Dynamic Monitoring of Safety Barriers in Petroleum Installations

Aida Omerovic & Atle Refsdal
SINTEF, Norway

Øyvind Rideng
Oilfield Technology Group, Norway

ABSTRACT: Accidents on petroleum installations can have huge consequences, resulting in loss of life, environmental damages as well as economic loss. A number of so called safety barriers are therefore from earlier implemented with the objective of reducing the risk. In order to assess the quality and risk level, a proper understanding of the ability of the barrier systems to perform as intended, is needed. However, due to the complexity of the barrier systems, this ability may depend on a multitude of technical and human factors. Furthermore, it may quickly change over time. In order to be able to perform corrective and preventative measures, early warnings should be captured and their implications interpreted. We argue that measurable indicators can be identified and aggregated, in order to calculate overall quality of a barrier system. Thus, the indicators can be exploited in a monitoring environment for purpose of predicting significant change of quality level, as well as for validation of the quality requirements. In this paper we present an approach to facilitate design of indicators for automated monitoring of the quality of safety barrier systems in petroleum installations. We moreover report on results and experiences from applying this approach in an industrial case study with a petroleum operator. The approach applied consists of a process and a tool-supported modeling language. The approach relies on relevant parts of PREDIQT and CORAS methods for quality prediction and risk analysis, respectively. The evaluation indicates that the approach facilitates development of an algorithm for monitoring barrier system quality for a given installation. The experiences from the case study moreover show that the presented approach is, to a large degree, well suited for its intended purpose, but it also points to areas in need for improvement.

1 INTRODUCTION

Accidents on petroleum installations may result in loss of life, huge environmental damages, and economic loss (Robertson & Krauss 2013). Such installations therefore implement a number of so called safety barriers in order to reduce the risk. A safety barrier is basically a set of measures directed towards a common goal of either reducing the likelihood of an initial triggering incident to occur at all, or preventing such incidents from escalating into a major accident. For example, an inspection and maintenance program may reduce the likelihood of a gas leak; a gas detection system may provide early warning in case a leak occurs; while an ignition source shutdown system may prevent the gas from igniting and leading to a major fire or explosion.

A proper understanding of the quality of the barrier systems, i.e. the degree to which they can be expected to perform as intended, is essential in order to assess the risk of accidents. However, barrier systems may be very complex and depend on many different technical and human factors. Quality by design is not sufficient, since human factors such as usage and compliance with routines may have a considerable impact on safety. Moreover, reliability of barrier systems cannot be guaranteed due to e.g. system deterioration. Analyzing the quality of the systems is therefore difficult, time-consuming and costly. Moreover, the degree to which each component is able to fulfill its role may quickly change over time. Hence, it is not feasible to maintain an up-to-date view of barrier quality based purely on manual analysis. Automated support for monitoring barrier system quality is therefore needed, in order to ensure the needed quality and frequency of the input. Although a lot of low-level data related to the current state of a specific part or aspect of a barrier system can be collected from any given petroleum installation, the real challenge is to transform this data into useful information that can be easily understood by human operators and decision makers. How this should be done will of course depend on the installation in question, as the implemented barrier systems and available data will differ between the installations.

This paper presents an approach to facilitate design of indicators for automated monitoring of the quality of safety barrier systems on petroleum installations, and reports on experiences from applying this approach in a realistic industrial case study with a petroleum operator. The approach consists of a process and a tool-supported modeling language to develop an algorithm for monitoring barrier system quality for a given installation. The aim is that the outputs of this algorithm can be presented to human operators in a suitable interface, thereby serving as a useful support for decision making. However, the presentation and interface to operators during run-time is out of scope for the approach we present here.

The approach applied in this study is heavily based on the PREDIQT method for model-based prediction of impacts of architectural design changes on system quality (Omerovic, 2012). While the PREDIQT method aims to support prediction of effects of architectural design changes on quality, based on automatic, semi-automatic and manual input, in this work we have focused on supporting automatic monitoring of all factors relevant for quality of a safety barrier. Moreover, PREDIQT has not so far been applied in the petroleum or safety domain. Compared to PREDIQT, the process we have applied is simplified and slightly adapted.

Moreover, the so called prediction models have been developed to varying degree: design models have been partially developed; no quality models have been developed; and Dependency Views have been developed in full scale. In addition, we have used the so called asset diagrams and threat diagrams from the CORAS method (Lund et al. 2011) to model risks to the barrier systems as well as to partially support the identification of the measurable indicators.

The system owner (i.e. the petroleum operator who was the case study provider) required confidentiality with respect to the results obtained. Thus, this paper reports mainly on the experiences obtained, describes the process undergone, the evaluation results, and the properties of the artifacts. The reported experiences and results have however provided valuable insight into the strengths and weaknesses of the approach.

The case study was conducted in the year 2012. The different parts of the PREDIQT method (the process, the tool, the modeling approach, and the traceability approach) were applied to various degree. The approach to uncertainty handling in PREDIQT (Omerovic & Stølen 2011b) was not applied in this study. The CORAS asset diagrams and CORAS threat diagrams have been applied too. In addition, the approach is assessed through a post-analysis review. All models were developed during the analysis and the entire target system (within the predefined scope) was analyzed. The analysis was performed in the form of five physical workshops and four intermediate (videoconference/teleconference) meetings in a fully realistic setting in terms of the scope, the objectives, the process, the prediction models and the participants.

The rest of this paper is organized as follows: In Section 2 we briefly present the research method and characterize the needs through a set of success criteria. The approach proposed is presented in Section 3. The instantiation of the approach in a case study is outlined in Section 4. The results of the evaluation are summarized in Section 5. In Section 6 we discuss the results, before concluding in Section 6.

2 RESEARCH METHOD

The objective of the study has been to propose and evaluate a method which facilitates monitoring of safety barrier quality. We have based our research method on the methodology presented by Solheim and Stølen (2007). This is an iterative method consisting of three steps, as illustrated by Figure 1.

The first step is the problem analysis, in which the researchers map a potential need for a new or improved artifact by interacting with potential users and stakeholders. The second step is the innovation, where the researchers try to identify an existing, or develop a new artifact that satisfies the needs characterized in the first step. The overall hypothesis will be that the artifact satisfies this need. Finally, the third step is the evaluation, where the researchers investigate whether the artifact actually satisfies the need. If this is found not to be the case, then a new iteration of the cycle must be initiated.

A number of different research strategies may be used for this investigation. A good overview is given by McGrath (1984). The strategies chosen for our evaluation can be classified as 1) a field experiment, as the researchers took part in implementing the procedure resulting from the innovation step and included their experiences and observations from this process in the evaluation, and 2) a judgment study, as responses where collected from the other participants by the use of questionnaires.

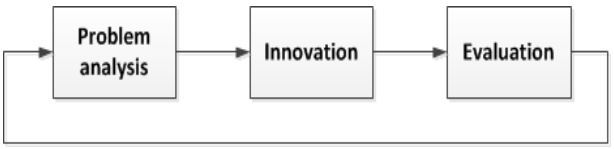

Figure 1. The research method applied. The figure is adopted from Solheim & Stølen (2007)

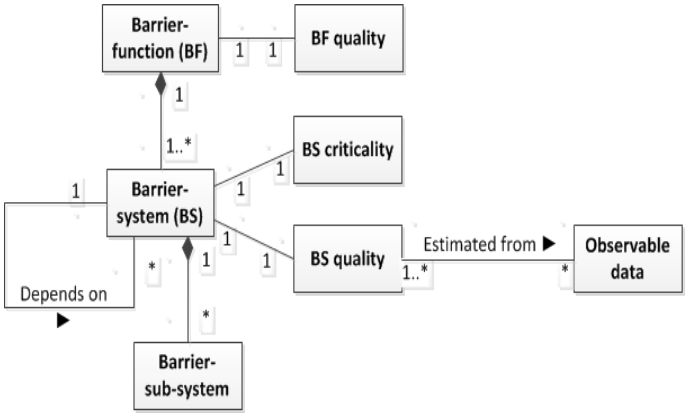

Figure 2. A conceptual model of the target domain

2.1 *Characterization of needs*

The initial step in the characterization of the needs was establishing a common understanding of the terminology and concepts. The initial meetings included presentation of the problem domain, the relevant procedures and the concepts. Three organizations have been involved: SINTEF as the research partner, Oilfield Technology Group (OTG) as the observer and advisor, and the petroleum operator as a case study provider. As a part of this step, a conceptual model of the target domain was developed, as shown by Figure 2. As illustrated, a barrier function consists of one or more barrier systems, while a barrier system may consist of several barrier sub-systems that contribute to fulfilling the function. Each barrier function is annotated with a quality rating, while each barrier system is assigned a level of criticality and a level of quality[1]. The latter is estimated from the observable data.

The problem analysis was carried out in cooperation between the three participating organizations. In the first meeting with representatives for all three participating organizations, the emphasis was on establishing a high-level understanding the needs of the petroleum operator and introducing the competences and technologies available in the participating organizations. After this meeting, the representatives of the petroleum operator prepared a proposal for a suitable case, focusing on their need to be able to monitor

[1] The industry has also established the term "performance" (of safety barriers), which is used for the same purpose as presented in this paper. In this paper, we solely use the term "quality", due to its broad meaning and in order to distinguish from the other (often more specific) definitions of performance.

the quality of safety barriers. These barriers are categorized into five different barrier functions. Figure 3 illustrates the decomposition of the barrier function "Prevent ignition" into barrier systems.

For each barrier system, there are a number of different aspects that influence its quality. Obtaining an overall view of the quality of barriers is therefore very difficult. The petroleum operator now performs periodic analysis of the barriers. This is a labour-intensive process not suited for capturing factors that may change from day to day. They therefore wanted support for automatic monitoring of barrier quality to supplement their existing analysis procedures. A lot of relevant data can be collected from installations, but the challenge is to develop an algorithm for turning these data into an overall quality assessment that can be understood by human operators.

In order to ensure that the approach and the results are generally applicable also for other petroleum operators and installations, the artefact to be developed would not be just one specific monitoring algorithm. Instead, we aimed for an approach that could be implemented by any petroleum operator to develop barrier quality monitoring algorithms tailored to their own installations, barriers and available data. This approach would include a procedure to develop such algorithms and modelling languages to support the process. In the following we characterize the needs in the form of success criteria for the approach.

1. All models can be easily understood by all involved actors.
2. The modelling languages have sufficient expressive power to capture all aspects that the domain experts consider relevant.
3. The application of the procedure results in an implementable monitoring algorithm.
4. When fed with correct input data, the resulting monitoring algorithm provides a correct evaluation of barrier system quality.
5. The approach is cost-effective, i.e. the benefits are well worth the effort.
6. The approach is sufficiently general to be applicable for all petroleum installations.

3 THE APPROACH TO ESTABLISH BARRIER EVALUATION ALGORITHM

The approach proposed prior to the case study is illustrated on Figure 4. The approach consists of four overall phases, and each one is divided into steps.

The objective of Phase 1 is to characterize the target of the analysis. Step 1 takes the characterization and the scope of the target of the analysis as input, and produces models of the target. In this step, the system is specified and modelled w.r.t. structure, dataflow, workflow, etc. The existing system models may be reduced. The models also specify the scope and the frames of the analysis. Step 2 takes the characterization of the target and the system models as input, and defines the notions of quality as output. The objective is to specify the quality concepts with respect to the target in

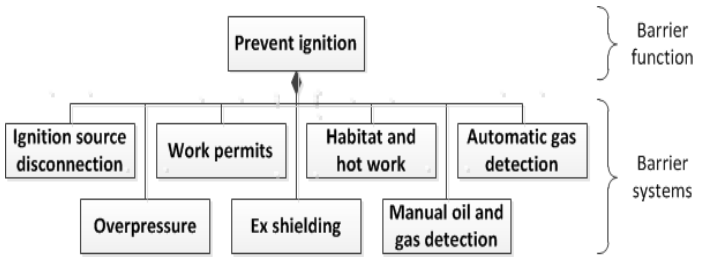

Figure 3. Barrier systems that contribute to fulfilling the barrier function "prevent ignition"

the form of a decomposed PREDIQT Quality Model, where each notion is defined qualitatively and quantitatively.

The objective of Phase 2 is to identify the indicators which have an impact on the quality. Basically, indicators are statements or functions that give a clue about the value of a model element/parameter. The values of indicators can be retrieved empirically and mapped to the relevant model elements. Step 1 takes the specification of system and quality form Phase 1 as input, and identifies threats on quality. It also annotates the threats with the indicators. The approach is as follows: the threats on quality are specified, before they are annotated with indicators. Step 2 takes the specification of system and quality, as well as the specification of threats on quality as input, and produces a dependency graph or tree, e.g. PREDIQT Dependency View (DV), w.r.t. quality. The Dependency Views are annotated with retrievable indicators which influence the quality. The indicators may overlap with the ones identified in relation to threats in Step 1. The modelling approach is as follows: develop models which show the dependencies of the system w.r.t. quality (as defined in Step 2 of Phase 1). The relevant elements of the models are annotated with indicators. The prior parameters of the models are estimated and propagated according to the propagation algorithm of the modelling approach. Note that the rationale for distinction between the indicators and the prior parameters of a model, is that the former may be related to model elements in a manner which is not supported by the modelling notation (e.g. due to mapping functions or repetitiveness), or that the indicators should have special treatments (e.g. assumptions) which need to be documented in the indicator specifications. The treatment of indicators and other kinds of trace-link information is in PREDIQT solved through the traceability approach (Omerovic & Stølen 2011a). The PREDIQT traceability approach however covers much broader needs for traceability beyond indicators, while this study has narrowed down the trace-link information to indicators only. We have identified the information needs for specifying the indicators in the petroleum domain, in addition to evaluating the feasibility of identifying the relevant ones and providing a useful quality assessment from a limited number of the indicators.

Phase 3 focuses on detailed specification and evaluation of the indicators. It also defines the relationship between the indicators and the elements of dependency models. This phase moreover specifies the algorithm for calculating the overall quality. Step 1 takes as input the threat specifications and the dependency models (developed in the previous phase), both annotated with indicators. The output of this step is specification of each indicator in the form of qualitative and quantitative interpretation, guidelines for data retrieval, units of measure and the assumptions. The approach is as follows: for each previously identified indicator, speci-

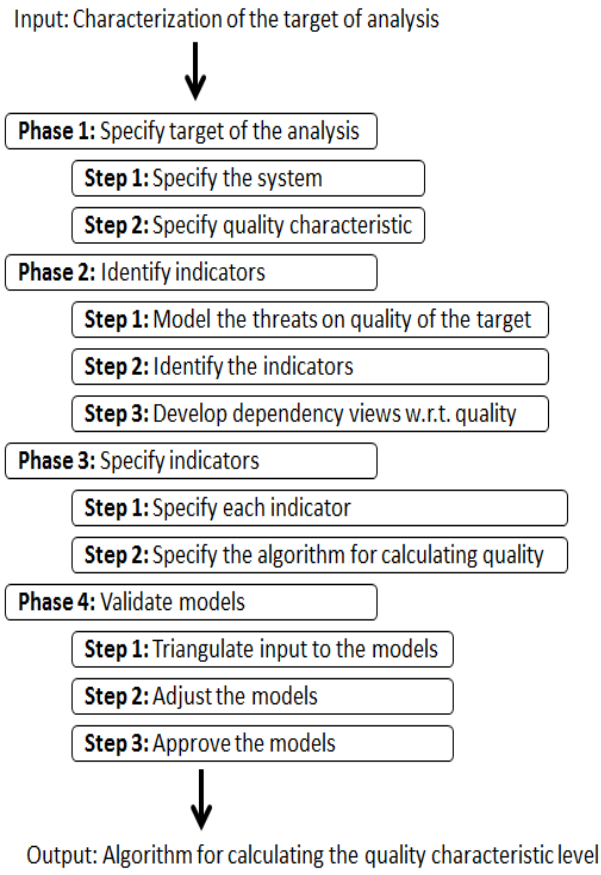

Figure 4. The process proposed prior to the case study

fy the properties, the mapping to the overall models, the values and the assumptions using a specification template. Step 2 takes the threat specifications, the dependency models and the indicator specifications as input, and produces an algorithm for calculating the overall quality level as output. The algorithm is deduced in the form of a function, based on the structure of the dependency models, the parameters of the dependency models, the propagation model of the approach applied for modelling the dependencies, and the indicator specifications.

Phase 4 aims at validating the models developed in the preceding phases. That is, the models should be exposed to various kinds of evaluation in order to ensure an acceptable level of uncertainty. The uncertainty may origin from insufficient information or knowledge, or from variability in context, usage, etc. Such factors influence validity and reliability and should therefore be explicitly expressed in the form of the uncertainty value as well as tolerance level. The uncertainty handling has not been included in this particular study, but the quality and particularly the validity of the models is addressed through the following three steps of Phase 4. Step 1 takes the dependency models including both the indicator specifications and the aggregation algorithm as the input, and provides the difference between the empirically measured and the initially estimated (and propagated) values, as output. During this strep, the overall quality level is compared and the deviation is considered. Step 2 takes the deviation as input and provides the updated models as the output. During this step, the models are adjusted according to the deviations revealed in Step 1. Based on consistency checks, the structure (including the mapping of the indicators) and the parameters of the models are modified. Consequently, also the aggregation algorithm for calculating the overall quality level is modified. The modified models are instantiated and the deviation evaluated again. In Step 3, the models are approved if the evaluation shows that the modified models are sufficiently complete, correct and certain.

4 THE INSTANTIATION OF THE APPROACH IN A CASE STUDY

This section outlines the process undergone during the case study, as well as the main properties and examples of the modelling artefacts produced. Note that the process undergone is to a large degree but not entirely an instantiation of the approach presented in Section 3. The reason for not covering the entire approach as planned is that the process has been revised during the analysis in order to meet the most prevailing needs and to apply as much as possible of the approach, within the limited resources assigned.

Table 1 summarizes the process undergone. For each workshop or intermediate meeting, we list the meeting number, the date, the participants, the meeting type, the meeting length, tasks, preparations, input and output.

Table 1. The process undergone during the case study

Meeting 1**; Date:** 19/6**; Participants:** 2 analysts, 2 domain experts, 2 observers**; Meeting type:** physical**; Meeting length:** 4h; **Tasks:** Identification of context, goals, scope and focus. The analysts presented their high-level understanding of the target system and customer goals. Establishing a common conceptual model. Deciding practical issues such as how to exchange confidential documents and fixing dates for meetings. **Preparations:** The customer prepared a high-level proposal for a case description and sent this, as well as documentation of the target of analysis, to the analysts. The analysts prepared a conceptual model, as well as their understanding of the case and some concrete questions for the customer, for presentation during the meeting. **Input:** Proposal for high-level case description and target description provided by the customer. Conceptual model proposed by the analysts. **Output:** Conceptual model.
Meeting 2**; Date:** 22/6**; Participants:** 2 analysts, 1 observer, 3 domain experts**; Meeting type:** video**; Meeting length:** 2h; **Tasks:** Ensuring that the analysts get a better and more detailed understanding of the target system. **Preparations:** The video conference was structured around a list of 17 questions that the analysts had prepared before the meeting. **Input:** Questions prepared by the analysts about the target system. **Output:** Notes with replies to the questions. List of further target documentation to be made available for the analysts.
Meeting 3**; Date:** 1/8**; Participants:** 2 analysts, 1 observer, 5 domain experts**; Meeting type:** video**; Meeting length:** 5.5h; **Tasks:** Finalize conceptual model, establish CORAS asset diagram, present detailed models of the target as understood by the analysts, introduce the CORAS risk modelling language, start developing CORAS threat diagrams.

Preparations: The analysts had prepared the following: updated conceptual model, CORAS asset diagram, models/descriptions of the target system extracted from the documentation received from the customer.
Input: Proposals for conceptual model, CORAS asset diagram, models/descriptions of target system.
Output: Comments and corrections to conceptual model and models/descriptions of the target system, initial and incomplete CORAS threat diagrams.

Meeting 4**; Date:** 16/8**; Participants:** 2 analysts, 1 observer, 5 domain experts**; Meeting type:** physical**; Meeting length:** 5.5h;
Tasks: Finalize models/descriptions of the target systems, develop CORAS threat diagrams and identify indicators.
Preparations: The analysts had prepared conceptual model and models/descriptions of the target system where the corrections from the previous meeting were implemented.
Input: Corrected conceptual model and models/descriptions of the target system.
Output: Accepted conceptual model and models/descriptions of the target system, CORAS threat diagrams with indicators.

Meeting 5**; Date:** 27/8**; Participants:** 2 analysts, 2 observers, 2 domain experts**; Meeting type:** video**; Meeting length:** 2.5h;
Tasks: Finalize CORAS diagrams with indicators, prepare next meeting by introducing PREDIQT dependency views.
Preparations: The analysts made minor changes to the CORAS threat diagrams from the previous meeting (structure and layout), and prepared a short introduction to the PREDIQT dependency views.
Input: CORAS threat diagrams resulting from the previous meeting, target models/descriptions.
Output: Corrected CORAS threat diagrams. A decision to measure QCF (Omerovic 2012) as a value in the interval [0,1] was also taken.

Meeting 6**; Date:** 5/9**; Participants:** 2 analysts, 1 observer, 4 domain experts**; Meeting type:** physical**; Meeting length:** 5h;
Tasks: Develop the PREDIQT dependency views including EI and initial QCF values. Present indicator specification form.
Preparations: The analysts prepared by going through some of the system documentation, but no new models or diagrams were produced before the meeting.
Input: Target models/descriptions, CORAS threat diagrams.
Output: PREDIQT dependency views with EI values and some initial QCF estimates for one of the barrier systems.

Meeting 7**; Date:** 27/9**; Participants:** 2 analysts, 1 observer, 4 domain experts**; Meeting type:** physical**; Meeting length:** 5h;
Tasks: Define algorithm for computing QCF for the combined barrier function from the QCF values for the relevant barrier systems. Initial sanity check of the first dependency view based on thought experiment. Review of dependency views developed by the customer representatives.
Preparations: The customer representatives prepared the missing PREDIQT dependency views based on the one that was made in the previous meeting. In addition, the customer representatives filled in indicator specification form for the first barrier system.
Input: PREDIQT Dependency views. One of these was prepared by the analysts and customer representatives in the previous meeting, and the rest were prepared by the customer representatives after the meeting.
Output: Fitted PREDIQT dependency views.

Meeting 8**; Date:** 31/10**; Participants:** 2 analysts, 2 observers, 5 domain experts**; Meeting type:** telco**; Meeting length:** 3h;
Tasks: Prepare for validation of quality algorithm (i.e. PREDIQT dependency views incl. indicator specifications) in the next meeting. Make adjustments and corrections to the indicator specification forms.
Preparations: The customer representatives prepared a qualitative interpretation of the QCF scale by dividing the interval [0,1] into non-overlapping sub-intervals and providing a natural language interpretation of each interval. They also filled in indicator specification forms for the remaining barrier systems. The analysts reviewed the indicator specification forms and identified issues for discussion. The values of all identified indicators became available based on measurements/retrievals/logs, etc.
Input: PREDIQT Dependency views and first version of indicator specification forms provided by the customer representatives.
Output: Indicator specification forms with corrections for all barrier systems.

Meeting 9**; Date:** 20/11**; Participants:** 2 analysts, 2 observers, 3 domain experts**; Meeting type:** physical**; Meeting length:** 4.5h;
Tasks: Validation of the models based on thought experiments.
In addition the participants were asked to evaluate the whole process by filling in questionnaires (note that this is not a part of the approach but a part of the evaluation of the case study).
Preparations: The customer representatives prepared test data in the form of fictitious values for basic indicator values. The analysts prepared thought experiments based on these data.
Input: Test data and prepared thought experiments.
Output: Results from thought experiments.

One of the observers who participated in the last meeting, was a consultant recently hired by the operator to assist in implementing the models involved into a tool, dedicated to their particular needs.

The barrier function "Prevent Ignition" and the underlying barrier systems were characterized as target and scope of the analysis. For modelling the target of the analysis in Step 1 of Phase 1, system models from existing documentation were reused. In addition, we developed UML class diagrams, and wrote functional descriptions and limitations of the barrier systems.

The definition of quality was in Step 2 of Phase 1 textually provided w.r.t. the barrier function on question. Interpretation of a categorical scale 1-6 for the barrier function quality, was also provided.

In Step 1 of Phase 2, one CORAS asset diagram was developed, and contained seven direct assets (quality of each barrier system) and one indirect asset (quality of the barrier function). For each asset, a dedicated threat diagram was

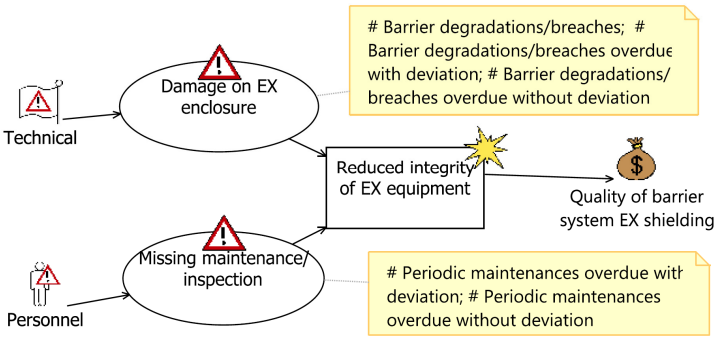

Figure 5. An extract of one of the CORAS threat diagrams.

developed in the CORAS tool. The threat diagrams were annotated with indicators. The respective threat diagrams contained the following number of indicators: 6, 12, 2, 6, 2, 7, and 5. An example of a CORAS threat diagram is shown in Figure 5.

In Step 2 of Phase 2, a PREDIQT Dependency View (DV) w.r.t. quality of each barrier system (BS) was developed using the existing PREDIQT tool (an eclipse-based DV editor). In addition to the seven BS specific DVs, a DV aggregating the BS specific DVs into a barrier function specific DV was developed. As a part of the DV development, indicators were assigned to the relevant parts of the DVs. The initial parameters of the DVs were then estimated. An illustrative example of a PREDIQT Dependency View, with fictitious values and fictitious structure, is shown in Figure 6. Values on nodes express quality characteristic fulfillment (QCF) and range between 0 and 1, while weights (EIs) on the arcs express the degree of dependency (ranging between 0 and 1) of a parent node, on its child node. Due to completeness property, the sum of weights on a subtree is 1. The propagation model is due to orthogonality and completeness properties of a DV, recursive bottom-op sum of products of related QCFs and EIs for each sub-tree. More details on DV-based modeling are presented in (Omerovic 2012).

In terms of the size of the BS specific DVs developed, the respective DVs consisted of the following numbers of nodes in total: 26, 52, 46, 14, 19, 49, and 46. The number of indicators annotated to each of the DVs, was respectively: 14, 32, 29, 8, 12, 29, and 32. These indicators were, once identified, compared with the ones identified with the CORAS threat diagrams, in order to verify completeness. The DV-related indicators were then specified using the specification template.

The indicator specification template was specifically customized for this case study and contained the following fields:

- Id: unique identifier for the indicator
- Name: a short indicator name
- Definition: qualitative and quantitative definition of the indicator and the variables/parameters. Also includes the definition of relationship between the indicator and the relevant model elements.
- Purpose: what purpose the indicator serves and which model elements it is related to.
- Measurement procedure: specifies how to retrieve the indicator values.
- Data source: specifies where to retrieve the indicator values from.
- Measurement frequency: specifies how often to retrieve the indicator values.
- Expected change frequency: specifies how often the indicator values are expected to change in reality (i.e. the dynamics of the indicator).
- Unit of measure: specifies the unit of measure for the indicator.
- Interpretation of the value measured: specifies the indicator values which are desirable, realistic but extreme, the normal area, and the edge to the unacceptable.
- Scale specifies the measurement scale for the indicator.
- Uncertainty: specifies uncertainty and the related sources of it. Can also be expressed in the form of interval, variance, etc.
- Value and measurement date: indicator value and the date of value retrieval.

The validation step was performed by comparing QCF values computed by the algorithm with values obtained from domain experts through a thought experiment. First the experts were provided with a list of indicator values or low-level QCF values as well as a resulting high-level QCF value computed by the algorithm reflecting the current situation. Then they were presented with a set of changes in indicator or low-level QCF values and asked to individually write down what they thought should be the new high-level QCF value after the change. The changes had been selected by the analysts from sets of data provided by the petroleum operator to reflect different imaginary points in time, as real historical data were not available.

During this process the experts did not have access to the models. Their task was to provide a QCF value that reflected the new state given the changes, rather than to try to predict the value computed by the algorithm. Each expert presented his/her estimate, and the experts then discussed until they agreed on a value. The values provided by the experts were then compared to the values computed by the algorithm. The deviation results from the 10 respective thought experiments were as follows: 0,016; 0.31; NA[2]; 0,52; 1,58; 0; 0,015; 0,063; 0,015; 0,089.

The deviations shown above are computed by the formula $|E\text{-}S|/E$, where E denotes the estimate from the group of experts and S denotes the corresponding value computed by the DV-based simulation. Note that this does not imply that we view the expert estimate (or the value computed by the algorithm) as the one correct QCF value that perfectly reflects the given state; both should be viewed as imperfect approximations. The formula above is used in order to take into account the varying impact of big and small changes on the high-level QCF value, as well as due to the confidentiality of the values obtained during the evaluation.

5 RESULTS OF THE EVALUATION

The evaluation of the case study was conducted in the form of a post-analysis review based on a pre-prepared questionnaire that the participants (excl. the analysts) filled out after the analysis. The questionnaire was designed w.r.t. the success criteria, and in such a manner that the responders can, if desired, be anonymous. Date of response, employer,

[2] Due to division by zero.

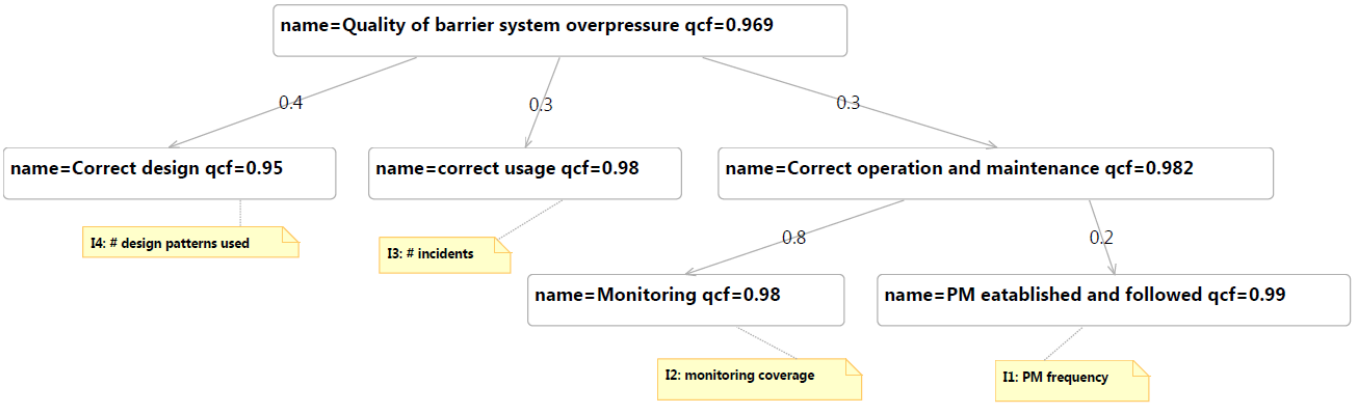

Figure 6. An illustrative example of a PREDIQT Dependency View, with fictitious values and fictitious structure

position, degree of education, years of professional experience, and role in the analysis, was also requested in the form. This section briefly summarizes the written responses received on the 9 concrete questions.

1. The method in general is well suited, well structured, comprehensible and easy to apply, at the same time as it is comprehensive and detailed. Interesting to apply a relatively established method in the oil/gas domain. The CORAS threat diagrams were less applicable then the PREDIQT Dependency Views (along with the indicator specifications), in this context. The models and the tools facilitate communication. The case study has improved the domain experts' understanding of the target and its quality.
2. The process undergone has been demanding but comprehensible. It has produced models that are easy to understand and can be used in the future. The CORAS threat diagram modeling was less suited than the PREDIQT DV modeling in this context. More data retrieval over a longer period of time should have been done. It is thorough and demands significant resources from the case study provider. It is demanding but worth the effort. The process became clearer after a while.
3. In terms of the models, the CORAS threat diagrams were much less comprehensible and more demanding than the PREDIQT DVs. The latter capture the relevant aspects and indicators. The DVs are rather complete and can become sufficiently precise if tuned over time. The models reflect the reality to a sufficient degree, but not 100%, mainly because it is generally impossible provided the limited time.
4. The analysis has provided better understanding of this BF and other BFs. This kind of analysis can be done on other installations and BFs as well. The models are comprehensible and can easily be explained to others. The models can be reused on other installations BFs. The structured process forces the participants to think about both the overall picture and the relevant details. The role of the different indicators is clear.
5. Will recommend the method to others. Will most probably work more on use of the method and present it to others within the organization. We have with 95% certainty decided to implement this method and a tool dedicated to our needs (data retrieval), within the organization. Will also use this in relation to barrier analysis. In our community (those working on quality of barriers within the organization), there is an agreement that this is a better method than what has been used so far.
6. The challenges: agree upon the weights on the DVs and the indicator specifications; complexity; data retrieval; support/follow up from the management; validity of the models; time and resources; tool support should be more available and guide on the method; integration of the tool with the overall software in the organization.
7. Benefits: better overview over the weaknesses; improved predictability; improved preparedness; improved safety; better documentation of the security and safety level; more efficient operation in the long term; easier to communicate the risk level to the platform personnel; the method provides better decision support; better understanding of complexity; a structured way of calculating the quality level of a barrier; less resource demanding due to automatized data retrieval; better barrier management; makes the barrier management dynamic.
8. Recommended improvements of the method: provide predefined DV patterns that can be adjusted; improved visualization; improve the tool support; standardize the interface to the relevant databases; provide symbols that are established within the oil/gas domain; more instantiations of the method in the oil/gas domain.
9. Overall comments: we have seen that a better visualization and communication of the risk picture in oil/gas sector, is possible. This will give long-term benefits. Positive overall impressions of the process and the role of the analysts.

In addition to the written feedback, the analysts have taken comprehensive notes of their observations as well as the time usage during the analysis. Due to the space limits, however, only the post-analysis review is reported here in the context of the evaluation.

6 DISCUSSION

The needs have earlier been characterized in the form of six success criteria (SC). In the following, we briefly elaborate, based on the evaluation results, to what degree the success criteria have been fulfilled. We also address the main threats to validity and reliability of the study.

The fact that all participants have actively participated in the analysis and agreed upon a set of models, indicated that the approach is comprehensible. Moreover, once one of the DVs had been developed during a workshop, the domain experts were able to independently provide the rest of the BS specific DVs. The indicator specifications were also provided by the domain experts independently from the ana-

lysts. These facts, in addition to the responses from post-analysis review, indicate that SC1 is, to a relatively high degree, fulfilled.

Regarding the expressive power of the modeling languages, the feedback received indicated that the domain experts were able to cover all relevant aspects and indicators in the DVs. During the workshops, all the concerns mentioned were possible to include in the DVs. One missing dimension that should have been included, is the PREDIQT Quality Model, in order to have a better common foundation for understanding and interpreting the QCF values on the DVs. The scale that was defined at meeting 8 had two weaknesses: it appeared late in the analysis, and it was a simplified specification trying to express both the interpretation and the acceptance levels of the QCF values.

The PREDIQT DVs include a defined syntax and semantics of the parameters, a propagation model and indicator specifications. The existing PREDIQT tool already includes automatic propagation. The PREDIQT traceability tool includes support for all needed trace-link information, including the indicator specification. As such, the propagation algorithm is fully implementable.

In terms of SC4, the validation performed in the last meeting showed too large a deviation between the DV-based simulations and the thought experiment based estimates. As such, the validation was not completed – only demonstrated. Further adjustment of models is necessary. The operator has expressed that, in spite of this, the possibility of observing the trends is useful. Moreover, thought experiments have obvious weaknesses compared to real data. Thus, SC4 is in the current state of the models not satisfied.

The case study has indicated that the approach is feasible in a fully realistic setting and within the limited resources allocated. The responses from post-analysis review also indicate that the analysis was well worth the effort. The operator has strong intention to implement the approach and dedicated tool support within the organization, which indicates the usefulness of the approach. They also intend to preserve much more relevant historical data then what has been the case so far. Although PREDIQT has been applied in multiple case studies before, this analysis was customized for this particular study. Moreover, CORAS asset and threat models were also developed (although post-analysis review indicates that this perhaps was not fully necessary). Hence, we expect less resource to be needed the second time this approach is applied in a similar domain. There is also a need for a baseline for comparing this approach with the alternative ones, in order to assess its cost-effectiveness. It should be a part of the future work.

SC6 is rather difficult to elaborate on, before more instantiations of the approach are performed on other petroleum installations and other BFs. The structure of the originally developed DV for one BS was quite general and was to a high degree reused on the overall BSs, but we need more such analyses in order to evaluate the generality of the approach.

Full documentation of the case study exists, but its availability is restricted due to confidentiality required by the customer. Hard evidence in the form of measurements to validate the correctness of the predictions would have been desirable, but this was unfortunately impossible within the frame of this case study. Instead, we have relied on extensive documentation and the domain expert group with solid background and diversity. Still, thought experiment-based validation of models has weaknesses compared to the measurement-based ones. Particularly, we cannot exclude that possible undocumented or inconsistent assumptions have been made in model development, although the active participation of the domain experts in all model development should prevent this. Statistical power was limited, due to low number of participants. The careful selection of experienced participants and the variety of the changes specified during model validation, compensated for some of this. Another weakness is that the same domain expert group has developed and validated the prediction models. However, given the complexity of the prediction models, the variation of the changes applied and variance of the deviation pattern obtained (between the simulations and the thought experiment-based estimates), we cannot see any indication of bias due to the same expert group. Although such threats to validity and reliability are present in such a study, we argue that the results indicate the feasibility and usefulness of the approach in a real-life setting.

7 CONCLUSIONS

The paper presents an approach that makes use of some of the artifacts of PREDIQT and CORAS methods in order to capture the architectural design of a barrier function along with its quality and risk aspects. This is done by modeling the target system and the dependencies w.r.t. risk and quality. The models facilitate a proactive and preventative approach where the dynamic aspects of a system are identified and can be monitored. As a result, the petroleum operator can identify trends and provide early warnings of changing system quality. We have reported on the experiences from using the approach for dynamic monitoring of safety barriers in petroleum installations in an industrial case study.

Apart from PREDIQT and CORAS methods, other modeling approaches based on weighted dependency trees are available and have been evaluated by Omerovic et al. (2011). Moreover, metrics estimation, system quality and the various notations for modeling system architecture, have received much attention in the literature (ISO/IEC 9126, Basili et al. 1994, Fenfor & Neil 1999, Fenton & Pfleeger 1998, Kazman et al. 1998, Mattsson et al. 2006, Neil et al. 2000)

The contributions of this paper include a presentation of the approach and its instantiation, as well as an evaluation of the performance of the approach in an industrial context. The experiences and results obtained indicate that the approach can be carried out with limited resources, in a real-life context and result in useful models that support dynamic monitoring of safety barriers. The study has also provided useful insight into the strengths and weaknesses of the approach, as well as suggested directions for future research. Particularly, the needs for including Quality Model, for a more streamlined process, and for even better tool support, have been highlighted.

Acknowledgments: This work has been conducted as a part of the Dynamic Risk Assistant project funded by the Research Council of Norway, as well as a part of the NESSoS network of excellence funded by the European Commission within the 7th Framework Programme.

REFERENCES

Basili, V. Caldiera, G. and Rombach, H. The Goal Question Metric Approach. Encyclopedia of Software Engineering, 1994.

Fenton, N. &Neil, M. A Critique of Software Defect Prediction Models. IEEE Transactions on Software Engineering, 25:675–689, 1999.

Fenton, N. &Pfleeger, S. Software Metrics: A Rigorous and Practical Approach . PWS Publishing Co., 1998

ISO/IEC 9126 - Software engineering – Product quality. 2004.

Kazman, R., Klein, M., Barbacci, M. Longstaff, T., Lipson, H. and Carriere, J. The Architecture Tradeoff Analysis Method. In Fourth IEEE International Conference on Engineering of Complex Computer Systems, pages 68–78, Aug 1998.

Lund, M. S., Solhaug, B., Stølen, K.: Model-Driven Risk Analsis – The CORAS Approach. Springer, 2011

Mattsson, M., Grahn, H., and Mårtensson, F.. Software Architecture Evaluation Methods for Performance, Maintainability, Testability and Portability. In Second International Conference on the Quality of Software Architectures, 2006.

McGrath, J. E., Groups: Interaction and Performance, Prentice Hall College Div, 1984.

Neil, M., Fenton, N., and Nielsen, L. Building Large-Scale Bayesian Networks. Knowledge Engineering Rev., 15(3):257–284, 2000

Omerovic, A. PREDIQT: A Method for Model-based Prediction of Impacts of Architectural Design Changes on System Quality. PhD thesis, University of Oslo, 2012

Omerovic, A. & Stølen, K. Traceability Handling in Model-based Prediction of System Quality. In Proc. 3rd International Conference on Advances in System Simulation (SIMUL'11), pages 79-88, IEEE Computer Society, 2011a.

Omerovic, A. & Stølen, K. A Practical Approach to Uncertainty Handling and Estimate Acquisition in Model-based Prediction of System Quality. In International Journal on Advances in Systems and Measurements, volume 4, pages 55-70, IARA, 2011b

Omerovic, A., Karahasanovic, A., Stølen, K. Uncertainty Handling in Weighted Dependency Trees – A Systematic Literature Review. In Dependability and Computer Engineering: Concepts for Software-Intensive Systems , pages 381–416. IGI Global, 2011.

Robertson, C. & Krauss, C, (2 August 2010). Gulf Spill Is the Largest of Its Kind, Scientists Say. The New York Times (The New York Times Company). Retrieved April 1. 2013

Solheim, I. & Stølen, K. Technology Research Explained. SINTEF A313, SINTEF ICT, 2007