Modeling, Identification and Control, Vol. 35, No. 4, 2014, pp. 333–346, ISSN 1890–1328

Verification and Examination Management of Complex Systems

Stian Ruud [1] Roger Skjetne [2]

[1] *Section for Control Systems, DNV GL, N-1322 Høvik, Norway, E-mail:* stian.ruud@dnvgl.com

[2] *Department of Marine Technology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway. E-mail:* roger.skjetne@ntnu.no

Abstract

As ship systems become more complex, with an increasing number of safety-critical functions, many interconnected subsystems, tight integration to other systems, and a large amount of potential failure modes, several industry parties have identified the need for improved methods for managing the verification and examination efforts of such complex systems. Such needs are even more prominent now that the marine and offshore industries are targeting more activities and operations in the Arctic environment. In this paper, a set of requirements and a method for verification and examination management are proposed for allocating examination efforts to selected subsystems. The method is based on a definition of a verification risk function for a given system topology and given requirements. The marginal verification risks for the subsystems may then be evaluated, so that examination efforts for the subsystem can be allocated. Two cases of requirements and systems are used to demonstrate the proposed method. The method establishes a systematic relationship between the verification loss, the logic system topology, verification method performance, examination stop criterion, the required examination effort, and a proposed sequence of examinations to reach the examination stop criterion.

Keywords: Verification management; Verification risk; Audit planning; Audit risk; Complex technical systems

1 Introduction

With the accelerated evolution of information and communication technology, the maritime industries have experienced in a short amount of time a significant change from conventional mechanical ships to modern computer-controlled ships. Ship system technology has to some extent developed and been taken into use faster than corresponding verification methods. In (Skjetne and Sørensen, 2004) a consortium of maritime industry partners expressed a need for further research to describe the experienced problems related to increased computer-based integration and software problems. As an answer to this, the need for managing the verification and examination efforts were identified by the industry in the years 2005-2010 by, for instance, DNV, Kongsberg Maritime, Statoil, Farstad, Marine Cybernetics, and Global Maritime during the development of advanced verification and certification methods for complex ship systems (such as DP systems), resulting among others in the DNV GL Recommended Practice (DNV, 2012). An example of a more advanced verification method in the maritime industry is the industry-established Hardware-In-the-Loop (HIL) simulation method (Skjetne and Egeland, 2006), which makes it possible to generate many more detailed and low-level test cases compared to the traditional Failure Mode and Effect Analysis (FMEA) testing of redundant ship systems. The reason for this, as explained by Skjetne and Egeland (2006), is that the HIL test

doi:10.4173/mic.2010.4.9

tool makes many more software and hardware functions testable in the target test system, that is, more functions are controllable (can be trigged and manipulated) and observable (the function behaviors can be observed and measured) by this test tool.

Besides control systems, also offshore operations are becoming more complex as new extreme frontiers are challenged. For instance, the reduced sea-ice extent in the Arctic due to global warming in recent years has provided new industrial opportunities. The shipping industry together with the Arctic countries has initiated development and increased use of the Northerns Sea Route for more efficient transportation from Europe to Asia. The offshore industry has shown an increased interest in integrated offshore operations for petroleum activities in Arctic ice-covered waters. Such operations are technically and physically more challenging than conventional open-water operations due to remoteness and general lack of infrastructure, low temperatures, darkness, and the presence of sea-ice and icebergs. Stationkeeping operations by position mooring (PM) or dynamic positioning (DP) are, as an example, challenging since ice forces are much stronger and rapidly varying compared to conventional open-water environmental forces, and the technical control systems have not been developed for the Arctic climate and ice loads. Despite this fact, it is said that the risk of offshore activities in the Arctics should not be higher than in the North Sea. Assuming then that consequences of an accident is higher in an ecological sensitive Arctic area, this means that the probability of an incidents must be reduced by additional technical and operational barriers. It follows from this that improved management of examination and verification methods of the new barriers is needed.

The new situation in computer-based control systems and more operations in extreme environments have raised some concerns, such as:

- The new system topologies are more integrated on a ship-wide scale and become more complex. The physical component topologies are well specified and can be assumed to be known for the verifiers, while software and computer-based communication topologies are not.

- The number of operational modes and combinations of user-allowed settings of the systems are increasing.

- Larger integrated operations in sensitive environments, like in the Arctics, require a better overall assessment of integrated functions, systems, and barriers against operational failure. Vessels should be verified as an integrated part of a larger system.

- There is a need for explicit assessment of the applicable failure modes, that is, to identify potential failure modes and to determine a relevant subset to verify while discriminating other failure modes.

- The potential verification scope of software functionality is very large, where the functionality is not specified in detail, and the inter-dependencies between software functions are unclear.

- There is a need to select the optimal sequence of examinations, to define adequate verification methods and sufficient quantities of examinations, under the assumption and general acceptance that verification cannot require complete coverage of all possible examinations. Hence, selection of the most beneficial examination method for a given requirement is needed.

- The system topologies or requirements structures may influence the need for examination effort, the selection of verification method, and the verification result.

In (Skjetne and Sørensen, 2004) it is stated that many different verification methods now exist for use in an industrial context, where testing by HIL simulation is one example of a verification method. Application of each such method require significant costs, especially if testing is to happen at sea trial for a vessel or within the Arctic environment. This further emphasizes the need for better management of verification.

Based on the above described situation, the industry partners of the project leading to (Skjetne and Sørensen, 2004) expressed a need for further research to obtain a more general understanding of how to describe and estimate verification contributions, and how to optimally put together a verification portfolio for a vessel or a system. In, for instance, the context of ship systems or integrated power systems, different types of representation and visual presentation of complex systems should be studied, and precise definitions of verification benefits should be proposed. Such definitions may then be used to collect empirical information about properties and relevance of existing verification methods and activities, for instance, standards like IEC 61508 (IEC, 2010), ISO 9000 types of assessments, class rules, certification, approval, manufacturing surveys, FMEAs, HIL simulation, software quality assurance techniques, dock trials, sea trials, field trials, and annual trials.

Given that sufficient knowledge related to the benefits of verification methods for various types of components or systems is established, the industry asks for a systematic approach to verification management. A

key issue in verification management is to find methods for allocation and sequencing different verification activities (traditional surveys, FMEAs, HIL, and other methods) in different parts of the lifecycle of the ship or system. To achieve this, relationships between system complexities, operational modes, verification volumes, and verification confidence levels must be studied, including the possible need to set limits on how complex systems can be built in order to ensure sustainable development with verification within reasonable use of resources.

In financial audits, the auditors are applying audit risk concepts in their planning of audits. The audit planning and the proposed verification management are in general representing the same type of considerations and parameters to be used for allocating examination resources as proposed in this paper. In the text box below are given brief quotes of the main concepts relating to the financial audit planning and audit risk (Arens et al., 2006; AICPA, 2006).

Audit risk *Audit risk (AR) is the risk that the auditor may unknowingly fail to appropriately modify his or her opinion on financial statements that are materially misstated.*
The model $AR = RMM \times DR$ expresses the general relationship of audit risk, the risks associated with the auditor's assessments of risk of material misstatement (RMM) (inherent and control risks), and the detection risk (DR).
Risk of material misstatement *Risk of material misstatement (RMM) is the product of inherent risk (IR) and control risk (CR).*
Control risk *Control risk (CR) is the risk that a misstatement that could occur in a relevant assertion and that could be material, either individually or when aggregated with other misstatements, will not be prevented or detected on a timely basis by the entity's internal control.*
Detection risk *Detection risk (DR) is the risk that the auditor will not detect a misstatement that exists in a relevant assertion that could be material, either individually or when aggregated with other misstatements. Detection risk is a function of the effectiveness of an audit procedure and of its application by the auditor. Detection risk cannot be reduced to zero because the auditor does not examine 100 percent of an account balance or a class of transactions and because of other factors.*

The objectives of this article are to explain the industrial need for examination and verification management of complex systems, and then to propose a novel method for verification and examination management based on industrial needs and some concepts and practices found in financial auditing.

The proposed method will be demonstrated through two case studies on how to manage the quantity and sequence of examination activities on subsystem level in order to reach an acceptable level of verification risk. The first case is one requirement to give a decision if the specified examination should be carried out. The second case illustrates verification management for a redundant system with a common component.

2 Problem formulation

2.1 The verification management approach

We define verification according to (IEC, 2010) as *"confirmation by examination and provision of objective evidence that the requirements have been fulfilled"*.

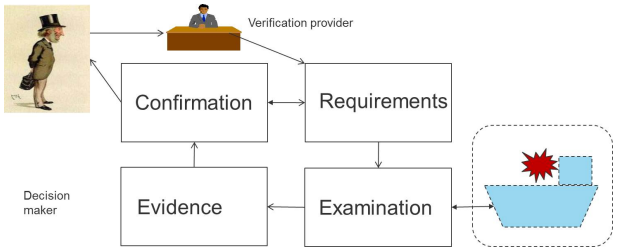

Figure 1: Illustration of the verification process.

We consider a verification on a system – subsystem – component level, where the system is constructed by a set of subsystems $\{A, B, C, \ldots\}$, and each subsystem is constructed by a set of components that all must function for the subsystem to function. The overall requirement is typically related to some main mode or system function (e.g. *stationkeeping* mode for a DP system) or an operation (e.g. *Ice Management operation* for an Arctic offshore operation). Correspondingly, we assume there is an overall requirement Y representing the overall system function. This is constructed by a set of main requirements $\{H_A, H_B, H_C, \ldots\}$ corresponding to subsystems. Each main requirement H_A is again constructed by a set of subrequirements $\{h_{A1}, h_{A2}, h_{A3}, \ldots\}$ on component level.

Verification management (VM) is proposed for cost-efficient verification of the overall system requirement Y. This involves establishing a system verification loss L, examination methods for each main requirement, an examination stop criterion Z for the overall verification process, and a decision rule of the verification result. Finally, VM must apply the decision rule for concluding the verification result – which is the result of the verification activity.

Examination management (EM) is a part of the verification management process. The examination management is to decide on zero examination effort or select a sequence of examination efforts $\{x_A, x_B, x_C, \ldots\}$ by corresponding methods for the main requirements by means of an estimate of the verification risk Ψ subject to minimization.

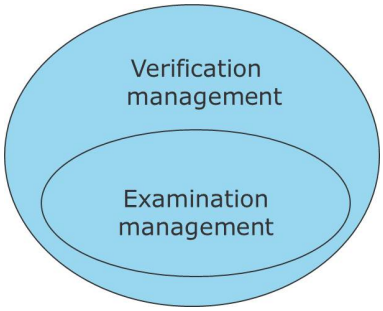

Figure 2: Verification and examination management.

Based on the background situation and industry needs a method for examination management is proposed, including the following main elements:

- A system is represented by an overall requirement Y. This is constructed by a set of main requirements, represented by the vector $H = (H_A, H_B, H_C, \ldots)$, through a Boolean structure function $Y = \Phi(H)$.
- Each main requirement H_j, $j = A, B, C, \ldots$, is assumed constructed from a set of subrequirements $\mathcal{H}_j = \{h_{j1}, h_{j2}, h_{j3}, \ldots\}$ that all must be satisfied for the corresponding main requirement to be satisfied, i.e. $H_j = h_{j1} \wedge h_{j2} \wedge h_{j3} \wedge \ldots$.
- Each subrequirement h_{ji} and each main requirement H_j take a state value (True) if the requirement is 'compliant', and (False) if the requirement is 'noncompliant'.
- Based on previous experience, statistics, or conservative estimates, we assume *a priori* (before examination; denoted by subscript 0) knowledge of the probability $p^c_{0,H_j} = P(H_j)$ that requirement H_j is compliant. Conversely, we assume the probability $p^{nc}_{0,H_j} = P(\neg H_j)$ that H_j is noncompliant.
- The examination of a main requirement H_j is characterised by the subset of *'examined subrequirements'* $\mathcal{E}_j(x_j) \subseteq \mathcal{H}_j$ and the subset of *'unexamined subrequirements'* $\mathcal{U}_j(x_j) \subseteq \mathcal{H}_j$, where x_j is some examination effort for H_j.
- For the overall verification activity there is a stop criterion Z for stopping further examination (Z is a Boolean expression).
- To the overall requirement Y is associated a potential loss L due to potentially wrong verification result (for instance that a noncompliant requirement is accepted).

Note the distinction between the main requirement H_j, which is a scalar Boolean state variable, and the set $\mathcal{H}_j$ that merely lists the "set of subrequirements" that H_j is constructed from. The restriction that all subrequirements in $\mathcal{H}_j$ must be compliant (true) for H_j to be compliant (true) means that we can relate the logical outcome of H_j to the "size" of the subset in $\mathcal{H}_j$ with equivalent outcome.

2.2 Illustrating example to clarify the concepts

Consider the situation where a forklift shall either be accepted or rejected by the buyer. The machine shall be accepted if the requirement Y is complied with. Assume that there is given an *a priori* probability $p^c_{0,Y} = P(Y)$ that the requirement Y is complied with, and conversely $p^{nc}_{0,Y} = P(\neg Y)$ that the requirement is not complied with.

The buyer is offered the possibility to carry out a verification and examination effort x, and this examination will clarify for certain if the requirements are complied with or not. The value of the machine is L, and in the case that the buyer accepts the machine, he has to pay the value L of the machine. This means that if the forklift is accepted without examination, the possible loss may be L for the buyer.

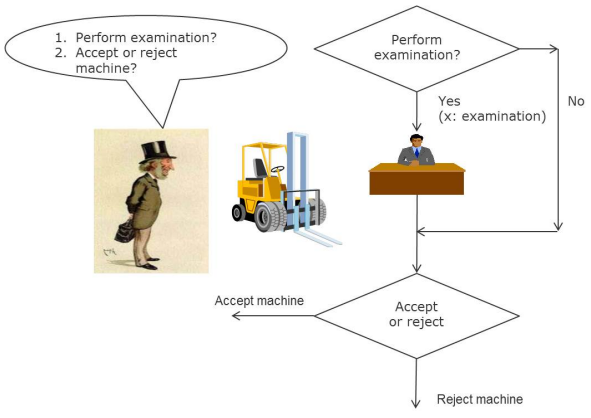

Figure 3: The buyer of the forklift may examine the forklift before accepting or rejecting it.

We will revisit this example in the first case study in order to demonstrate the detailed steps of the verification management method.

2.3 Abbreviations and notations

Boolean operators are $\wedge$ for logical AND, $\vee$ for logical OR, and $\neg$ for negation. In addition we have the following nomenclature:

- $A, B, C, \ldots$: Applied names of subsystems.
- h_{ji}: Subrequirements for the subsystem $j = A, B, C, \ldots$, $i = 1, 2, \ldots, M_j$, typically representing components or functions. Each subrequirement is assumed statistically mutually independent and takes a state value 'True' if the requirement is compliant, and 'False' if it is noncompliant.
- H_j: Main requirements, $j = A, B, C, \ldots$, constructed by a series of subrequirements. Each main requirement takes a state value 'True' if the requirement is compliant and 'False' if it is noncompliant. All main requirements are collected in a state vector $H = (H_A, H_B, H_C, \ldots)$.
- Y: Overall system requirement constructed by a Boolean structure function of main requirements $Y = \Phi(H)$.
- $\mathcal{H}_j$: Set of subrequirements for H_j, $j = A, B, C, \ldots$, for instance $\mathcal{H}_A = \{h_{A1}, h_{A2}, h_{A3}\}$.
- $\mathcal{H}_j^c$: The subset of subrequirements, $\mathcal{H}_j^c \subseteq \mathcal{H}_j$, that are compliant.
- $\mathcal{H}_j^{nc}$: The subset of subrequirements, $\mathcal{H}_j^{nc} \subseteq \mathcal{H}_j$, that are noncompliant. This gives $\mathcal{H}_j^c \cup \mathcal{H}_j^{nc} = \mathcal{H}_j$.
- x_j: Examination effort by means of an examination method for requirement H_j. We collect all efforts into a vector $x = (x_A, x_B, x_C, \ldots)$.
- $\mathcal{E}_j(x_j)$: Examined set of subrequirements for H_j as a function of examination effort x_j. Note: $\mathcal{E}_j(x_j) \subseteq \mathcal{H}_j$.
- $e_j(x_j)$: Scalar measure of examined subrequirements in $\mathcal{H}_j$. We collect all examination functions into a vector $e(x) = (e_A(x_A), e_B(x_B), e_C(x_C), \ldots)$.
- $\mathcal{U}_j(x_j)$: Unexamined set of subrequirements for H_j as a function of examination effort x_j. Note: $\mathcal{U}_j(x_j) \subseteq \mathcal{H}_j$, $\mathcal{E}_j(x_j) \cup \mathcal{U}_j(x_j) = \mathcal{H}_j$, and $\mathcal{E}_j(x_j) \cap \mathcal{U}_j(x_j) = \emptyset$.
- $u_j(x_j)$: Scalar measure of unexamined subrequirements in $\mathcal{H}_j$. We collect all the functions into a vector $u(x) = (u_A(x_A), u_B(x_B), u_C(x_C), \ldots)$.
- L: The loss which may follow as a consequence of wrong verification decision for requirement Y.
- $\Psi(x)$: Verification risk for the system requirement Y. Initial verification risk is denoted Ψ_0.
- $\frac{\partial \Psi}{\partial x_j}$: Sensitivity of verification risk Ψ with respect to effort x_j for the respective verification method.
- p_{0,H_j}^c, p_{0,H_j}^{nc}: *a priori* probability $p_{0,H_j}^c = P(H_j)$ and $p_{0,H_j}^{nc} = P(\neg H_j)$ before examination work, where $p_{0,H_j}^c + p_{0,H_j}^{nc} = 1$.
- Z: Examination stop criterion, assumed to take a logic value (true or false).

Note that scripted notation is used for sets of requirements, e.g. $\mathcal{H}$, $\mathcal{E}$, and $\mathcal{U}$. If we in the text discuss a single main requirement, we typically use $j = A$ and H_A (without loss of generality). In most of the paper, this is the case.

3 Proposed examination and verification management method

Verification or examination management shall recommend if and how much examination effort x one should perform before accepting or rejecting conformance to the given requirement Y. We make the assumptions:

- Subsystems, components, and examination of subrequirements are assumed to be statistically mutually independent.
- In a complex system consisting of a number of subsystems $(A, B, C, \ldots)$ the examination is assumed to be performed at the subsystem level and the result of the examination can be aggregated to the top system level by means of standard risk and reliability methods. All requirements are possible to be examined.
- If all subrequirements are examined and thus all main requirements verified, then a completely correct verification decision will be made for Y. This assumes that a selected verification method gives perfect certainty of compliance if performed.
- An examination of a requirement shall establish the state of the requirement, being either compliant or non-compliant.
- The state of a requirement is assumed to be unchanged due to an examination effort. This implies that examination must be nondestructive. We note, however, that for some types of examination the testing may influence the state of the equipment to be improved or possibly to become worse (destructive testing).

- In the case that a noncompliant requirement is identified by the examination, it has to be decided if the state of the requirement (or the component) should be fixed, or if the state shall remain noncompliant as identified.

Note that the verification risk will be the same in both cases if the noncompliant requirement is restored or not, since verification risk is only related to the knowledge of the requirement's state for making the correct verification decision.

3.1 Requirements and outcome of verification decision

The overall requirement Y may be a single requirement, $Y = H_A$, or a complex Boolean expression containing a number of requirements. This is generally represented by a Boolean structure function $Y = \Phi(H)$, e.g. $Y = \Phi(H) = (H_A \vee H_B) \wedge H_C$. Each main requirement H_A is constructed by a series of subrequirements to be satisfied, that is, $H_A = h_{A1} \wedge h_{A2} \wedge h_{A3} \wedge \ldots$.

The corresponding set of subrequirements, e.g. $\mathcal{H}_A = \{h_{A1}, h_{A2}, h_{A3}\}$, may originate from standards, class rules, recommended practices, functional specifications, or be agreed between the user of the verification result and the verifier.

The verifier's task is then to decide on an examination scope for the requirements, perform the planned examinations, and thereby provide evidence for accepting or rejecting the requirements. To be able to make a decision regarding whether to accept or reject the requirement, the verifier must choose a sufficient set of examinations. Since the main requirement H_A logically needs all respective subrequirements to be compliant, a scope of examinations of the subrequirements is needed. The examination may include verification of the complete set of subrequirements in $\mathcal{H}_A$, a subset of $\mathcal{H}_A$, or no examination at all. The outcome of such a process may in principle be four different scenarios, as indicated in Figure 4:

1. H_A is actually true and verifier accepts requirements: Right decision, green box.
2. H_A is actually true but verifier rejects requirements: Wrong decision, yellow box, Type I error.
3. H_A is actually false but verifier accepts requirements: Wrong decision, red box, Type II error.
4. H_A is actually false and verifier rejects requirements: Right decision, green box.

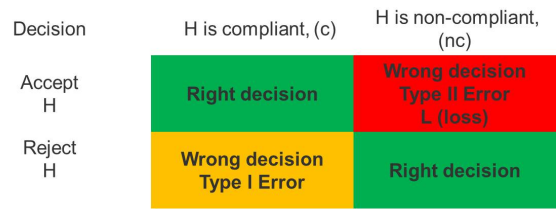

Figure 4: Possible outcomes of verification decision.

3.2 Verification loss

In this paper the outcome that a wrong verification decision, by accepting a noncompliant requirement, is denoted as a Type II error, which is similar to the terminology used in hypothesis testing (NIST/SEMATECH, 2014). Obviously, the critical requirements are those that in reality are noncompliant and unexamined, that is, those in the set $\mathcal{H}_A^{nc} \cap \mathcal{U}_A(x_A)$.

Verification loss L is the worst-case loss that may follow as a result of a wrong verification decision for a Type II error. L is the consequence of the verification decision for a given stakeholder (owner, yard, verifier, etc.) in the case that requirement Y is wrongly accepted in the verification decision. This value should be established by the verifier together with the stakeholders applying the verification result. The loss due to a Type I error is assumed smaller and, thus, not considered in this paper.

The verification loss parameter L may be described by consequence classes related to, for instance:

- Fatalities/injuries from accidents.
- Environmental consequences (pollution).
- Loss of facilities.
- Income losses due to operational unavailability.
- System operational risk; risk without barriers or risk with barriers.
- Insurance coverage.

The idea in this context is to assume that the verification loss L can be established upfront the verification work and be given as a prerequisite for the verification and examination management. At the start of the verification, no examination work x has been done and the potential verification loss is L.

For a given loss value L, a conservative definition of overall verification risk Ψ can be proposed as:

$$\Psi = L. \tag{1}$$

In this paper only the potential loss of a Type II error is considered, but it is possible to include more error types and loss effects.

In the conservative definition of the verification risk it is assumed that no effects of *a priori* knowledge, or examination, or other effects (degradation) have been taken into account in the estimate Ψ. Compared to a normal risk equation, this definition assumes by its conservative nature that the *a priori* probability of Y being noncompliant (false) is equal to 1.

3.3 *A priori* probability of state of requirement

In most cases the verification manager has some presumed and possibly conservative *a priori* information about the expected (probable) outcome of an examination of H_A. The verifier may decide to not perform any examination and only base the verification decision on the *a priori* probability $p^c_{0,H_A} = P(H_A)$. Another general assumption for verification and examination management is that exhaustive examination of a complex system and complex requirements is often not possible. This means that the practical verification results may only be possible based on a partial examination or no examination at all. Another situation is when the requirement H_A consists of many subrequirements $\{h_{A1}, h_{A2}, h_{A3}, \ldots\}$, where the verifier must evaluate which subrequirement that shall be given the highest examination priority and which shall be given less effort.

We assume that the set of subrequirements can be divided into a set of compliant requirements $\mathcal{H}^c_A$ and a set of noncompliant requirements $\mathcal{H}^{nc}_A$, as illustrated in Figure 5.

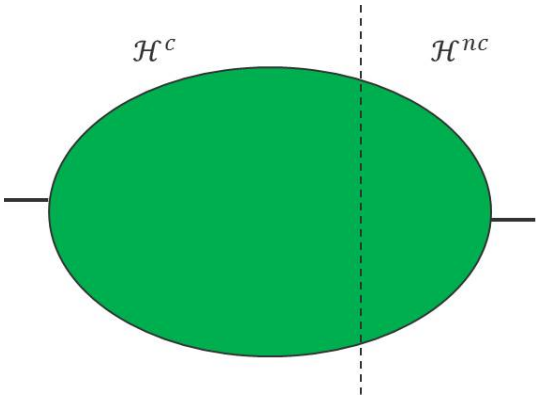

Figure 5: Illustration of the set of subrequirements $\mathcal{H}_A = \mathcal{H}^c_A \cup \mathcal{H}^{nc}_A$ for the main requirement H_A. The *a priori* assumed initial state of H_A could be given by the estimated probabilities $P(H_A) = p^c_{0,H_A}$ and $P(\neg H_A) = p^{nc}_{0,H_A}$. This is used for initiating the verification and examination management process.

There is then a need for describing some *a priori* knowledge of the initial (or unexamined) state of the requirements. This *a priori* information is given as a probability that the requirement H_A is initially compliant, that is,

$$p^c_{0,H_A} = P(H_A) = P(\mathcal{H}^c_A) \tag{2}$$

where the initial estimate could, for instance, be that $p^c_{0,H_A} = 0.5$. Conversely, the *a priori* information could be given as the probability of H_A being noncompliant, that is,

$$\begin{aligned} p^{nc}_{0,H_A} &= P(\neg H_A) = P(\mathcal{H}^{nc}_A) \\ &= P(\mathcal{H}_A) - P(\mathcal{H}^c_A) = 1 - p^c_{0,H_A}. \end{aligned} \tag{3}$$

We let Φ^c_a be the algebraic function relating the probabilities of the main requirements H being compliant to the probability of the Boolean function $Y = \Phi(H)$ being compliant. Equivalently, we let Φ^{nc}_a relate the probabilities of the main requirements H being noncompliant to Y being noncompliant. This is simply obtained by the substitutions

$$P(H_a \wedge H_b) = P(H_a)\, P(H_b) \tag{4}$$

$$P(H_a \vee H_b) = P(H_a) + P(H_b) - P(H_a)\, P(H_b) \tag{5}$$

$$P(\neg H_a) = 1 - P(H_a). \tag{6}$$

Then we get

$$p^c_{0,Y} = P(Y) = P(\Phi(H)) = \Phi^c_a(p^c_{0,H}) \tag{7}$$

$$p^{nc}_{0,Y} = P(\neg Y) = P(\neg\Phi(H)) = \Phi^{nc}_a(p^{nc}_{0,H}), \tag{8}$$

where $p^*_{0,H} = (p^*_{0,H_A}, p^*_{0,H_B}, p^*_{0,H_C}, \ldots)$.

Based on the specified verification loss and *a priori* given probabilities, we propose to estimate the *a priori* verification risk as a fraction of the verification loss by

$$\Psi_0 = L \cdot P(\neg Y) = L \cdot p^{nc}_{0,Y} = L \cdot \Phi^{nc}_a(p^{nc}_{0,H}). \tag{9}$$

This means that initially with no examination ($x = 0$), the verification risk is equal to the risk of the overall system requirement Y being noncompliant.

3.4 Examination of requirements

In a verification and examination activity the verifier will normally examine the specified sets of requirements $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C, \ldots$ before concluding the verification decision (accept/reject) for Y. For each main requirement H_A, the examined set of subrequirements is expressed by $\mathcal{E}_A(x_A)$, which grows with increasing examination effort x_A. The unexamined part is the set $\mathcal{U}_A(x_A)$ of subrequirements that have not yet been examined. This set decreases with increasing examination effort x_A. It follows that $\mathcal{H}_A = \mathcal{E}_A(x_A) \cup \mathcal{U}_A(x_A)$ and $\mathcal{E}_A(x_A) \cap \mathcal{U}_A(x_A) = \emptyset$ as illustrated in Figure 6.

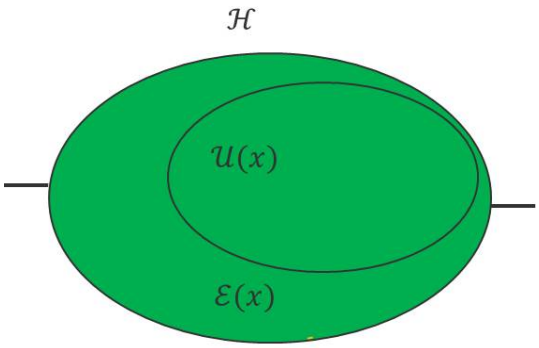

Figure 6: The set $\mathcal{H}_A$ of subrequirements is divided into examined and unexamined requirements. Initially, all requirements in $\mathcal{H}_A$ are unexamined such that $\mathcal{U}_A(0) = \mathcal{H}_A$ and $\mathcal{E}_A(0) = \emptyset$. When effort x_A is increased, the set $\mathcal{E}_A(x_A)$ will grow and $\mathcal{U}_A(x_A)$ will decrease.

At the start of the examination ($x_A = 0$) the examination status of the requirement H_A is 'not examined'. Let a scalar characteristics function $u_A(x_A)$ describe how the unexamined set $\mathcal{U}_A(x)$ decreases for increasing examination effort, that is, $u_A(x_A) \in [0, 1]$ with $u_A(0) = 1$ and $u_A(x_A) = 0$ when the requirement H_A has been completely examined – in which case $\mathcal{U}_A(x_A) = \emptyset$. Similarly, we define the scalar characteristics function $e_A(x_A)$ to describe how the set $\mathcal{E}_A(x_A)$ grows with increased examination effort. Without loss of generality, we let $e_A(x_A) \in [0, 1]$ with $e_A(0) = 0$ and $e_A(x_A) = 1$ when the requirement H_A has been completely examined. We choose $u_A(x_A)$ and $e_A(x_A)$ such that $u_A(x_A) + e_A(x_a) = 1$.

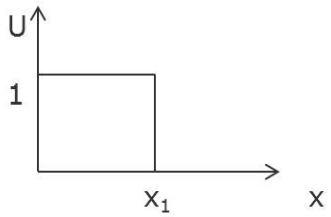

Figure 7: A requirement is unexamined $u_A(0) = 1$ (or $e_A(0) = 0$) for $x_A = 0$ and examined $e_A(x_A) = 1$ (or $u_A(x_A) = 0$) at $x_A = x_1$.

For complex systems and complex requirements, exhaustive examinations of all requirements and subrequirements can in practice not be accomplished as the costs of the examinations will typically be too large compared to the possible benefits. To quantity sufficient examinations of the requirements in order to make the verification process cost-efficient is the key result to be established by the verification and examination management process.

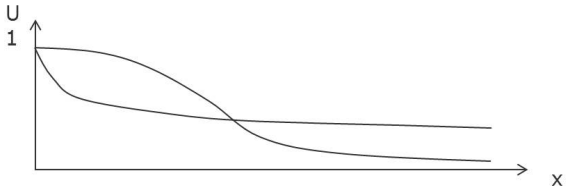

Figure 8: Examples of functions for unexamined parts of $\mathcal{H}$. Such curves may be relevant and applicable for complex requirements (upper curve could e.g. be HIL, while the lower curve could be FMEA).

The characteristics function $u_A(x_A)$ may take different shapes. If, for instance, a single test can verify the status of a single requirement H_A, then $u(x_A)$ may take the shape illustrated in Figure 7. If H_A is constructed by a series of subrequirements, then $u_A(x)$ may take the form of a staircase function linearly stepping down from one towards zero as all subrequirement are tested. However, typically a requirement will consist of a large number of subrequirements, each subrequirement will possibly need several tests, examination will have an initial cost and need preparations, and results will need post-analysis. Thus, $u_A(x_A)$ will more generally be characterized by some curve as illustrated in Figure 8.

Correspondingly, we assume that the characteristics function $u_A(x_A)$ describing the examination of a requirement H_A is a continuous function that monotonically decreases with increasing examination effort x_A.

An important assumption made, is that if $u_A(x_A) = 0$ and the verifier has completely examined the set of requirements $\mathcal{H}_A$, then the verification decision will always be correct – either accepting or rejecting H_A (green boxes outcome in Figure 4). This means that the intended effect of increased examination is to reduce the number of unexamined subrequirements that may lead to a verification loss; see Figure 9. However, since some examinations (efforts) will give larger verification risk reduction than others, it is important in verification and examination management to prioritize the examinations based on cost-benefit considerations.

In order to estimate how the expected verification risk changes with examination effort, we note that the probability of Type II errors becomes

$$P\left(\mathcal{H}_A^{nc} \cap \mathcal{U}_A(x_A)\right). \tag{10}$$

Using the characteristics function $u_A(x_A)$ as the mea-

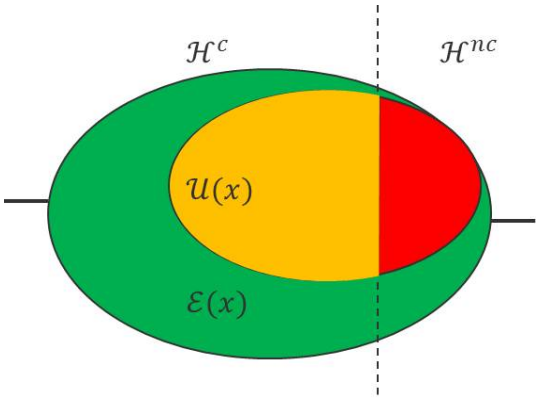

Figure 9: Illustration of the sets of subrequirements $\mathcal{H}_A^c$, $\mathcal{H}_A^{nc}$, and how these are overlapped with $\mathcal{E}_A(x_A)$ and $\mathcal{U}_A(x_A)$. The verification risk is estimated by means of the intersection (red part) of the unexamined set $\mathcal{U}_A(x_A)$ and the set of noncompliant requirements $\mathcal{H}_A^{nc}$.

sure of $\mathcal{U}_A(x_A)$, we analytically calculate (10) by

$$P\left(\mathcal{H}_A^{nc} \cap \mathcal{U}_A(x_A)\right) \approx p_{0,H_A}^{nc} \cdot u_A(x_A). \tag{11}$$

Recalling the function $\Phi_a^{nc}(\cdot)$ in (8), the proposed measure of the verification risk as a function of examination effort is then

$$\Psi(x) = L \cdot \Phi_a^{nc}(p_{0,H}^{nc} \circ u(x)), \tag{12}$$

where $\circ$ denotes the element-wise product between the two corresponding vectors, that is

$$p_{0,H}^{nc} \circ u(x) = \begin{bmatrix} p_{0,H_A}^{nc} \cdot u_A(x_A) \\ p_{0,H_B}^{nc} \cdot u_B(x_B) \\ \vdots \end{bmatrix}. \tag{13}$$

As illustrated in Figure 10, this shows that the verification risk initially (with no examination $x = 0$ such that $u(x) = (1, 1, 1, \ldots)$) takes the value of the *a priori* estimated risk Ψ_0. Then the verification risk reduces with increased examination effort according to the examination function $u(x)$. If for some effort x_1 the requirement Y is completeley examined, then $u(x_1) = (0, 0, 0, \ldots)$ and $\Psi(x_1) = 0$.

3.5 Marginal verification risk

In order to determine the effect of a specific examination effort x_j, one could elaborate $\Psi(x)$ by calculating $\partial\Psi/\partial x_j$, and use this expression for marginal verification risk efficiency with regard to the examination effort x_j of requirements in H_j, that is

$$\frac{\partial \Psi(x)}{\partial x_j} = L \cdot \frac{\partial \Phi_a^{nc}(p_{0,H}^{nc} \circ u)}{\partial u_j} \cdot \frac{\partial u_j(x_j)}{\partial x_j}. \tag{14}$$

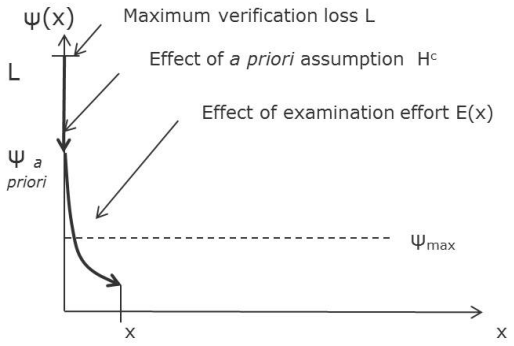

Figure 10: The verification risk measure is based on the verification loss L, the *a priori* estimated state of Y, and the effect $u(x)$ of increasing examination effort x by the given examination methods.

For complex systems or requirements consisting of many subrequirements (for instance related to different subsystems or components $\{A, B, C \ldots\}$), the expression $\partial\Psi(x)/\partial x_j$ can be further elaborated in order to find the marginal verification risk (also called the Birnbaums measure (Rausand and Høyland, 2004)) for examination of a given requirement.

Verification and examination management is now to determine the sequence of examinations that should be carried out among the main requirements H_j before concluding the outcome of verification. The sequence of examinations can be decided by selecting the requirements that achieve largest risk reduction in $\Psi(x)/\partial x_j$ (steepest decent) for a given examination effort x_j.

3.6 Stop examination criteria

The verification management model handles verifications that potentially may contain large quantities of effort. In order to limit the examination effort, the model proposes a criterion for stopping the examination and concluding the verification. Two examples of possible stop criteria are illustrated in Figure 11.

The first criterion Z_1 illustrated in Figure 11 is related to the marginal change of the verification risk function. For example, Z_1 ='true' if $\partial\Psi(x)/\partial x_j > -1$ for a given method and corresponding effort x_j then the examination should be stopped. This means that the number -1 is an example threshold indicating when the marginal verification risk reduction is less than or equal to the marginal examination effort x_j.

The other proposed criterion Z_2 is related to the achieved level of verification risk. For example, Z_2 ='true' if $\Psi(x) \leq 20\ 000$ after an effort x, where the number 20 000 is an example threshold to be selected.

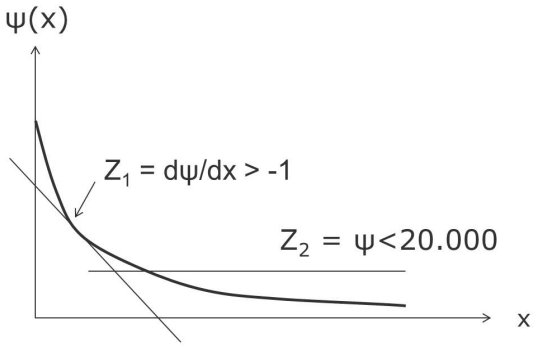

Figure 11: Examples of examination stop criteria Z_1 and Z_2, where Z_1 dictates a stop when the marginal value of verification risk $(\partial\Psi/\partial x)$ reduction is lower than the marginal examination effort ∂x, and Z_2 dictates a stop when the overall verification risk is below the value 20.000.

4 Case studies

4.1 Case 1: Criteria for performing examination of forklift

We return to the illustrating example of Section 2.2, where a forklift shall either be accepted or rejected by the buyer. The forklift shall be accepted if the single requirement Y is complied with. Assume that there is given *a priori* probability $p^c_{0,Y} = P(Y)$ that the requirement Y is complied with, or $p^{nc}_{0,Y} = P(\neg Y)$ that the requirement is not complied with.

The buyer is offered the possibility to carry out verification and examination at cost x_1, where this examination will clarify for certain if the requirement is complied with or not; see Figure 3. The value of the forklift is L. In the case that the buyer accepts the forklift, he has to pay the value L. The verification management question is to decide whether the examination and verification shall be carried out at an effort of x_1; see Figure 7.

In this case we have $\Phi^{nc}_a(p^{nc}_{0,Y}) = p^{nc}_{0,Y}$ and

$$\Psi(x) = L \cdot p^{nc}_{0,Y} \cdot u(x). \tag{15}$$

Let the examination stop criterion Z ='true' be to stop examination when the examination effort exceeds the verification risk, that is, $Z = \{x_1 > \Psi(x)\}$. Inserting the expression (15) for the verification risk gives the stop criterion

$$Z = Z(x) = \left\{x_1 > L \cdot p^{nc}_{0,Y} \cdot u(x)\right\}. \tag{16}$$

Assume that *a priori* it is a 10% probability that the requirement Y is noncompliant, that is, $p^{nc}_{0,Y} = P(\neg Y) = 0.1$. Initially, we then get the condition on examination effort $(x_1 \leq 0.1L)$ for performing the examination. This means that if the cost of performing the examination is less than 10% of the cost L of the machine, then the buyer should decide to perform the examination of Y.

4.2 Case 2. Verification and examination management of a redundant system with one common component

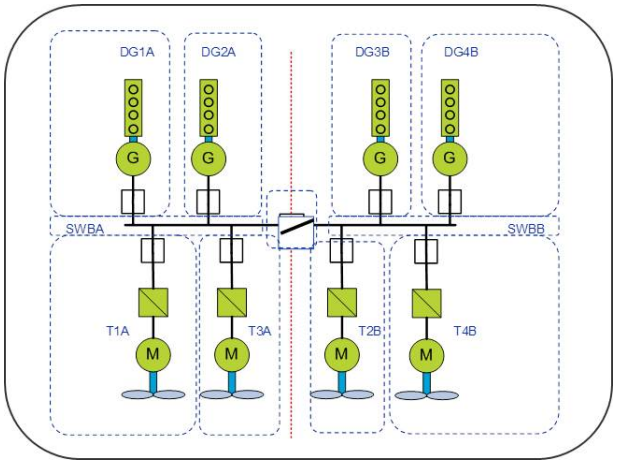

Figure 12: A DP vessel is arranged with 4 trusters and 4 diesel-generators. The A-side (port) and B-side (starboard) consist of switchboard $SWBA$ and switchboard $SWBB$, each with 2 connected diesel-generators and 2 thrusters, respectively. The switchboards $SWBA$ and $SWBB$ are connected with a bus-tie breaker X, labeled subsystem C.

We are given a redundant power generating and thruster system as indicated in the block diagram in Figure 12. The requirement to this system may be described by

$$Y = (H_A \vee H_B) \wedge H_C, \tag{17}$$

and by de Morgan's theorem we get the negated requirement

$$\neg Y = (\neg H_A \wedge \neg H_B) \vee \neg H_C. \tag{18}$$

The probability that the requirement is noncompliant becomes

$$\begin{aligned} p^{nc}_{0,Y} &= P(\neg Y) = P\left((\neg H_A \wedge \neg H_B) \vee \neg H_C\right) \\ &= P(\neg H_A) \cdot P(\neg H_B) + P(\neg H_C) \\ &\quad - P(\neg H_A) \cdot P(\neg H_B) \cdot P(\neg H_C) \\ &= p^{nc}_{0,H_A} p^{nc}_{0,H_B} + p^{nc}_{0,H_C} - p^{nc}_{0,H_A} p^{nc}_{0,H_B} p^{nc}_{0,H_C} \\ &=: \Phi^{nc}_a(p^{nc}_{0,H}), \end{aligned} \tag{19}$$

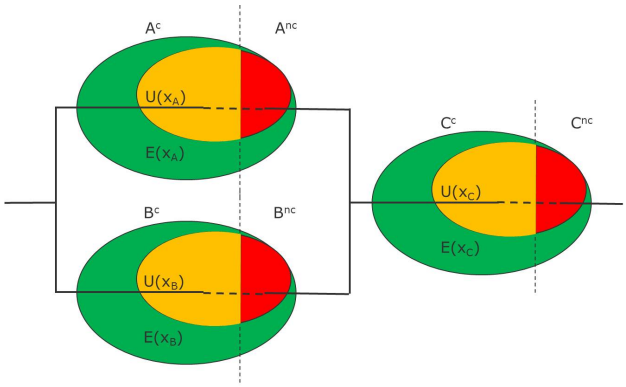

Figure 13: Reliability diagram of complex system $(A \vee B) \wedge C$. The *a priori* assumptions for the compliant parts and the unexamined parts for each subsystem (or subrequirements) are indicated. The verification and examination management objective is to estimate the individual x_A, x_B, x_C examination efforts in order to fulfill the stop criterion (Z).

where

$$p_{0,H}^{nc} = \begin{bmatrix} p_{0,H_A}^{nc} \\ p_{0,H_B}^{nc} \\ p_{0,H_C}^{nc} \end{bmatrix} \tag{20}$$

are the *a priori* probabilities that respective requirements H_A, H_B, H_C are not initially complied with.

Assume that the verification loss is of Type II and this loss is represented by L, and let x_A, x_B, x_C be the respective examination efforts on subsystems A, B, C. Based on the above formula for $P(\neg Y)$ and (12) the verification risk now becomes

$$\begin{aligned} \Psi(x) &= L \cdot \Phi_a^{nc}(p_{0,H}^{nc} \circ u(x)) \qquad (21) \\ &= L \cdot [p_{0,H_A}^{nc} u_A(x_A) \cdot p_{0,H_B}^{nc} u_B(x_B) + p_{0,H_C}^{nc} u_C(x_C) \\ &\quad - p_{0,H_A}^{nc} u_A(x_A) \cdot p_{0,H_B}^{nc} u_B(x_B) \cdot p_{0,H_C}^{nc} u_C(x_C)] \end{aligned}$$

where it is assumed that the examinations of the subsystems A, B, C are statistically independent. The marginal verification risk (Birnbaums measure) with regards to examination efforts can now be described analytically by

$$\frac{\partial \Psi}{\partial x_A} = L \cdot p_{0,H_A}^{nc} \frac{\partial u_A}{\partial x_A} \left[p_{0,H_B}^{nc} u_B - p_{0,H_B}^{nc} u_B \cdot p_{0,H_C}^{nc} u_C \right] \tag{22}$$

$$\frac{\partial \Psi}{\partial x_B} = L \cdot p_{0,H_B}^{nc} \frac{\partial u_B}{\partial x_B} \left[p_{0,H_A}^{nc} u_A - p_{0,H_A}^{nc} u_A \cdot p_{0,H_C}^{nc} u_C \right] \tag{23}$$

$$\frac{\partial \Psi}{\partial x_C} = L \cdot p_{0,H_C}^{nc} \frac{\partial u_C}{\partial x_C} \left[1 - p_{0,H_A}^{nc} u_A \cdot p_{0,H_B}^{nc} u_B \right]. \tag{24}$$

The stop examination criterion Z is defined as

$$Z = \{\Psi(x) < 50\ 000\}, \tag{25}$$

and the verification loss is $L = 1000\ 000$ (NOK). Let the *a priori* assumptions of compliance (c) and non-compliance (nc) of the requirements H_A, H_B, H_C be

$$p_{0,H_A}^{c} = 0.8, \qquad p_{0,H_A}^{nc} = 0.2 \tag{26}$$

$$p_{0,H_B}^{c} = 0.8, \qquad p_{0,H_B}^{nc} = 0.2 \tag{27}$$

$$p_{0,H_C}^{c} = 0.9, \qquad p_{0,H_C}^{nc} = 0.1, \tag{28}$$

and assume that the subsystems, requirements, and examinations for A, B, C are statistically independent.

The examination characteristics functions are all assumed expressed as

$$u_j(x_j) = \frac{1}{1 + x_j}, \qquad x_j \in [0, \infty), \tag{29}$$

with $j = A, B, C$, where x_j for instance corresponds to days of examination. This gives

$$\frac{\partial u_j}{\partial x_j} = \frac{-1}{(1 + x_j)^2}, \tag{30}$$

which is inserted into the verification risk and marginal verification risk equations for the A, B, C subsystems.

The model has been implemented in an Excel spreadsheet, which is used for proposing the examination efforts and the sequence of the efforts. Table 1 shows the sequence of the examinations $\{x_C, x_A, x_C, x_B\}$ selected on the basis of the marginal verification risk for the requirements. The examination is stopped when the verification risk is below 50 000 as $\{Z = 43000 < 50000\}$ satisfies the given stop examination criterion. At this stage the verifier will make the verification decision either to accept or reject the Y requirement.

When the stop criterion is reached, the result is a proposed examination effort of $x_A = 1$ day on the A-system, $x_B = 1$ day on the B-system, and $x_C = 2$ days on the C-system, in total 4 days of examinations.

5 Results and discussion

This paper reports the expressed original industry need for examination and verification management as formulated in 2011. In the RCN research projects "D2V" (RCN project no. 210670) and "Arctic DP" (RCN project no. 199567) a conceptual examination and verification management (VM/EM) model was proposed in the years 2012-2014 and presented in this paper.

The main properties of the VM/EM model are:

Table 1: Verification and examination management applied to find an effective sequence of examination for the subsystems A, B, C. The table shows the proposed sequence of examination calculated from a steepest decent approach. The red numbers illustrate the information used for selecting examination. The resulting examination effort becomes: $(x_A = 1, x_B = 1, x_C = 2)$.

Step	x_A	$\frac{\partial \Psi}{\partial x_A}$	x_B	$\frac{\partial \Psi}{\partial x_B}$	x_C	$\frac{\partial \Psi}{\partial x_C}$	$\Psi(x)$
0	0	-36000	0	-36000	0	-96000	136000
1	0	-38000	0	-38000	1	-24000	88000
2	1	-9500	0	-19000	1	-24500	69000
3	1	-9667	0	-19833	2	-10888	52666
4	1	-4833	1	-4833	2	-11000	43000

- The VM/EM method is based on a holistic top-down approach and based on examination of combinations of subsystems. A top-down recursive method to any level of detail of subsystems is presented.

- Based on possible consequences of Type II errors, the concept of 'verification risk' has been proposed as the key parameter to be reduced by means of examination efforts. Verification risk is to be used as the main parameter for selecting and managing examination efforts.

- The VM/EM method establishes relations between the initial verification loss, the complex system design topology, verification method performance, and the required examination effort to achieve the examination stop criterion.

- Establishment of u-functions for describing relation between the examined parts of a requirement and the examination effort. An assumption made is that different u-functions may have different forms, and that some barriers might be more efficient to examine than the others in order to reduce the overall verification risk with lowest use of examination resources.

- The VM/EM method allows for complex system descriptions using standard Boolean and reliability methods, such as de Morgans theorem, Birnbaums measure, reliability block diagrams (RBD), event tree analysis (ETA), and fault tree analysis (FTA).

- Requirement for an examination stop criterion to clarify when examination should be stopped.

- The Excel implementation model is quite simple and intuitive, although the background analytical equations for a medium size system might be quite large and complex.

The paper provides two simple verification management cases based on an Excel implementation. The cases demonstrate how the verification and examination management process of examination efforts may work in a general and in a practical manner. In future, the steepest decent algorithm to find the sequence of examinations should be implemented in an optimization-based framework.

The model requires that the user must be more explicit on specification of the verification task to be carried out, compared to a traditional planning of examination efforts. These issues are normally taken implicitly into consideration in the approval or verification processes. However, in the proposed verification risk model it is required to be more explicit on topics like:

- Potential verification consequence loss L.

- Specification of complex requirements through the Boolean structure function $Y = \Phi(H)$.

- Examination stop criterion Z on how accurate and how much cost or effort that should be spent on reaching the examination result and then make the verification decision.

The model allows for scaling up to a large number of components and subsystems and their corresponding requirements. This will also make it possible to select parts of the model that could be modeled in more detail.

Examination and verification management as indicated in this paper is mathematical and will produce the same result every time. However, the initial assumptions of *a priori* values and u-functions will be based on various types of considerations that could result in different results from time to time.

The model seems to be robust, and the resulting distribution of examination efforts in the different subsystems will probably depend on the component position in the system topology. The design of the system topology strongly influences the efficiency of examination efforts in the different subsystems.

In the approach described in this paper, the negative concept 'verification risk' has been chosen to be the key concept for selecting examination efforts. In the very beginning of this development process, the positive concept of 'verification benefit' was proposed as the key concept. However, at a certain stage the conformity to audit risk concepts motivated the verification risk concept, and the mathematical expressions for verification risk was perceived to be easier to handle in calculations than the expressions for confirmation and benefit.

Complex systems often have a high number of system operational modes that should be verified. Such advanced modeling could be included in future development of the models.

The estimation method demonstrated in this paper is based on a manual selection method of one single step. It is obviously possible to carry out the estimation with more automatic methods and this should be considered in the future when more basic knowledge about the prerequisites discussed above have been elaborated and justified.

6 Conclusion

The main result of the work reported in this paper is a proposal of a conceptual framework for verification and examination management as requested by the verification industry in 2010-2011. The main requested issue at that time was to manage how much examination and verification effort should be allocated on the various parts of a system and how the sequence of such efforts should be distributed on the subsystem parts. Another important issue was to estimate the effect of different verification and examination methods. The main lesson learned by the work is that it is possible to model and organize the selection of the examination effort for complex systems in the proposed manner with the given examination and verification risk management assumptions.

The VM/EM modeling is a simple and flexible approach that is similar to existing models for expressing reliability and risk. An extended verification risk model may make it possible to model a wide range of system requirements and system topologies.

The conceptual model clarifies and specifies key concepts regarding verification and examination management. The verification risk concept also has similarities to the audit risk method applied in planning of financial auditing.

Complex systems or operations have complex logic relations between subsystems and subfunctions. These relations have effects on the overall system properties (e.g. reliability, availability, restoration time, risk, etc.) The relations also have effects on the required examination efforts that may be needed in order to achieve an overall goal on verification risk at system level.

Acknowledgments

Aleks Karlsen and Karl Hovden in the section for approval of Control Systems in DNV GL have provided valuable input and comments from the industry to clarify the need for verification and examination management. We appreciate the constructive comments to the paper from Torbjørn Skramstad (DNV GL, NTNU) and Nils Albert Jenssen (Kongsberg Maritime) during the development and review of the text.

Research partly funded by Research Council of Norway (RCN) project no. 199567: KMB "Arctic DP", with partners Kongsberg Maritime, Statoil, and DNV GL, partly by RCN project 203471: CRI SAMCoT, and partly by RCN project 210670: KPN D2V.

Disclaimer

The results in this paper are the result of research projects at NTNU and carried out by the authors. NTNU and the sponsoring companies have accepted the publication of the paper; however, NTNU nor the sponsoring companies cannot be held liable for any use or reference to this paper.

References

AICPA. Audit risk and materiality in conducting an audit. Statements on Auditing Standards AU sec. 312 (SAS No. 107), American Inst. Certified Public Accountants, 2006. URL http://www.aicpa.org/research/standards/auditattest/pages/sas.aspx#SAS100. Internet, visited 2014/11/13.

Arens, A., Elder, R., and Beasley, M. *Auditing and Assurance Services: An Integrated Approach*. Pearson Prentice Hall, 2006. URL http://books.google.no/books?id=JcWuHAAACAAJ.

DNV. Failure Mode and Effect Analysis (FMEA) of Redundant Systems. DNV-RP-D102, DNV-GL, 2012. URL https://exchange.dnv.com/publishing/codes/download.asp?url=2012-01/rp-d102.pdf. Internet, visited 2014/11/13.

IEC. Functional safety of electrical/ electronic/ programmable electronic safety-related systems. IEC 61508, Int. Electrotech. Comm., 2010. URL http://www.iec.ch/functionalsafety.

NIST/SEMATECH. *e-Handbook of Statistical Methods*. 2014. URL http://www.itl.nist.gov/div898/handbook. Internet, visited 2014/11/18.

Rausand, M. and Høyland, A. *System Reliability Theory: Models, Statistical Methods, and Applications*. John Wiley & Sons Ltd, New Jersey, 2 edition, 2004.

Skjetne, R. and Egeland, O. Hardware-in-the-loop testing of marine control systems. *Modeling, Identification and Control*, 2006. 27(4):239–258. doi:10.4173/mic.2006.4.3.

Skjetne, R. and Sørensen, A. J. Computer-based systems on ships and offshore vessels: The software problem ++. Report, Marine Cybernetics AS, Trondheim, Norway, 2004. Joint report by Marine Cybernetics, DNV, Statoil, Smedvig Offshore, Kongsberg Maritime, Norsk Hydro, Stolt Offshore, Eidesvik, Subsea 7, Solstad Offshore, Ulstein, ABB, PSA Norway, and Norwegian Maritime Directorate.