et al. (2018). Application of system-theoretic process analysis to the isolation of subsea wells: Opportunities and challenges of applying STPA to subsea operations. In *Offshore Technology Conference*. Offshore Technology Conference.
- Kim, H., M. A. Lundteigen, A. Hafver, F. B. Pedersen, G. Skoffeland, C. Holden, and S. J. Ohrem (2018). Application of systems-theoretic process analysis to a subsea gas compression system. In *Safety and Reliability – Safe Societies in a Changing World – Proceedings of the 28th International European Safety and Reliability Conference, ESREL 2018*, pp. 1467–1476. CRC Press/Balkema.
- Leveson, N. (2011). *Engineering a safer world: systems thinking applied to safety*. Engineering systems. Cambridge, MA: The MIT Press.
- Leveson, N. (2020). *An Improved Design Process for Complex, Control-Based Systems Using STPA and a Conceptual Architecture*. Draft retrieved 2020–02–11. Unpublished.
- Leveson, N. and J. Thomas (2018). *STPA handbook*.
- Rachman, A. and R. C. Ratnayake (2015). Implementation of system-based hazard analysis on physical safety barrier: A case study in subsea HIPPS. In *2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pp. 11–15. IEEE.
- Steinhauser, E. P. (2019). Addressing the challenges of implementing safety instrumented systems in multi-product batch processes. *Journal of Loss Prevention in the Process Industries* 57, 164–173.
- Teikari, O. (2014). CORSICA task 4.1 hazard analysis method of digital I&C systems. *Technical Report VTT-R-03821-14*.
- Thomas, J., F. Lemos, and N. Leveson (2012). Evaluating the safety of digital instrumentation and control systems in nuclear power plants. *NRC Technical Research Report 2013*.
- Zhang, J., H. Kim, Y. Liu, and M. A. Lundteigen (2019). Combining system-theoretic process analysis and availability assessment: A subsea case study. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 520–536.
- Zikrullah, N. A., M. J. P. van der Meulen, H. Kim, and M. A. Lundteigen (2019). Clarifying implementation of safe design principles in IEC 61508: Challenges of novel subsea technology development. In *Proceedings of the 29th European Safety and Reliability Conference (ESREL)*, pp. 2928–2936. Research Publishing.

Article IV

N. A. Zikrullah, M. J. P. van der Meulen, M. A. Lundteigen, Finite-state automata modeling pattern of systems-theoretic process analysis results, *Reliability Engineering & System Safety*. (n.d.), under review.

This article is awaiting publication and is therefore not included.

ISBN 978-82-326-5459-8 (printed ver.)
ISBN 978-82-326-6561-7 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)



NTNU

Norwegian University of
Science and Technology